MATHÉMATIQUES

Analyse et algèbre

Cours de mathématiques de deuxième année avec exercices corrigés et illustrations avec Maple



Stéphane Balac
Laurent Chupin

PRESSES POLYTECHNIQUES ET UNIVERSITAIRES ROMANDES

Trouver la bibliothèque ici : https://1024terabox.com/s/1mg0wxl22G8NjM8MA2RzgFw

Trouver la bibliothèque ici : https://1024terabox.com/s/1mg0wxl22G8NjM8MA2RzgFw

Analyse et algèbre

Cours de mathématiques de deuxième année avec exercices corrigés et illustrations avec Maple

Collection des sciences appliquées de l'INSA de Lyon

Analyse et algèbre

Cours de mathématiques de deuxième année avec exercices corrigés et illustrations avec Maple

Stéphane Balac
Laurent Chupin

PRESSES POLYTECHNIQUES ET UNIVERSITAIRES ROMANDES

Egalement disponible chez le même éditeur

Eléments de mathématiques discrètes
Louis Frécon

Algèbre et Analyse
Cours de mathématiques de première année avec exercices corrigés
Stephane Balac, Frédéric Sturm

Introduction à la statistique
Stephan Morgenthaler

Initiation aux probabilités
Sheldon M. Ross

Algèbre linéaire
Aide-mémoire, exercices et applications
Robert C. Dalang, Amel Chaabouni

Analyse
Recueil d'exercices et aide-mémoire (vol. 1 et 2)
Jacques Douchet

Aide-mémoire d'analyse
Heinrich Matzinger

Les Presses polytechniques et universitaires romandes sont une fondation scientifique dont le but est principalement la diffusion des travaux de l'Ecole polytechnique fédérale de Lausanne et de l'Institut National des Sciences Appliquées de Lyon, ainsi que d'autres universités et écoles d'ingénieurs francophones.
Le catalogue de leurs publications peut être obtenu par courrier aux Presses polytechniques et universitaires romandes, EPFL – Centre Midi, CH-1015 Lausanne, par E-Mail à ppur@epfl.ch, par téléphone au (0)21 693 41 40, ou par fax au (0)21 693 40 27.

http://www.ppur.org

ISBN 978-2-88074-782-4
© Presses polytechniques et universitaires romandes, 2008
CH – 1015 Lausanne.
Tous droits réservés.
Reproduction, même partielle, sous quelque forme
ou sur quelque support que ce soit,
interdite sans l'accord écrit de l'éditeur.

Imprimé en Italie

Préface

La curiosité vous aura ainsi poussé à lire, au-delà de la quatrième de couverture, ces quelques lignes posées en préface de ce livre. Aussi, lecteur ou lectrice inconnu-e, profiterais-je de ce trait de caractère – qui est bien plus un indispensable moteur qu'un vilain défaut – pour vous faire part ici de deux choses sur ce livre.

La première est l'intérêt qu'un étudiant de premier cycle (ou de second cycle désireux de revenir sur quelques notions de base) peut avoir à tenir ce livre sur le coin de sa table.

Stéphane BALAC et Laurent CHUPIN enseignent depuis plusieurs années à l'INSA de Lyon dans la filière ASINSA du premier cycle. Leur modestie dût-elle en souffrir, je dois écrire qu'ils font partie de ces enseignants du supérieur qui, à côté d'une activité reconnue de recherche, maîtrisent l'art difficile de la pédagogie. Ils sont même au nombre de ces enseignants « solides » sur lesquels un responsable peut se reposer du souci d'une équipe ou d'une filière. En un mot, ils connaissent leur métier, de la théorie mathématique à la pratique des étudiants. L'utilisation de cette qualité ne s'est pas limitée à une seule filière du premier cycle et quelques années ont suffi à les doter d'une solide expérience. L'ouvrage que vous tenez dans les mains est quelque part une conséquence heureuse, mais somme toute naturelle, de cette expérience.

Les auteurs se sont ainsi proposés d'écrire un livre qui synthétise cette pratique, et permette à un public plus large que celui – déjà vaste – de l'INSA d'en bénéficier. Il s'agit donc d'une « somme », non au sens d'un exposé qui voudrait plus profiter à la gloire des auteurs qu'à l'usage des étudiants, mais plutôt à celui d'un recueil dont l'exhaustivité autorise plus d'un mode de lecture. En effet, ce livre peut s'utiliser :

- comme manuel, qui accompagnera avec profit un étudiant des premières années d'enseignement supérieur, dans l'apprentissage quotidien par nombre d'exemples et d'illustrations, d'exercices corrigés et de remarques ;
- d'un livre de référence, qui à côté d'une introduction et d'exemples donnant à voir le côté concret d'un thème, offre un niveau de formalisme et une exhaustivité suffisants pour répondre à tous les besoins ;
- d'un livre d'exercices permettant de mieux cerner et perfectionner son niveau de compréhension et de pratique.

La seconde chose, mais vous l'aurez déjà deviné, est qu'à la lecture de cet ouvrage, vous devriez sans doute partager un peu du plaisir et de l'intérêt que j'ai eu à fréquenter ses auteurs. Ce livre leur ressemble, je l'espère, un peu :

sans rien rabattre sur le contenu mathématique, il se situe à hauteur d'étudiant. Vous tenez donc un livre qui peut, sans nul doute, leur être utile.

Bernard ROUX
Directeur du Pôle de Mathématiques
INSA de Lyon

Le contexte dans lequel cet ouvrage a été réalisé est particulier puisqu'il s'agit de la filière de premier cycle internationale ASINSA de l'INSA de Lyon. Elle accueille pour moitié des étudiants français et pour moitié des étudiants originaires d'Asie (Chine, Vietnam, Malaisie, Singapour...). Le fait que les deux auteurs de ce livre enseignent dans une telle filière n'est pas neutre : Stéphane BALAC et Laurent CHUPIN possèdent en effet ce mélange de rigueur et de qualité d'écoute qui font de l'enseignant un « passeur ». Ces qualités ont trouvé dans le contexte multiculturel d'Asinsa un terrain naturel d'expression et leur ouvrage porte l'empreinte de cette richesse : sur le chemin menant aux concepts mathématiques développés dans ce livre, les auteurs ont déposé pour le lecteur des balises extrêmement claires, élaborées à la lumière de cet éclairage multiculturel.

Il est à noter que le présent livre *Analyse et algèbre, cours de mathématiques de deuxième année* fait logiquement suite à l'ouvrage *Algèbre et analyse, cours de mathématiques de première année* publié en 2003 par Stéphane BALAC et Frédéric STURM. L'ensemble formé par ces deux tomes constitue un outil de référence clair et rigoureux, adapté à tout étudiant de premier cycle scientifique.

Jérôme DUPUIS
Directeur de la filière ASINSA
INSA de Lyon

Denis FRIBOULET
Directeur de la filière ASINSA de 2003 à 2007
INSA de Lyon

Avant-propos

Ce *Cours de mathématiques de deuxième année* est issu de l'enseignement dispensé par les auteurs en deuxième année du cycle préparatoire de l'INSA de Lyon dans la filière de premier cycle international ASINSA. Tant dans le fond que dans la forme, il reprend les choix qui avaient été retenus lors de la rédaction du *Cours de mathématiques de première année* publié dans la collection Sciences Appliquées de l'INSA de Lyon aux Presses polytechniques et universitaires romandes[1]. Ce *Cours de mathématiques* dépasse par certains développements le cadre strict du programme habituellement couvert en deuxième année de premier cycle de l'enseignement supérieur. Nous avons souhaité en effet en faire un document de référence que l'élève ingénieur puisse utiliser dans la suite de ses études pour approfondir ou revoir les notions d'algèbre ou d'analyse utilisées dans les enseignements de mathématiques appliquées pour l'ingénieur.

Nous nous sommes attachés dans cet ouvrage à donner des définitions précises et à présenter des raisonnements rigoureux sans toutefois chercher l'exhaustivité. Ainsi, les démonstrations « techniques » sont omises au profit de démonstrations pouvant améliorer la compréhension du résultat énoncé, illustrant l'utilisation de notions déjà introduites ou mettant en avant des idées ou méthodes susceptibles d'être réutilisées par la suite. Celles-ci sont soigneusement détaillées et commentées et une attention toute particulière a été apportée à leur rédaction. Nous ne saurions que trop insister sur l'importance d'une bonne compréhension de ces démonstrations[2], élément fondamental pour l'assimilation des notions mathématiques introduites. Par ailleurs, nous avons dans la mesure du possible cherché à motiver les notions introduites et à les illustrer par des exemples, des remarques et des mises en garde afin de rendre l'apprentissage plus dynamique. Nous avons également cherché à illustrer certaines notions mathématiques introduites dans cet ouvrage en ayant recours au logiciel de calcul formel MAPLE[3]. Chaque chapitre contient de courts exercices visant à tester la bonne compréhension des notions abordées. Il se termine par des exercices de synthèse qui font appel à la fois aux résultats présentés dans le chapitre concerné et aux notions acquises dans les chapitres antérieurs. Ces exercices, souvent issus de devoirs et d'interrogations écrites, doivent permettre d'assimiler des méthodes de raisonnement et des techniques de calcul. Tous les

---

[1] *Algèbre et analyse, cours de mathématiques de première année*, S. BALAC, F. STURM. Collection INSA de Lyon, Presses polytechniques et universitaires romandes, 2003.

[2] Ce qui nécessite une appropriation de la démonstration passant le plus souvent par une reformulation de celle-ci. On peut considérer qu'il s'agit du meilleur moyen de s'assurer de l'assimilation des éléments antérieurs du cours.

[3] MAPLE est un produit de Maplesoft, `www.maplesoft.com`.

exercices sont intégralement corrigés en fin de chaque chapitre. Nous avons tenu à apporter le plus grand soin à la rédaction de ces corrigés en y incluant tous les détails utiles à une bonne assimilation. Enfin, nous avons essayé de fournir quelques éléments biographiques sur les mathématiciens cités dans cet ouvrage à travers formules et théorèmes afin de mieux situer les résultats présentés dans leur contexte historique.

L'enseignement des mathématiques durant les deux années du cycle préparatoire formant un tout cohérent, l'enseignement de seconde année fait largement appel aux notions mathématiques introduites au cours de la première année. Dans cet ouvrage, nous avons mentionné par des notes de bas de page les références aux résultats du cours de première année lorsqu'une connaissance précise de l'énoncé de ceux-ci est indispensable à la compréhension du développement effectué. Les numéros de définition ou théorème mentionnés font référence au *Cours de mathématiques de première année* publié dans la même collection[(2)].

La rédaction de cet ouvrage a été menée à bien au sein du Pôle de mathématiques de l'INSA de Lyon où nous avons pu bénéficier de différentes ressources pédagogiques et matérielles. Il doit aussi sa réalisation aux encouragements de la Direction de la filière internationale ASINSA à mener à bien de tels projets. Nous remercions Messieurs Bernard ROUX, Jérôme DUPUIS et Denis FRIBOULET, qui par ailleurs ont bien voulu préfacer cet ouvrage. Cet ouvrage a bénéficié de la relecture et des commentaires d'Hélène RICARD et de Pascale STÉPHAN, Professeurs agrégés à l'INSA de Lyon, de Frédéric BAYART, Maître de conférences à l'Université Bordeaux 1, de Laurent CARROT, Professeur agrégé en classes préparatoires à Saint-Étienne, de Grégory GINOT, Maître de conférences à l'Université Pierre et Marie Curie, de Sébastien MARTIN, Maître de conférences à l'Université Paris-Sud et de Frédéric STURM, Maître de conférences à l'INSA de Lyon. Nous tenons à leur exprimer ici toute notre gratitude pour cet important travail. Pour l'accueil compréhensif et la confiance qui nous a été accordée, nous souhaitons enfin remercier Bernard BALLAND, Directeur de la Collection des Sciences Appliquées de l'INSA de Lyon, et Olivier BABEL, Directeur des Presses polytechniques et universitaires romandes.

Les auteurs recueilleront avec intérêt toute remarque ou suggestion concernant cet ouvrage.

Stéphane BALAC
ENSSAT
Université de Rennes 1
F-22305 Lannion

Laurent CHUPIN
Pôle de Mathématiques
INSA de Lyon
F-69621 Villeurbanne

Introduction

Ce *Cours de mathématiques de deuxième année* est structuré en quatre parties, la première consacrée aux séries, la deuxième aux espaces vectoriels normés et à l'algèbre bilinéaire, la troisième au calcul différentiel pour les fonctions de plusieurs variables réelles et la dernière à l'intégration dans le plan et dans l'espace. La partie ayant trait aux séries est organisée de la façon suivante. Nous commençons par un chapitre consacré aux séries numériques à termes réels ou complexes. Ce chapitre s'appuie fortement sur le cours de première année consacré aux suites numériques. Les deux chapitres suivants concernent les suites et les séries de fonctions. Nous avons choisi de ne considérer dans un premier temps que les fonctions réelles de la variable réelle, repoussant momentanément l'étude des suites et séries de fonctions complexes de la variable complexe. Ce parti pris est guidé par le souci de ne pas multiplier les notions nouvelles dans un même chapitre. Le chapitre 4 est lui consacré exclusivement aux fonctions complexes de la variable complexe. On y introduit les notions nouvelles de limite, de continuité et de dérivabilité et on y apporte des compléments aux chapitres consacrés aux suites et séries de fonctions. Afin de manipuler les notions liées aux fonctions complexes de la variable complexe, le chapitre 5 est consacré aux séries entières complexes. Les séries entières réelles sont évoquées en insistant sur les propriétés qui diffèrent des séries entières complexes. Le dernier chapitre de cette partie traite des séries de Fourier.

La deuxième partie du cours, constituée des chapitres 7 et 8, est consacrée aux notions d'espace métrique, d'espace vectoriel normé et d'espace euclidien. Nous avons fait le choix d'introduire le concept général de distance et d'espace métrique plutôt que de se restreindre à celui de norme et d'espace vectoriel normé compte tenu de son caractère plus général sans que cela ne se traduise par des difficultés d'assimilation supplémentaires des notions introduites. La première partie du chapitre 7 est consacrée à généraliser dans le cadre d'un espace métrique un certain nombre de notions de topologie énoncées préalablement dans le cas de la droite réelle (cours de première année) et dans le cas du plan complexe (chapitre 4). L'accent est mis sur les notions de topologie directement utiles dans la suite du cours pour introduire de façon claire et précise les théorèmes du calcul différentiel. La seconde partie du chapitre 7 est consacrée à préciser la notion de suite dans un espace métrique en insistant sur la notion de complétude d'un espace métrique et en faisant le lien avec différentes notions topologiques. Le chapitre 8, au contenu assez dense, regroupe des notions relevant de l'algèbre bilinéaire dans le cadre des espaces vectoriels réels. On introduit les notions de forme bilinéaire et de forme quadratique, insistant sur la représentation matricielle d'une forme bilinéaire en dimension finie et mettant en parallèle cette vision matricielle avec celle présentée en première

année à propos des applications linéaires. Cette partie est l'occasion de rappeler et de compléter un certain nombre de connaissances en calcul matriciel concernant en particulier les matrices symétriques et les matrices orthogonales. La notion d'orthogonalité est reprise dans la deuxième partie du chapitre où sont abordées les notions de produit scalaire et d'espace euclidien. La dernière partie du chapitre d'algèbre bilinéaire traite de la réduction des coniques et des quadriques.

La troisième partie du cours est consacrée au calcul différentiel, partant du cadre général de fonctions définies sur un espace métrique et allant vers les techniques spécifiques aux fonctions définies sur $\mathbb{R}^n$. Le chapitre 9 traite des notions de limite et de continuité. Le théorème du point fixe, aux nombreuses applications en mathématiques, est abordé ainsi que les propriétés de continuité des applications linéaires requises au chapitre 10 pour définir la notion de différentielle. Dans le chapitre 10, nous avons fait le choix d'introduire la notion générale d'application différentiable dans le cadre d'un espace vectoriel normé, même si assez rapidement le cas de la dimension finie est abordé et la notion de dérivée partielle introduite. Cette notion d'application différentiable sera reprise dans le dernier chapitre du livre où seront introduites les notions de formes différentielles largement utilisées en physique. Toujours en lien avec les utilisations en physique, sont présentés dans ce chapitre les opérateurs différentiels usuels ainsi que le théorème d'inversion local illustré dans le cas des changements de coordonnées courants. La dernière partie du chapitre est consacrée à la formule de Taylor et à son utilisation pour l'étude des extrema d'une fonction de plusieurs variables réelles.

La dernière partie de l'ouvrage est consacrée au calcul intégral pour les fonctions de plusieurs variables réelles. Le chapitre 11 aborde l'étude des courbes et des surfaces. Il permet de préciser et de relier entre elles les différentes notions de courbes et de surfaces (paramétrée, cartésienne et implicite) et est l'occasion de manipuler les résultats de calcul différentiel introduits dans la partie précédente, en particulier la notion de gradient, le théorème d'inversion locale et le théorème des fonctions implicites. Il est un prérequis au chapitre 13 consacré à l'intégrale curviligne et à l'intégrale surfacique. Nous introduisons au chapitre 12 la notion d'intégrale multiple en reprenant, pour une meilleure compréhension, l'approche qui avait été utilisée dans le *Cours de première année* pour introduire l'intégrale de Riemann d'une fonction de la variable réelle. L'intégration sur des courbes et des surfaces est l'objet du dernier chapitre. Ce chapitre est également l'occasion de préciser certaines notions liées aux formes différentielles permettant d'énoncer le théorème très général de Stokes dont certains corollaires sont les formules de Green et d'Ostrogradski très utilisées en physique.

Table des matières

Préface v

Avant-propos vii

Introduction ix

Table des matières xi

Table des notations xxi

1 Séries numériques 1

1.1 Définitions et premières propriétés 1

1.1.1 Définitions 1

1.1.2 Critère de Cauchy 9

1.1.3 Premières propriétés 10

1.2 Séries réelles à termes positifs 13

1.2.1 Propriétés générales 13

1.2.2 Critères de comparaison 14

1.2.3 Séries de référence 23

1.3 Séries absolument convergentes 28

1.3.1 Généralités 28

1.3.2 Règles de Cauchy et de D'Alembert 30

1.3.3 Produit de Cauchy de deux séries 37

1.4 Séries semi-convergentes 40

1.4.1 Définition 40

1.4.2 Séries alternées 41

1.4.3 Règle d'Abel 46

1.4.4 Séries trigonométriques 47

1.4.5 Séries ne relevant pas de la règle d'Abel 48

1.5 Groupement de termes et changement de l'ordre des termes . . 49

1.6 Exercices de synthèse 53

1.7 Solutions des exercices 55

2 Suites de fonctions **67**

2.1 L'espace vectoriel des applications bornées 67
2.2 Convergence simple et convergence uniforme 70
2.2.1 Généralités 70
2.2.2 Convergence simple 70
2.2.3 Convergence uniforme 73
2.2.4 Propriétés 78
2.2.5 Critère de Cauchy pour la convergence uniforme 82
2.3 Propriétés de la limite d'une suite d'applications réelles 84
2.3.1 Limite simple d'une suite d'applications 84
2.3.2 Limite d'une suite d'applications continues 86
2.3.3 Limite d'une suite d'applications intégrables 89
2.3.4 Limite d'une suite d'applications dérivables 92
2.4 Plan d'étude de la convergence d'une suite d'applications . . . 94
2.5 Exercices de synthèse 95
2.6 Solutions des exercices 97

3 Séries de fonctions **113**

3.1 Convergence d'une série de fonctions 113
3.1.1 Définitions 113
3.1.2 Propriétés 122
3.2 Critères de convergence uniforme 125
3.2.1 Convergence normale 125
3.2.2 Convergence uniforme des séries alternées 130
3.2.3 Règle d'Abel pour la convergence uniforme 133
3.3 Propriétés de la somme d'une série de fonctions 134
3.4 Exercices de synthèse 144
3.5 Solutions des exercices 147

4 Fonctions complexes de la variable complexe **159**

4.1 Fonctions complexes d'une variable réelle 160
4.1.1 Limite et continuité 160
4.1.2 Dérivabilité 161
4.1.3 Intégrabilité 162
4.1.4 Applications bornées 162
4.1.5 Inégalité des accroissements finis 163

4.2 Topologie du plan complexe 164
4.2.1 Ouverts et fermés de $\mathbb{C}$ 164
4.2.2 Intérieur et adhérence d'un ensemble 168
4.2.3 Arcs géométriques dans le plan complexe 170
4.3 Fonctions complexes d'une variable complexe 173
4.3.1 Généralités 173
4.3.2 Quelques exemples 175
4.3.3 Représentations graphiques 176
4.4 Limite d'une fonction de la variable complexe 178
4.4.1 Définition de la limite 178
4.4.2 Étude de la limite 181
4.4.3 Propriétés de la limite 184
4.5 Continuité d'une fonction de la variable complexe 189
4.5.1 Définitions et premières propriétés 189
4.5.2 Opérations algébriques sur les applications continues . . 191
4.6 Dérivabilité d'une fonction de la variable complexe 193
4.6.1 Fonctions holomorphes et fonctions différentiables . . . 193
4.6.2 Propriétés des fonctions holomorphes 198
4.6.3 Dérivées successives 202
4.7 Fonctions usuelles de la variable complexe 204
4.7.1 La fonction exponentielle 204
4.7.2 Fonctions trigonométriques circulaires et hyperboliques 208
4.7.3 La fonction logarithme 213
4.8 Suites et séries de fonctions de la variable complexe 216
4.8.1 Suites de fonctions de la variable complexe 216
4.8.2 Séries de fonctions de la variable complexe 217
4.8.3 Continuité des fonctions usuelles 218
4.9 Exercices de synthèse 220
4.10 Solutions des exercices 222

5 Séries entières **231**
5.1 Série entière de la variable complexe 231
5.1.1 Généralités sur les séries entières 231
5.1.2 Rayon de convergence d'une série entière 232
5.1.3 Détermination du rayon de convergence 237
5.1.4 Convergence uniforme d'une série entière 246

5.2 Fonctions définies par une série entière 248
5.2.1 Continuité de la fonction somme 248
5.2.2 Dérivabilité de la fonction somme 249
5.2.3 Fonction développable en série entière 253
5.3 Série entière de la variable réelle 260
5.3.1 Propriétés des séries entières réelles 260
5.3.2 Développement en série entière 264
5.3.3 Quelques utilisations des séries entières 267
5.3.4 Développement en série entière des fonctions usuelles . . 270
5.4 Séries entières et équations différentielles 271
5.4.1 Résolution d'une équation différentielle à l'aide de séries 271
5.4.2 La méthode de Frobenius 275
5.4.3 Équations différentielles et développement en série . . . 279
5.5 Exercices de synthèse 280
5.6 Solutions des exercices 284

6 Séries de Fourier **301**
6.1 Séries trigonométriques 301
6.1.1 Convergence d'une série trigonométrique 301
6.1.2 Propriétés de la somme d'une série trigonométrique . . 302
6.2 Décomposition en série de Fourier 305
6.2.1 Coefficients de Fourier 305
6.2.2 Propriétés des coefficients de Fourier 314
6.2.3 Convergence d'une série de Fourier 320
6.2.4 Quelques exemples 325
6.3 Exercices de synthèse 329
6.4 Solutions des exercices 332

7 Espaces métriques et espaces vectoriels normés **345**
7.1 Distance et espace métrique 345
7.2 Normes et espaces vectoriels normés 356
7.3 Topologie dans un espace métrique 369
7.3.1 Ensembles ouverts et ensembles fermés 369
7.3.2 Intérieur et adhérence d'un ensemble 377
7.3.3 Ensembles bornés 382
7.3.4 Espaces connexes 384

7.3.5 Distance à un ensemble 386
7.4 Suites dans un espace métrique 388
7.4.1 Définitions et propriétés 388
7.4.2 Suites extraites 395
7.4.3 Suites de Cauchy 398
7.4.4 Caractérisation des fermés dans un espace métrique . . 410
7.4.5 Ensembles compacts 412
7.5 Exercices de synthèse 416
7.6 Solutions des exercices 418

8 Algèbre bilinéaire **437**
8.1 Formes bilinéaires et formes quadratiques 437
8.1.1 Formes bilinéaires 437
8.1.2 Cas d'un espace vectoriel en dimension finie 440
8.1.3 Changement de base 442
8.1.4 Formes bilinéaires symétriques 446
8.1.5 Formes quadratiques 448
8.1.6 Orthogonalité 454
8.1.7 Bases orthogonales 458
8.1.8 Décomposition en somme de carrés de formes linéaires . 463
8.1.9 Classification des formes quadratiques 471
8.2 Produit scalaire sur un espace vectoriel 476
8.2.1 Produit scalaire 476
8.2.2 Espace euclidien 484
8.2.3 Procédé d'orthogonalisation de Schmidt 486
8.3 Isométries et matrices orthogonales 490
8.3.1 Isométries 490
8.3.2 Matrices orthogonales 492
8.3.3 Description géométrique des isométries 498
8.4 Diagonalisation des matrices symétriques réelles 503
8.5 Réduction des coniques et quadriques 506
8.5.1 Étude des coniques 506
8.5.2 Classification des coniques 509
8.5.3 Description des différents genres de coniques 510
8.5.4 Étude des quadriques 514
8.5.5 Classification des quadriques 515
8.6 Exercices de synthèse 520
8.7 Solutions des exercices 523

9 Limite et continuité **535**

9.1 Limite et continuité dans un espace métrique 535
9.1.1 Limite d'une fonction 535
9.1.2 Continuité d'une fonction 546
9.1.3 Continuité uniforme 556
9.1.4 Applications lipschitziennes 558
9.1.5 Applications continues et topologie 561
9.2 Limite et continuité des fonctions de $\mathbb{R}^n$ dans $\mathbb{R}^p$ 565
9.3 Continuité des applications linéaires 572
9.4 Le théorème du point fixe 584
9.5 Exercices de synthèse 592
9.6 Solutions des exercices 594

10 Calcul différentiel **609**

10.1 Différentiabilité dans un espace vectoriel 609
10.1.1 Applications différentiables 609
10.1.2 Opérations sur les applications différentiables 618
10.1.3 Utilisation des différentielles en physique 624
10.2 Différentiabilité en dimension finie 627
10.2.1 Dérivation selon un vecteur 627
10.2.2 Dérivées partielles 630
10.2.3 Matrice jacobienne 635
10.2.4 Cas des applications à valeurs dans $\mathbb{R}$ 641
10.3 Applications de classe $\mathcal{C}^q$ 645
10.3.1 Notion de dérivées d'ordre supérieur 645
10.3.2 Dérivations partielles d'ordre supérieur 647
10.3.3 Propriétés des applications de classe $\mathcal{C}^q$ 651
10.3.4 Cas des applications à valeurs dans $\mathbb{R}$ 653
10.3.5 Opérateurs différentiels 654
10.3.6 Théorèmes d'inversion locale et globale 663
10.3.7 Changements de coordonnées 666
10.4 Inégalité des accroissements finis ; formules de Taylor 677
10.4.1 Inégalité des accroissements finis 677
10.4.2 Formule de Taylor 683
10.5 Extrema d'une fonction réelle de plusieurs variables 687
10.5.1 Conditions nécessaires d'existence d'un extremum . . . 688

10.5.2 Conditions suffisantes d'existence d'un extremum 692
10.6 Exercices de synthèse 697
10.7 Solutions des exercices 699

11 Courbes et surfaces **713**
11.1 Trois définitions de courbes et de surfaces 713
11.1.1 Le règne des paramètres 713
11.1.2 Ensembles implicites 719
11.1.3 Courbes et surfaces cartésiennes 721
11.1.4 Premiers liens entre les différentes représentations . . . 722
11.1.5 Quelques courbes et surfaces de référence 724
11.2 Surfaces cartésiennes dans l'espace 727
11.2.1 Courbes de niveau et lignes coordonnées 728
11.2.2 Plan tangent à une surface cartésienne 730
11.2.3 Position d'une surface par rapport à son plan tangent . 732
11.3 Étude des arcs paramétrés 735
11.3.1 Classification des arcs 735
11.3.2 Propriétés des arcs paramétrés 743
11.3.3 Étude des arcs paramétrés dans le plan 750
11.3.4 Étude des arcs paramétrés dans l'espace 772
11.4 Étude des nappes paramétrées 788
11.4.1 Classification des nappes paramétrées 789
11.4.2 Régularité des nappes paramétrées 791
11.4.3 Étude des surfaces paramétrées dans l'espace 797
11.4.4 Arcs tracés sur une nappe 801
11.5 Théorème d'inversion locale en géométrie 804
11.5.1 Application à l'étude des courbes paramétrées 804
11.5.2 Application à l'étude des surfaces paramétrées 807
11.6 Théorème des fonctions implicites 809
11.6.1 Cas des fonctions de $\mathbb{R}^2$ dans $\mathbb{R}$ 809
11.6.2 Étude des courbes implicites du plan 814
11.6.3 Cas des fonctions de $\mathbb{R}^{n+q}$ dans $\mathbb{R}^q$ 819
11.6.4 Étude des surfaces implicites dans l'espace 826
11.6.5 Étude des courbes implicites dans l'espace 829
11.7 Exercices de synthèse 832
11.8 Solutions des exercices 836

12 Intégrale multiple **855**
12.1 Définition et premières propriétés 855
12.1.1 Fonctions en escalier sur un pavé de $\mathbb{R}^n$ 855
12.1.2 Intégrale d'une fonction en escalier 858
12.1.3 Fonctions Riemann-intégrables sur un pavé 861
12.1.4 Fonctions Riemann-intégrables sur un ensemble borné . 865
12.1.5 Conditions d'intégrabilité 868
12.1.6 Propriétés des fonctions Riemann-intégrables 873
12.2 Calcul des intégrales multiples 875
12.2.1 Le théorème de Fubini 875
12.2.2 Intégration par pile et par tranche 881
12.3 Changement de variables pour une intégrale multiple 893
12.3.1 Changement de variables dans $\mathbb{R}^n$ 893
12.3.2 Changement de variables polaire 900
12.3.3 Changement de variables cylindrique 905
12.3.4 Changement de variables sphérique 908
12.3.5 Exploitation de symétries 911
12.4 Intégrale multiple généralisée 915
12.5 Exercices de synthèse 921
12.6 Solutions des exercices 923

13 Intégration sur des courbes et des surfaces **937**
13.1 Intégrale curviligne 937
13.1.1 Intégrer le long d'une courbe 938
13.1.2 Longueur d'une courbe 945
13.2 Circulation et formes différentielles 956
13.2.1 Formes différentielles de degré 1 956
13.2.2 Utilisation du formalisme des 1-formes en physique . . . 966
13.2.3 Circulation d'une 1-forme 968
13.3 Intégrale de surface 973
13.3.1 Intégrer sur une surface 973
13.3.2 Aire d'une surface 979
13.4 Formule de Stokes 985
13.4.1 Généralités sur les formes de degré p 985
13.4.2 Intégrale d'une 2-forme 992
13.4.3 Intégrale d'une p-forme 994

13.4.4 Différentielle extérieure 996
13.4.5 Théorème de Stokes 998
13.5 Exercices de synthèse 1009
13.6 Solutions des exercices 1013

Bibliographie **1027**

Index **1029**

Table des notations

$\|\cdot\|$	norme sur un espace vectoriel
$\langle\cdot,\cdot\rangle$	produit scalaire sur un espace vectoriel
$\mathbf{u}\wedge\mathbf{v}$	produit vectoriel des deux vecteurs $(\mathbf{u},\mathbf{v})\in(\mathbb{K}^3)^2$
$\omega\wedge\widetilde{\omega}$	produit extérieur des deux formes différentielles ω et $\widetilde{\omega}$
${}^t\mathrm{A}$	transposée de la matrice $\mathrm{A}\in\mathcal{M}_{m,n}(\mathbb{K})$
A^*	adjointe de la matrice $\mathrm{A}\in\mathcal{M}_{m,n}(\mathbb{K})$
$A^\perp$	orthogonal de l'ensemble A
$\mathcal{A}(U,V)$	ensemble des applications dont l'ensemble de définition est U et l'image est (un sous-ensemble de) V
$\mathrm{Arg}(z)$	argument principal du complexe z ; $\mathrm{Arg}(z)\in]-\pi,\pi]$
$\mathcal{B}(U,\mathbb{K})$	ensemble des applications bornées sur U à valeurs dans $\mathbb{K}$
$B(0,R)$	dans un espace vectoriel normé $(E,\|\cdot\|)$, boule ouverte centrée à l'origine et de rayon R : $\{\mathbf{x}\in E\ ;\ \|\mathbf{x}\|<R\}$
$\overline{B}(0,R)$	dans un espace vectoriel normé $(E,\|\cdot\|)$, boule fermée centrée à l'origine et de rayon R : $\{\mathbf{x}\in E\ ;\ \|\mathbf{x}\|\leqslant R\}$
$\mathcal{C}_n^k$	coefficients du binôme de Newton : $\mathcal{C}_n^k=\frac{n!}{k!\,(n-k)!}$
$\mathcal{C}^0(U,\mathbb{K})$	ensemble des applications de U dans K continues sur U
$\mathcal{C}^k(U,\mathbb{K})$	ensemble des applications de U dans $\mathbb{K}$ possédant une dérivée k–ième continue sur U
$\mathcal{C}^\infty(U,\mathbb{K})$	ensemble des applications de U dans $\mathbb{K}$ indéfiniment dérivables sur U
$C(q)$	cône isotrope de la forme quadratique q
$\complement_E X$	complémentaire de l'ensemble X dans E
$\mathrm{d}f$	différentielle de l'application f
$\mathrm{d}^k f$	différentielle k-ième de l'application f

$\mathrm{d}_{\mathbf{v}} f_{\mathbf{a}}$	dérivée de l'application f en $\mathbf{a}$ dans la direction du vecteur $\mathbf{v}$
$\partial_x f, \partial_y f, \partial_z f$ $\partial_1 f, \partial_2 f, \partial_3 f$	dérivées partielles premières de l'application f aussi notées $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$
Δf	laplacien de l'application f
δ_{ij}	symbole de Kronecker ; vaut 1 si $i = j$ et 0 sinon
$D(0, R)$	disque ouvert centré à l'origine et de rayon R dans le plan complexe : $\{z \in \mathbb{C} \ ; \ \lvert z\rvert < R\}$
$\overline{D}(0, R)$	disque fermé centré à l'origine et de rayon R dans le plan complexe : $\{z \in \mathbb{C} \ ; \ \lvert z\rvert \leqslant R\}$
$\det(\mathrm{A})$	déterminant de la matrice A
div $\mathbf{f}$	divergence de l'application vectorielle $\mathbf{f}$
$\mathrm{E}(x)$	partie entière du réel x
E^*	dual algébrique de l'espace vectoriel E : $E^* = \mathcal{L}(E, \mathbb{K})$
$\varepsilon(q)$	signature de la forme quadratique q
$f^{(k)}$	dérivée k–ième de l'application f
f^{-1}	bijection réciproque de l'application bijective f
$\mathrm{GL}_n(\mathbb{K})$	ensemble des matrices inversibles de $\mathcal{M}_n(\mathbb{K})$
H_f	matrice hessienne de l'application f
$\mathrm{i_d}$	application identité $x \mapsto x$
I_n ou Id	matrice identité dans $\mathcal{M}_n(\mathbb{K})$
$\mathrm{Im}(f)$	image de l'application linéaire f
J_f	matrice jacobienne de l'application f
$\mathcal{J}_f$	jacobien de l'application f : $\mathcal{J}_f = \det(\mathrm{J}_f)$
$\mathrm{Ker}(f)$	noyau de l'application linéaire f
$\mathbb{K}$	indifféremment le corps $\mathbb{R}$ ou le corps $\mathbb{C}$
$\mathbb{K}^n$	produit carthésien $\mathbb{K} \times \cdots \times \mathbb{K}$ de $\mathbb{K}$ par lui-même n fois
$\mathbb{K}[X]$	ensemble des polynômes à coefficients dans $\mathbb{K}$
$\mathbb{K}(X)$	ensemble des fractions rationnelles à coefficients dans $\mathbb{K}$

$\mathbb{K}_n[X]$	ensemble des polynômes de $\mathbb{K}[X]$ de degré au plus n
$\mathcal{S}(\mathbb{K})$	ensemble des suites numériques à coefficients dans $\mathbb{K}$
$\ell^1(\mathbb{K})$	ensemble des séries numériques à coefficients dans $\mathbb{K}$ absolument convergentes
$\mathcal{L}(E,F)$	ensemble des applications linéaires de E dans F
$\mathcal{L}_c(E,F)$	ensemble des applications linéaires continues de E dans F
$\mathcal{L}_2(E\times E,\mathbb{R})$	ensemble des formes bilinéaires sur E
$\Lambda^p(U)$	ensemble des p-formes différentielles sur U
$\mathcal{M}_{m,n}(\mathbb{K})$	espace vectoriel des matrices rectangulaires à m lignes et n colonnes dont les éléments appartiennent à $\mathbb{K}$
$\mathcal{M}_n(\mathbb{K})$	ensemble des matrices carrées $n\times n$ dont les éléments appartiennent à $\mathbb{K}$
$m(P)$	mesure du pavé P de $\mathbb{R}^n$
$N(q)$	noyau de la forme quadratique q
∇f	gradient de l'application f
$(O;\mathbf{i},\mathbf{j},\mathbf{k})$	repère canonique de l'espace
$\mathbb{R}\setminus 2\pi\mathbb{Z}$	$\{x\in\mathbb{R}\ ;\ \forall k\in\mathbb{Z}\quad x\neq 2k\pi\}$
$\mathbb{R}^+$	ensemble des réels positifs
$\mathbb{R}^*$	ensemble des réels non nuls
$\mathbb{R}_+^*$	ensemble des réels strictement positifs
$\mathbb{R}_*^n$	$=\mathbb{R}^n\setminus\{(0,\dots,0)\}$
$\mathrm{rg}(q)$	rang de la forme quadratique q
$\mathrm{rot}\,\mathbf{f}$	rotationnel de l'application vectorielle $\mathbf{f}$
$\Sigma(\mathbb{K})$	ensemble des séries numériques à coefficients dans $\mathbb{K}$
$\mathcal{S}(U,\mathbb{K})$	ensemble des suites de fonctions définies sur $U\subset\mathbb{K}$ à valeurs dans $\mathbb{K}$
$\Sigma(U,\mathbb{K})$	ensemble des séries de fonctions définies sur $U\subset\mathbb{K}$ à valeurs dans $\mathbb{K}$
$\mathcal{S}_F(f)$	série de Fourier de l'application f

$\mathcal{S}_n(\mathbb{R})$	ensemble des matrices symétriques de $\mathcal{M}_n(\mathbb{R})$
$\mathrm{Tr(A)}$	trace de la matrice $\mathrm{A} \in \mathcal{M}_n(\mathbb{K})$
$(u_n)_n$	suite de terme général u_n
$\sum_n u_n$	série de terme général u_n
$\overline{X}$	adhérence de l'ensemble X
$\mathrm{Int}\,(X)$ ou $\mathring{X}$	intérieur de l'ensemble X
$\overline{z}$	complexe conjugué de z dans $\mathbb{C}$

CHAPITRE 1

Séries numériques

Les séries numériques, considérées en tant que limites de progressions géométriques, sont apparues très tôt, dès Aristote, dans l'histoire des mathématiques mais au 18–ième siècle la manipulation des séries numériques, en particulier leurs propriétés de convergence, faisait encore l'objet de nombreux débats et travaux. L'un des célèbres paradoxes de Zénon d'Élée, philosophe grec du 5^e siècle avant J.-C., connu sous le nom d'*Achille et la tortue*, a pu être résolu en ayant recours aux séries numériques par le mathématicien belge Grégoire de Saint-Vincent au 17^e siècle. Le paradoxe d'*Achille et la tortue* est le suivant : un jour, le héros grec Achille dispute une course à pied avec une tortue à laquelle il laisse un certaine avance. Zénon d'Élée met en avant le fait que le temps qu'Achille parcourt la distance qu'il a laissé d'avance à la tortue, celle-ci s'est déplacée. Le temps qu'Achille parcourt cette nouvelle distance qui le sépare de la tortue, celle-ci a encore un petit peu avancé, Achille doit donc encore couvrir cette distance avant de rattraper la tortue. Mais une fois encore, le temps qu'il parcourt cette distance, la tortue a à nouveau avancé. En définitive, Zénon affirme qu'Achille ne pourra jamais rattraper la tortue... Ce paradoxe est lié au fait qu'intuitivement on peut penser qu'en sommant un nombre infini de termes positifs, ici le temps mis par Achille pour parcourir à chaque étape la distance le séparant de la tortue, on obtient nécessairement un nombre infini. Ce qui n'est pas le cas.

Nous allons dans ce chapitre introduire un certain nombre de notions liées aux séries numériques, en insistant sur les critères permettant d'établir qu'une série possède une somme finie et sur les méthodes permettant de calculer cette somme.

1.1 Définitions et premières propriétés

1.1.1 Définitions

On désigne par $\mathbb{K}$ le corps $\mathbb{R}$ ou le corps $\mathbb{C}$ et on note $\mathcal{S}(\mathbb{K})$ l'espace vectoriel sur $\mathbb{K}$ des suites numériques[1] à valeurs dans $\mathbb{K}$ (suites réelles si $\mathbb{K} = \mathbb{R}$ ou suites complexes si $\mathbb{K} = \mathbb{C}$).

[1] Voir le chap. 5 du *Cours de première année*.

Si l'on considère une suite numérique $(u_n)_n$, il est assez naturel, à l'instar de ce qui est fait avec une suite géométrique, de sommer les n premiers termes de cette suite et d'étudier la nouvelle suite $(U_n)_n$ ainsi construite qui est définie par

$$\forall n \in \mathbb{N} \qquad U_n = \sum_{k=0}^{n} u_k. \tag{1}$$

qui est appelée *suite des sommes partielles associée à la suite* $(u_n)_n$.

DÉFINITION 1.1 *Soit $(u_n)_n$ une suite numérique; on appelle* série numérique de terme général u_n *et l'on note $\sum_n u_n$, le couple $((u_n)_n, (U_n)_n)$ d'éléments de $\mathcal{S}(\mathbb{K}) \times \mathcal{S}(\mathbb{K})$ où $(U_n)_n$ est la suite des sommes partielles associée à la suite $(u_n)_n$ définie par la relation (1).*

Remarques

1. À une suite numérique $(u_n)_n$ on peut associer sa suite des sommes partielles $(U_n)_n$ et par conséquent la série $\sum_n u_n$. Inversement, à une suite numérique $(S_n)_n$ on peut associer une série $\sum_n u_n$ dont elle est la suite des sommes partielles de la manière suivante :

$$u_0 = S_0 \quad \text{et} \quad \forall n \in \mathbb{N}^* \qquad u_n = S_n - S_{n-1}.$$

2. Deux séries numériques $\sum_n u_n$ et $\sum_n v_n$ sont égales si pour tout $n \in \mathbb{N}$ on a $u_n = v_n$. ◇

Structure d'espace vectoriel

On définit sur l'ensemble $\Sigma(\mathbb{K})$ des séries numériques sur $\mathbb{K}$ une loi de composition interne $+$ de la manière suivante : si $\sum_n u_n$ et $\sum_n v_n$ sont deux séries numériques alors $\sum_n u_n + \sum_n v_n$ est la série numérique de terme général $u_n + v_n$. On peut également munir $\Sigma(\mathbb{K})$ d'une loi de composition externe $\cdot$ de la manière suivante : pour toute série numérique $\sum_n u_n$ et pour tout scalaire $\lambda \in \mathbb{K}$ la série $\lambda \cdot \sum_n u_n$ est la série de terme général λu_n.

L'ensemble des séries numériques $\Sigma(\mathbb{K})$ muni de ces deux lois $+$ et $\cdot$ est un espace vectoriel[(2)] sur $\mathbb{K}$.

DÉFINITION 1.2 (Nature d'une série)

✗ *On dit que la série* $\sum_n u_n$ converge dans $\mathbb{K}$ *si la suite des sommes partielles $(U_n)_n$ converge dans $\mathbb{K}$.*

✗ *Si $\mu \in \mathbb{K}$ désigne la limite de la suite $(U_n)_n$, on dit que μ est la* somme de la série $\sum_n u_n$ *et on note $\mu = \sum_{n=0}^{+\infty} u_n$.*

✗ *Si la suite des sommes partielles $(U_n)_n$ diverge dans $\mathbb{K}$, on dit que la série* $\sum_n u_n$ diverge.

[(2)] Voir le chapitre 8 du *Cours de première année*.

Remarque Si la suite $(u_n)_n$ n'est définie qu'à partir d'un rang $n_0 \in \mathbb{N}^*$, on peut prolonger cette suite en une suite $(\tilde{u}_n)_n$ définie sur $\mathbb{N}$ en posant

- $\tilde{u}_n = u_n$ pour tout $n \in \mathbb{N}$ avec $n \geqslant n_0$;
- $\tilde{u}_n = 0$ pour tout $n \in \mathbb{N}$ avec $n < n_0$.

La série $\sum_n \tilde{u}_n$ sera notée $\sum_{n \geqslant n_0} u_n$.

Inversement, à la série $\sum_n u_n$ et à l'entier non nul n_0, on peut associer la série $\sum_{n \geqslant n_0} u_n$ qui est qualifiée de série déduite de $\sum_n u_n$ par troncature.

Dans ce chapitre et dans les chapitres suivants, les propriétés concernant les séries seront énoncées en supposant que $n_0 = 0$ mais on adaptera sans difficulté l'énoncé de ces propriétés dans le cas où $n_0 \in \mathbb{N}^*$. ◇

Exercice 1.1 ***Procédé télescopique***
Soient $(u_n)_n$ une suite numérique et $\sum_n v_n$ la série de terme général $v_n = u_{n+1} - u_n$. Montrer que la série $\sum_n v_n$ converge si et seulement si la suite $(u_n)_n$ converge. Quelle relation a-t-on entre la limite de la suite $(u_n)_n$ et la somme de la série $\sum_n v_n$?

Les exemples qui suivent vont fournir les principales techniques d'étude d'une série numérique (sa nature et le calcul de sa somme) basées sur l'étude de la suite des sommes partielles. Ces exemples nous fourniront également des séries de référence pour illustrer les propriétés des séries numériques qui sont l'objet de ce chapitre.

Exemples

1. La série géométrique de raison r ($r \in \mathbb{K}$) et de premier terme α ($\alpha \in \mathbb{K}$) est la série de terme général $u_n = \alpha r^n$. Pour tout $n \in \mathbb{N}$, on a[3]

$$U_n = \sum_{k=0}^{n} u_k = \begin{cases} \alpha \dfrac{1 - r^{n+1}}{1 - r} & \text{si } r \neq 1 \\ (n+1)\,\alpha & \text{si } r = 1 \end{cases}.$$

Si $|r| < 1$, la suite $(U_n)_n$ converge vers $\alpha(1-r)^{-1}$ et si $|r| \geqslant 1$ la suite $(U_n)_n$ diverge. On a donc la propriété suivante.

La série géométrique de raison r et de premier terme α converge si et seulement si $|r| < 1$; elle a pour somme $\displaystyle\sum_{n=0}^{+\infty} \alpha r^n = \alpha(1-r)^{-1}$.

Revenons sur le paradoxe *d'Achille et la tortue* ; supposons qu'Achille court à la vitesse de 10 m/s, que la tortue se déplace à 0,1 m/s et qu'Achille ait laissé 100 m d'avance à la tortue. Achille met 10 s à parcourir les 100 m que la tortue avait d'avance ; la tortue a parcouru pendant ce temps 1 m qu'Achille

[3] Voir le *Cours de première année*, proposition 5.21 p. 195.

met 10^{-1} s à couvrir. La tortue a pendant ce temps avancé de 10^{-2} m et Achille met 10^{-3} s à parcourir cette distance. Le temps total que court Achille dans sa poursuite de la tortue est, en seconde,

$$T = 10 + 10^{-1} + 10^{-2} + \cdots = \sum_{n=0}^{+\infty} 100\ 10^{-k} = \frac{10}{1-10^{-2}} = \frac{1000}{99} = 10,1010\ldots$$

Achille met donc un peu moins de 11 s pour rattraper la tortue. La tortue aura au total parcouru une distance, en mètre, de

$$D = 1 + 10^{-2} + 10^{-4} + \cdots = \sum_{n=0}^{+\infty} 10^{k} = \frac{1}{1-10^{-2}} = \frac{100}{99} = 1,1010\ldots$$

avant qu'Achille ne la rejoigne.

2. Montrons que la série $\sum_{n\geqslant 2} u_n$ de terme général $u_n = \dfrac{1}{n(n-1)}$ converge et que sa somme vaut 1. On a la décomposition en éléments simples[4] suivante dans $\mathbb{R}(X)$

$$\frac{1}{X(X-1)} = \frac{1}{X-1} - \frac{1}{X}.$$

qui permet d'écrire le terme général de la suite des sommes partielles sous la forme

$$U_n = \sum_{k=2}^{n} u_k = \sum_{k=2}^{n} \left(\frac{1}{k-1} - \frac{1}{k}\right) = \sum_{k=1}^{n-1} \frac{1}{k} - \sum_{k=2}^{n} \frac{1}{k} = 1 - \frac{1}{n}.$$

La suite des sommes partielles converge vers 1 et par conséquent la série $\sum_{n\geqslant 2} u_n$ est convergente et a pour somme $\sum_{n=2}^{+\infty} u_n = 1$.

3. La série de terme général $u_n = (-1)^n \frac{1}{n}$ $(n \in \mathbb{N}^*)$ est appelée *série harmonique alternée.* Elle converge et sa somme vaut $-\ln(2)$. Pour tout $k \in \mathbb{N}^*$ on a en effet $\frac{1}{k} = \int_0^1 t^{k-1}\,\mathrm{d}t$, de sorte que le terme général de la suite des sommes partielles associée à la série harmonique alternée s'écrit :

$$U_n = \sum_{k=1}^{n} \frac{(-1)^k}{k} = -\sum_{k=1}^{n} \int_0^1 (-t)^{k-1}\,\mathrm{d}t = -\int_0^1 \sum_{k=1}^{n} (-t)^{k-1}\,\mathrm{d}t.$$

Or, $\displaystyle\sum_{k=1}^{n} (-t)^{k-1} = \frac{1-(-t)^n}{1+t}$ car il s'agit de la somme des n premiers termes d'une progression géométrique[4] de raison $-t$. On en déduit que :

$$\begin{aligned} U_n &= -\int_0^1 \frac{1-(-t)^n}{1+t}\,\mathrm{d}t = -\int_0^1 \frac{1}{1+t}\,\mathrm{d}t + \int_0^1 \frac{(-t)^n}{1+t}\,\mathrm{d}t \\ &= -\ln(2) + \int_0^1 \frac{(-t)^n}{1+t}\,\mathrm{d}t. \end{aligned}$$

[4] Voir le chap. 7 du *Cours de première année.*

Finalement, on obtient pour tout $n \in \mathbb{N}^*$

$$|U_n + \ln(2)| = \left| \int_0^1 \frac{(-t)^n}{1+t}\,\mathrm{d}t \right| \leqslant \int_0^1 \frac{t^n}{1+t}\,\mathrm{d}t \leqslant \int_0^1 t^n\,\mathrm{d}t = \frac{1}{n+1}.$$

On conclut alors par le théorème d'encadrement[5] que la suite des sommes partielles converge vers $-\ln(2)$.

La série harmonique alternée $\sum_n (-1)^n \frac{1}{n}$ converge et a pour somme $-\ln(2)$.

On pourra consulter l'exercice 1.2 p. 7 pour une autre méthode permettant d'établir ce résultat.

4. Montrons que la série de terme général $u_n = 1/n!$ converge et que sa somme vaut e. La formule de Taylor-Lagrange[6] appliquée à l'ordre n à la fonction exponentielle entre 0 et 1 donne :

$$\exists c_n \in]0,1[\qquad \mathrm{e} = \sum_{k=0}^{n} \frac{1}{k!} + \frac{1}{(n+1)!}\,\exp(c_n).$$

Le terme général de la suite $(U_n)_n$ des sommes partielles associée à cette série vérifie donc :

$$U_n = \sum_{k=0}^{n} \frac{1}{k!} = \mathrm{e} - \frac{1}{(n+1)!}\,\exp(c_n).$$

Puisque $c_n \in]0,1[$ on a $1 < \exp(c_n) < \mathrm{e}$ et

$$|U_n - \mathrm{e}| = \frac{1}{(n+1)!}\,\exp(c_n) < \frac{\mathrm{e}}{(n+1)!}.$$

Par le théorème d'encadrement[5], on en déduit que la suite des sommes partielles converge vers e et par conséquent que la série $\sum_n u_n$ est convergente, de somme $\sum_{n=1}^{+\infty} u_n = \mathrm{e}$. Plus généralement, en utilisant la même méthode, on montre le résultat suivant.

Pour tout réel x la série de terme général $u_n = x^n/n!$ converge et sa somme vaut e^x.

5. La série de terme général $u_n = \ln(1 + \frac{1}{n})$ diverge puisque le terme général de la suite des sommes partielles qui s'écrit :

$$U_n = \sum_{k=1}^{n} \ln(1+\tfrac{1}{k}) = \sum_{k=1}^{n} \ln\left(\frac{k+1}{k}\right) = \sum_{k=1}^{n} \ln(k+1) - \sum_{k=1}^{n} \ln(k) = \ln(n+1)$$

[5] Voir le théorème 5.1 p. 179 du *Cours de première année*.

[6] Voir le *Cours de première année*, théorème 16.4 p. 750.

tend vers $+\infty$.

6. Pour déterminer la nature de la série de terme général $u_n = (-1)^n/\sqrt{n}$, intéressons-nous aux suites $(S_n)_n$ et $(T_n)_n$ extraites de la suite des sommes partielles $(U_n)_n$ associée à cette série et correspondant aux termes d'indices pairs et aux termes d'indices impairs :

$$S_n = U_{2n+2} = \sum_{k=1}^{2n+2} u_k, \qquad T_n = U_{2n+1} = \sum_{k=1}^{2n+1} u_k.$$

La suite $(U_n)_n$, et par conséquent la série $\sum_{n\geqslant 1} u_n$, convergent si et seulement si les deux suites extraites $(S_n)_n$ et $(T_n)_n$ convergent vers une même limite[7]. Cette limite est alors la limite de la suite $(U_n)_n$ et la somme de la série $\sum_{n\geqslant 1} u_n$. Pour tout $n \in \mathbb{N}$, on a

$$S_{n+1} - S_n = \sum_{k=1}^{2n+4} u_k - \sum_{k=1}^{2n+2} u_k = u_{2n+4} + u_{2n+3} = \frac{1}{\sqrt{2n+4}} - \frac{1}{\sqrt{2n+3}} < 0,$$

$$T_{n+1} - T_n = \sum_{k=1}^{2n+3} u_k - \sum_{k=1}^{2n+1} u_k = u_{2n+3} + u_{2n+2} = -\frac{1}{\sqrt{2n+3}} + \frac{1}{\sqrt{2n+2}} > 0,$$

ce qui signifie que la suite $(S_n)_n$ est strictement décroissante et la suite $(T_n)_n$ est strictement croissante. Par ailleurs,

$$S_n - T_n = \sum_{k=1}^{2n+2} u_k - \sum_{k=1}^{2n+1} u_k = u_{2n+2} = \frac{1}{\sqrt{2n+2}}.$$

On en déduit que les deux suites $(S_n)_n$ et $(T_n)_n$ sont des suites adjacentes[8] ; elles convergent vers une même limite ℓ qui vérifie :

$$\forall n \in \mathbb{N} \qquad T_n \leqslant \ell \leqslant S_n.$$

La série $\sum_{n\geqslant 1} \frac{(-1)^n}{\sqrt{n}}$ est donc une série convergente. On peut obtenir une valeur approchée de la somme de cette série en calculant les valeurs de S_n et T_n pour n « assez grand » : ℓ vaut approximativement $\frac{1}{2}(S_n+T_n) = S_n+\frac{1}{2}u_{2n+2}$ et l'erreur commise est inférieure à $\frac{1}{2}(S_n - T_n)$. Pour obtenir une valeur approchée de ℓ avec une erreur inférieure à 10^{-3} (trois chiffres significatifs), il faut choisir n de sorte que :

$$\frac{1}{2}(S_n - T_n) = \frac{1}{2\sqrt{2n+2}} \leqslant 10^{-3}.$$

On a $\frac{1}{2\sqrt{2n+2}} \leqslant 10^{-3}$ pour $n \geqslant 124\,999$. Cette méthode pour calculer une valeur approchée de la somme se révèle un peu coûteuse en calculs. Une valeur approchée de la somme de la série est : -0.604.

---

[7] Voir la proposition 5.19 p. 189 du *Cours de première année*.

[8] Voir la définition 5.6 et le théorème 5.2 p. 185 du *Cours de première année*.

Exercice 1.2

1 - Soit $(S_n)_n$ la suite des sommes partielles associée à la série $\sum_{n\geqslant 1}(-1)^n\frac{1}{n}$.

a) Montrer que la suite extraite de $(S_n)_n$ correspondant aux termes d'indices pairs et la suite extraite correspondant aux termes d'indices impairs sont deux suites adjacentes.

b) En déduire que la série $\sum_{n\geqslant 1}(-1)^n\frac{1}{n}$ converge.

2 - On considère l'application $f : x \in [0, 1] \mapsto \ln(1 + x)$.

a) Justifier que f est indéfiniment dérivable et calculer sa dérivée n–ième pour tout $n \in \mathbb{N}$.

b) En appliquant la formule de Taylor-Lagrange, déterminer la somme de la série $\sum_{n\geqslant 1}(-1)^n\frac{1}{n}$.

Contrairement à ce que pourraient laisser croire les exemples précédents, il est rare que l'on puisse calculer la valeur de la somme d'une série numérique de manière explicite. Dans la plupart des cas, on peut assez facilement établir, à l'aide des différents critères qui seront présentés dans ce chapitre, que la série converge mais il est beaucoup plus difficile d'en calculer la somme. En général, on doit se contenter d'une valeur approchée de la somme de la série. Celle-ci peut être obtenue sur le principe suivant : puisque la suite des sommes partielles $(U_n)_n$ associée à la série $\sum_n u_n$ converge vers la somme μ de la série, une approximation de μ est donnée par $U_N = \sum_{n=0}^{N} u_n$ pour N assez (ou très) grand. Le problème qui se pose alors est de déterminer, avant d'effectuer le calcul, quelle doit être la valeur de N à prendre pour que l'erreur soit inférieure à une valeur donnée (généralement 10^{-3} ou 10^{-6}).

Reprenons l'exemple de la série $\sum_n 1/n!$ qui a pour somme e. Nous avons établi que pour tout $n \in \mathbb{N}$ on avait $|U_n - \mathrm{e}| < \mathrm{e}/(n + 1)!$. Cela signifie que $U_N = \sum_{n=0}^{N} 1/n!$ fournit une valeur approchée de la somme de la série avec une erreur qui est plus petite que $\mathrm{e}/(N + 1)!$. En général, la question se pose dans le sens inverse : pour obtenir une valeur approchée de la somme de la série avec une erreur inférieure à 10^{-6}, quelle valeur de N faut-il prendre ? Dans cet exemple, il faut que N soit tel que $\mathrm{e}/(N + 1)! \leqslant 10^{-6}$. On peut essayer d'avoir recours à MAPLE pour obtenir la valeur de N correspondante. Une première tentative avec la commande `solve( exp(1)/N !=10**(-6), N)` permet de s'apercevoir que MAPLE ne sait pas résoudre formellement cette équation. L'instruction[9] `fsolve(exp(1)/N !=10**(-6), N=0..100)`, permet d'obtenir pour solution approchée 9,876846513. On retient donc $N = 10$.

[9] On observera que l'instruction `fsolve(exp(1)/N !=10**(-6), N)` fournit une solution négative (−2.000000368) qui très clairement ne nous intéresse pas ici. Dans l'instruction `fsolve(exp(1)/N !=10**(-6), N=0..100)` on impose à MAPLE de rechercher une solution comprise entre 0 et 100, donc positive.

Effectuons le calcul avec MAPLE :

```
> evalf(sum( 1/n!,n=0..10));
> evalf(abs(%-exp(1)));
                                2.718281801
                                  2.7E-8
```

En prenant $N = 10$ on a donc une valeur approchée de la somme avec une erreur qui vaut $2,7\,10^{-8}$. L'erreur commise est très inférieure à 10^{-6} ; cela est lié au fait que la quantité $e/(N+1)!$ à la base de notre estimation est un majorant de l'erreur commise (mais n'est pas égal à l'erreur commise).

DÉFINITION 1.3 *Étant donné une série numérique $\sum_n u_n$ convergente, on appelle* reste d'ordre N *de la série et on note R_N la somme de la série $\sum_{n\geqslant N+1} u_n$:*

$$R_N = \sum_{n=N+1}^{+\infty} u_n.$$

Pour tout $N \in \mathbb{N}$, on a

$$\mu = \sum_{n=0}^{+\infty} u_n = \sum_{n=0}^{N} u_n + \sum_{n=N+1}^{+\infty} u_n = U_N + R_N.$$

Comme nous venons de la voir, une valeur approchée de la somme de la série, lorsque celle-ci n'est pas explicitement calculable, est donnée par U_N pour N assez grand. La détermination de la valeur de N à prendre en considération pour obtenir une erreur inférieure à une valeur fixée au préalable nécessite d'être en mesure d'estimer ou de majorer le reste d'ordre N de la série.

Exercice 1.3 Déterminer la nature et calculer le cas échéant la somme des séries suivantes :

$$1\text{-}\sum_{n\geqslant 1}\frac{1}{n(n+1)} \qquad 2\text{-}\sum_{n}\frac{(-1)^n}{n!} \qquad 3\text{-}\sum_{n\geqslant 1}\frac{e^{-n}}{n}.$$

Dans ce chapitre consacré à l'étude des séries numériques, nous allons introduire des critères permettant d'établir qu'une série numérique converge. Nous établirons également des résultats permettant, pour différentes classes de séries numériques, d'obtenir une estimation du reste d'ordre N de la série afin de pouvoir procéder au calcul approché de la somme de la série lorsque celle-ci n'est pas calculable explicitement.

CAUCHY, Augustin-Louis (1789, Paris - 1857, Sceaux).



Augustin-Louis Cauchy commence sa carrière comme ingénieur militaire. En 1816, il obtient un poste de professeur à la Faculté des Sciences de Paris et à l'École Polytechnique et entre à l'Académie des Sciences. L'œuvre de Cauchy est considérable, surtout en analyse où il a su donner le cadre rigoureux nécessaire à son développement. Il a introduit une notion précise de continuité et a élaboré une définition rigoureuse de l'intégrale. Son travail concerne tous les domaines des mathématiques, en particulier les équations différentielles, la théorie des groupes et l'algèbre linéaire.

1.1.2 Critère de Cauchy

THÉORÈME 1.1 (Critère de Cauchy)
Une condition nécessaire et suffisante pour que la série numérique $\sum_n u_n$ *converge est qu'elle vérifie le* critère de Cauchy *:*

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N \in \mathbb{N} \quad \forall (m,n) \in \mathbb{N}^2 \quad (n \geqslant m \geqslant N \Longrightarrow |\sum_{k=m}^{n} u_k| \leqslant \varepsilon).$$

Démonstration La série numérique $\sum_n u_n$ converge si et seulement si la suite des sommes partielles $(U_n)_n$ qui lui est associée converge. Or, une suite numérique converge si et seulement s'il s'agit d'une suite de Cauchy[10]. Ainsi la série numérique $\sum_n u_n$ converge si et seulement si la suite $(U_n)_n$ de terme général $U_n = \sum_{k=0}^{n} u_k$ est une suite de Cauchy, c'est-à-dire[11] si

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists \widetilde{N} \in \mathbb{N} \quad \forall (p,n) \in \mathbb{N}^2 \quad \left(n \geqslant p \geqslant \widetilde{N} \Longrightarrow |U_n - U_p| \leqslant \varepsilon\right). \qquad (2)$$

Soit ε un réel strictement positif et $N = \widetilde{N}+1$ où $\widetilde{N}$ est défini par l'assertion (2). Pour tous réels m et n tels que $n \geqslant m \geqslant \widetilde{N}+1$ on a

$$\left|\sum_{k=m}^{n} u_k\right| = |U_n - U_{m-1}| \leqslant \varepsilon$$

d'après l'assertion (2) où $p = m-1$. Le résultat est établi. □

[10] Voir le *Cours de première année*, théorème 5.4 p. 191.
[11] Voir le *Cours de première année*, définition 5.8 p. 191.

Exemple La série $\sum_{n\geqslant 1}\frac{1}{n}$, appelée série harmonique, diverge. Pour l'établir, montrons que cette série ne satisfait pas au critère de Cauchy. Il s'agit de vérifier que[12]

$$\exists\varepsilon\in\mathbb{R}_+^*\quad\forall N\in\mathbb{N}\quad\exists(m,n)\in\mathbb{N}^2\quad\Big((n\geqslant m\geqslant N)\ \text{ et }\ \Big|\sum_{k=m}^{n}\frac{1}{k}\Big|>\varepsilon\Big).\qquad(3)$$

Étant donné $N\in\mathbb{N}$, considérons les entiers $m=2^N$ et $n=2^{N+1}$; on a

$$\sum_{k=m}^{n}\frac{1}{k}=\sum_{k=2^N}^{2^{N+1}}\frac{1}{k}\geqslant\left(2^{N+1}-2^N+1\right)\ \frac{1}{2^{N+1}}>\left(2^{N+1}-2^N\right)\ \frac{1}{2^{N+1}}=\frac{1}{2},$$

car dans la somme, il y a $2^{N+1}-2^N+1$ termes positifs tous supérieurs à $\frac{1}{2^{N+1}}$. Cela permet d'affirmer que $\sum_{k=m}^{n}\frac{1}{k}>\frac{1}{2}$. On a donc montré l'assertion (3) avec $\varepsilon=\frac{1}{2}$. On en déduit que la série $\sum_{n\geqslant 1}\frac{1}{n}$ ne vérifie pas le critère de Cauchy ; il s'agit donc d'une série divergente. Il est par ailleurs aisé de vérifier que la suite des sommes partielles $(S_n)_n$ est à termes positifs et strictement croissante (et non majorée, sans quoi elle convergerait). La suite $(S_n)_n$ tend donc vers $+\infty$.

1.1.3 Premières propriétés

Proposition 1.1 *Si les deux suites $(u_n)_n$ et $(v_n)_n$ ne diffèrent que par un nombre fini de termes, c'est-à-dire si*

$$\exists N\in\mathbb{N}\quad\forall n\in\mathbb{N}\quad(n\geqslant N\Longrightarrow v_n=u_n)$$

alors les séries $\sum_n u_n$ et $\sum_n v_n$ sont de même nature :

- *ou bien elles convergent toutes les deux ;*
- *ou bien elle divergent toutes les deux.*

Démonstration Les suites des sommes partielles $(U_n)_n$ et $(V_n)_n$ associées aux deux séries $\sum_n u_n$ et $\sum_n v_n$ vérifient pour $n>N$:

$$\begin{aligned}U_n&=\sum_{k=0}^{n}u_k=\sum_{k=0}^{N}u_k+\sum_{k=N+1}^{n}u_k=U_N+\sum_{k=N+1}^{n}u_k,\\ V_n&=\sum_{k=0}^{n}v_k=\sum_{k=0}^{N}v_k+\sum_{k=N+1}^{n}v_k=V_N+\sum_{k=N+1}^{n}v_k.\end{aligned}$$

Puisque $u_k=v_k$ pour tout entier $n\geqslant N$, on en déduit que $U_n-V_n=U_N-V_N$ et par conséquent que $U_n=V_n+(U_N-V_N)$ Comme la quantité U_N-V_N est indépendante de n, on en conclut que[13] :

---

[12] Cette assertion quantifiée est la négation de celle définissant le critère de Cauchy dans la proposition 1.1.

[13] Voir le *Cours de première année*, proposition 5.8 p. 175.

- si la suite $(V_n)_n$ converge vers ν alors $(U_n)_n$ converge vers $\nu + U_N - V_N$;
- si la suite $(U_n)_n$ converge vers μ alors $(V_n)_n$ converge vers $\mu - U_N + V_N$;
- si l'une des deux suites diverge, l'autre diverge aussi.

Le résultat est démontré. □

Exemple Les séries $\sum_{n\geqslant 1} \ln(n-\frac{1}{2})/n^2$ et $\sum_{n\geqslant 1} |\ln(n-\frac{1}{2})|/n^2$ sont de même nature : comme $\ln(x) \geqslant 0$ pour tout $x \geqslant 1$, elles ne diffèrent que par leur premier terme qui vaut respectivement $-\ln(2)$ et $\ln(2)$.

Si les deux suites $(u_n)_n$ et $(v_n)_n$ ne diffèrent que par un nombre fini de termes alors les séries $\sum_n u_n$ et $\sum_n v_n$ sont de même nature mais si elles convergent leur somme est en général différente.

PROPOSITION 1.2 *Une condition nécessaire pour que la série $\sum_n u_n$ converge est que la suite $(u_n)_n$ converge vers 0.*

Démonstration Supposons que la série $\sum_n u_n$ converge et que sa somme est égale à μ où $\mu \in \mathbb{K}$. D'après la définition 1.2, la suite des sommes partielles $(U_n)_n$ converge vers μ. De plus, la suite de terme général $S_n = U_{n-1}$ $(n \in \mathbb{N}^*)$ converge aussi vers μ (cette suite est formée des mêmes termes que la suite $(U_n)_n$, seul l'indicage des termes est modifié). Puisque pour tout $n \in \mathbb{N}^*$ on a $u_n = U_n - U_{n-1}$, on en déduit que la suite $(u_n)_n$ converge vers 0. □

Remarque On dispose là d'un critère simple permettant d'établir qu'une série diverge. Si le terme général de la série ne tend pas vers 0 quand n tend vers $+\infty$, alors la série diverge nécessairement. On qualifie dans ce cas la divergence de la série de *divergence grossière.* ◇

Exemples

1. La série $\sum_n \sin(n)$ diverge grossièrement car la suite de terme général $\sin(n)$ diverge.
2. La série $\sum_n \frac{n}{n+1}$ diverge grossièrement car $\lim_{n\to+\infty} \frac{n}{n+1} = 1$.

Le fait que le terme général d'une série doive tendre vers 0 est une condition nécessaire mais pas suffisante pour que la série converge. Nous avons établi précédemment que la série de terme général $\ln(1+\frac{1}{n})$ divergeait et pourtant $\lim_{n\to+\infty} \ln(1+\frac{1}{n}) = 0$.

PROPOSITION 1.3

✗ *Soient $\sum_n u_n$ et $\sum_n v_n$ deux séries numériques convergentes de sommes respectives μ et ν et λ un scalaire. On a les propriétés suivantes :*

- *la série $\sum_n(u_n + v_n)$ converge et sa somme vaut $\mu + \nu$;*
- *la série $\sum_n(\lambda u_n)$ converge et sa somme vaut $\lambda\mu$.*

✗ *Soient $\sum_n u_n$ une série numérique convergente, $\sum_n v_n$ une série numérique divergente et λ un scalaire non nul. On a les propriétés suivantes :*

- *la série $\sum_n(u_n + v_n)$ diverge ;*
- *la série $\sum_n(\lambda v_n)$ diverge.*

Démonstration Ces propriétés résultent de manière immédiate des propriétés des suites numériques[14] en considérant les suites des sommes partielles associées à ces séries. □

Remarques

1. Il résulte de la proposition 1.3 que l'ensemble des suites numériques dont la série associée converge est un sous-espace vectoriel[15] de l'espace vectoriel des suites convergentes.

2. Par ailleurs, si λ est un scalaire *non nul* alors les séries $\sum_n(\lambda u_n)$ et $\sum_n u_n$ sont de même nature. Cette propriété est très fréquemment utilisée lors de l'étude de la nature d'une série numérique. Par exemple, on peut affirmer que la série $\sum_n \frac{2}{n!}$ converge car on a établi que la série $\sum_n \frac{1}{n!}$ convergeait. De plus, $\sum_{n=0}^{+\infty} \frac{2}{n!} = 2\sum_{n=0}^{+\infty} \frac{1}{n!} = 2\mathrm{e}$.

3. Enfin, la proposition 1.3 permet d'affirmer qu'une série complexe $\sum_n u_n$ converge si et seulement si les deux séries réelles $\sum_n \mathcal{R}\mathrm{e}(u_n)$ et $\sum_n \mathcal{I}\mathrm{m}(u_n)$ convergent. De plus, dans ce cas $\sum_{n=0}^{+\infty} u_n = \sum_{n=0}^{+\infty} \mathcal{R}\mathrm{e}(u_n) + \mathrm{i} \sum_{n=0}^{+\infty} \mathcal{I}\mathrm{m}(u_n)$. ◇

Si la série $\sum_n u_n$ converge et si la série $\sum_n v_n$ diverge alors la série $\sum_n(u_n + v_n)$ diverge. Mais si les deux séries $\sum_n u_n$ et $\sum_n v_n$ divergent, la série $\sum_n(u_n + v_n)$ peut, selon les cas, ou bien converger ou bien diverger (voir les exemples suivants).

Exemples

1. On a vu précédemment que la série $\sum_{n\geqslant 1} \frac{(-1)^n}{n}$ converge et a pour somme $-\ln(2)$ et que la série $\sum_{n\geqslant 1} \frac{1}{n}$ diverge. D'après la proposition 1.3, la série $\sum_{n\geqslant 1}(\frac{(-1)^n}{n} + \frac{1}{n})$ diverge.

2. Les séries $\sum_{n\geqslant 1} \frac{1}{n}$ et $\sum_{n\geqslant 1} \frac{-1}{n}$ divergent. La série $\sum_{n\geqslant 1} \left(\frac{1}{n} + \frac{-1}{n}\right)$ est la série nulle qui converge et qui a pour somme 0.

---

[14] Voir le *Cours de première année*, proposition 5.8 p. 175.

[15] Voir le chap. 8 du *Cours de première année*.

1.2 Séries réelles à termes positifs

1.2.1 Propriétés générales

On a vu que l'on ne change pas la nature d'une série lorsque l'on modifie un nombre fini de ses termes. Les résultats énoncés dans cette partie concernant la nature des séries réelles à termes positifs seront donc également valables pour des séries à termes positifs à partir d'un certain rang.

Par ailleurs, si $\sum_n u_n$ est une série à termes négatifs alors d'après la proposition 1.3, la série $\sum_n -u_n$ est une série à termes positifs qui est de même nature que la série $\sum_n u_n$. Les résultats qui seront énoncés dans cette partie s'appliqueront donc d'une manière plus générale aux séries de signe constant à partir d'un certain rang. On sera toutefois vigilant au fait que pour une série à termes négatifs, dans les résultats faisant intervenir des relations de comparaison $(\leqslant, \geqslant)$, celles-ci doivent être inversées.

PROPOSITION 1.4 *Une condition nécessaire et suffisante pour qu'une série $\sum_n u_n$ à termes positifs converge est que la suite des sommes partielles associée à cette série soit majorée, i.e. que*

$$\exists M \in \mathbb{R}^+ \quad \forall n \in \mathbb{N} \quad \sum_{k=0}^{n} u_k \leqslant M.$$

Démonstration Soit $\sum_n u_n$ une série à termes positifs. La suite $(U_n)_n$ des sommes partielles associée à cette série est une suite croissante puisque :

$$\forall n \in \mathbb{N} \qquad U_{n+1} - U_n = u_{n+1} \geqslant 0.$$

Or, toute suite croissante et majorée converge[16]. Donc si la suite des sommes partielles est majorée, elle converge et par conséquent la série $\sum_n u_n$ aussi.

Par ailleurs[17], toute suite croissante et non majorée tend vers $+\infty$. Donc si la suite des sommes partielles n'est pas majorée, elle tend vers $+\infty$ et la série $\sum_n u_n$ diverge. L'équivalence est établie. □

COROLLAIRE 1.1 *Une condition nécessaire et suffisante pour qu'une série à termes positifs diverge est que la suite des sommes partielles associée à cette série ne soit pas majorée.*

Démonstration Il s'agit de la négation de l'équivalence établie à la proposition 1.4. □

[16] Voir le *Cours de première année*, proposition 5.16 p. 183.

[17] Voir le *Cours de première année*, proposition 5.17 p. 184.

Pour une série numérique à termes positifs, seuls deux cas sont possibles : soit elle converge (et a pour somme un réel positif), soit elle diverge en ayant une somme infinie (*i.e.* la suite des sommes partielles associée à cette série tend vers $+\infty$).

1.2.2 Critères de comparaison

THÉORÈME 1.2 *Soient $(u_n)_n$ et $(v_n)_n$ deux suites réelles* à termes positifs *telles que*

$$\exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad (n \geqslant N \Longrightarrow u_n \leqslant v_n).$$

- *Si la série $\sum_n v_n$ converge alors la série $\sum_n u_n$ converge.*
- *Si la série $\sum_n u_n$ diverge alors la série $\sum_n v_n$ diverge.*

Démonstration $\trianglerighteq$ Supposons que la série $\sum_n v_n$ converge et notons ν sa somme. Pour montrer que la série $\sum_n u_n$ converge, montrons que la suite des sommes partielles $(U_n)_n$ associée à cette série converge. Par hypothèse,

$$\exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad (n \geqslant N \Longrightarrow u_n \leqslant v_n).$$

Pour tout entier $n \geqslant N$, on a

$$U_n = \sum_{k=0}^{N-1} u_k + \sum_{k=N}^{n} u_k \leqslant \sum_{k=0}^{N-1} u_k + \sum_{k=N}^{n} v_k \leqslant \sum_{k=0}^{N-1} u_k + \nu.$$

La dernière inégalité résulte du fait que la série $\sum_n v_n$ étant une série à termes positifs et convergente, de somme ν, on a

$$\sum_{k=N}^{n} v_k \leqslant \sum_{k=N}^{+\infty} v_k \leqslant \sum_{k=0}^{+\infty} v_k = \nu.$$

La suite des sommes partielles $(U_n)_n$ associée à la série à termes positifs $\sum_n u_n$ est donc majorée. D'après la proposition 1.4, cela permet de conclure que la série $\sum_n u_n$ converge.

$\trianglerighteq$ La seconde assertion est la contraposée de la première. Elle est donc vraie[18]. □

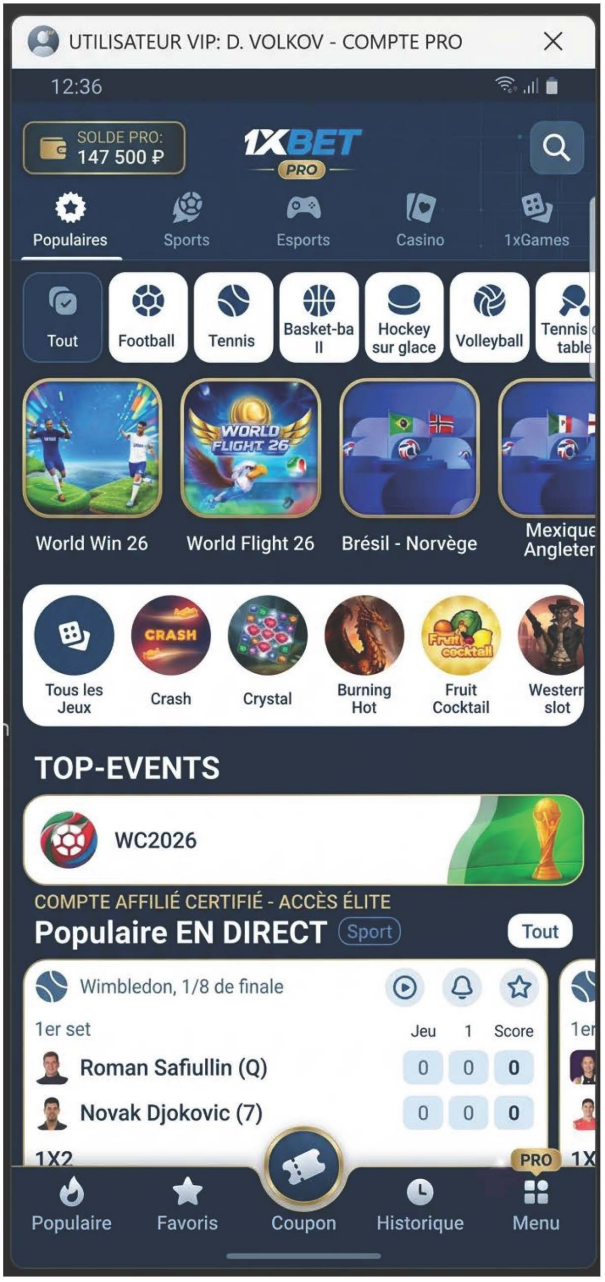
Remarque Dans le cas de séries toutes les deux à termes négatifs, la conclusion du théorème 1.2 est encore vraie sous l'hypothèse suivante :

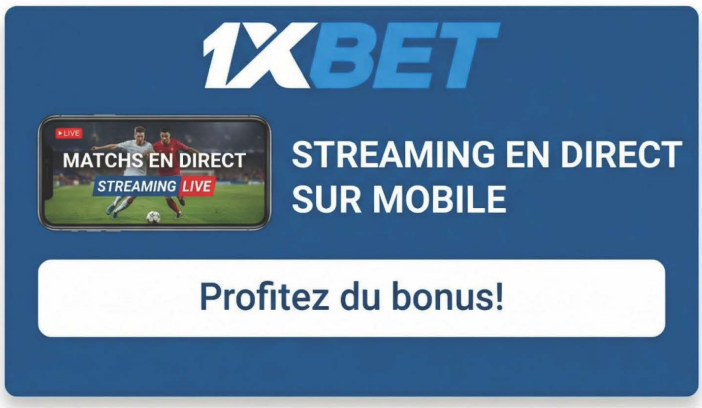
$$\exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad (n \geqslant N \Longrightarrow v_n \leqslant u_n).$$

Noter que le sens de l'inégalité est inversé. ◇

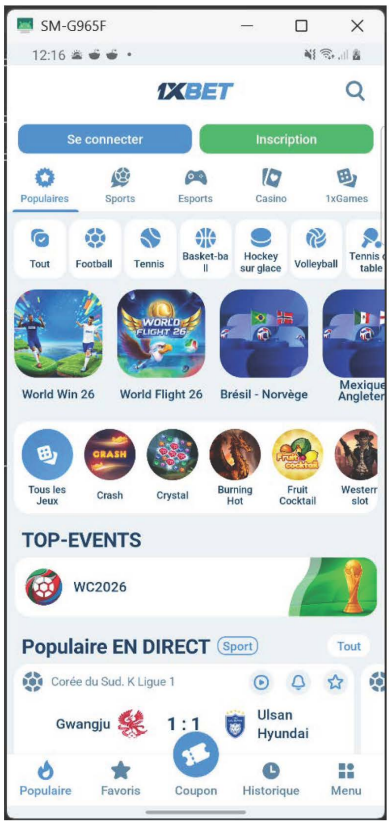
[18] Voir le *Cours de première année*, proposition 1.2 p. 10.

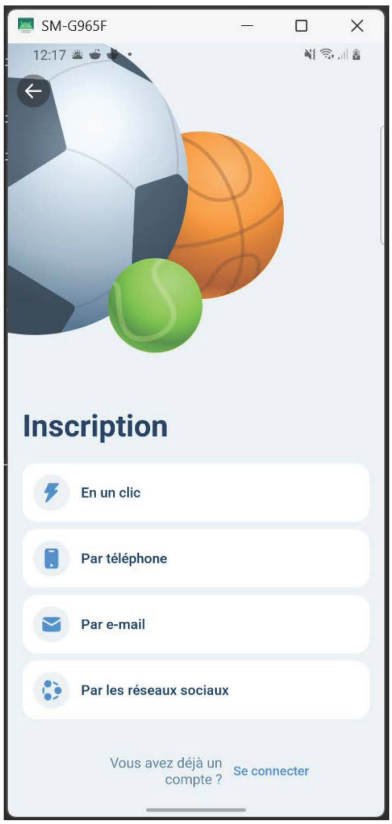

ÉTAPES D'INSCRIPTION
You Tube
pour voir comment creer un compte
CLICK TO WATCH

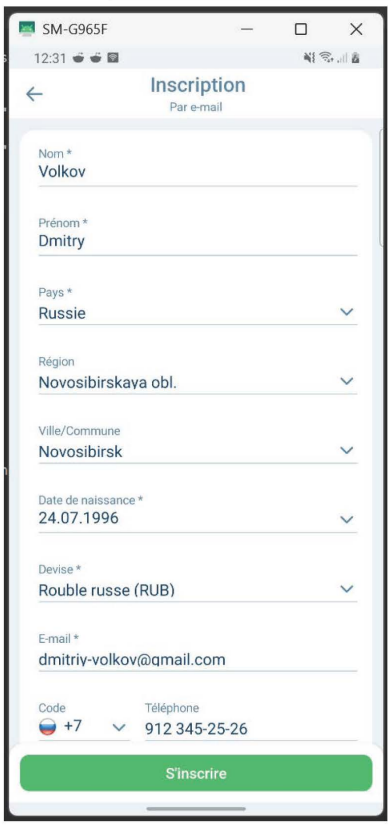

UTILISATEUR VIP: D. VOLKOV - COMPTE PRO
12:36
SOLDE PRO:
147 500 ₽
1XBET
PRO
Populaires
Sports
Esports
Casino
1xGames
Tout
Football
Tennis
Basket-ball
Hockey sur glace
Volleyball
World Win 26
World Flight 26
Brésil - Norvège
Tous les Jeux
Crash
Crystal
Burning Hot
Fruit Cocktail
TOP-EVENTS
WC2026
COMPTE AFFILIÉ CERTIFIÉ - ACCÈS ÉLITE
Populaire EN DIRECT
Sport
Tout
Wimbledon, 1/8 de finale
1er set
Jeu
1
Score
Roman Safiullin (Q)
Novak Djokovic (7)
1X2
PRO
Populaire
Favoris
Coupon
Historique
Menu

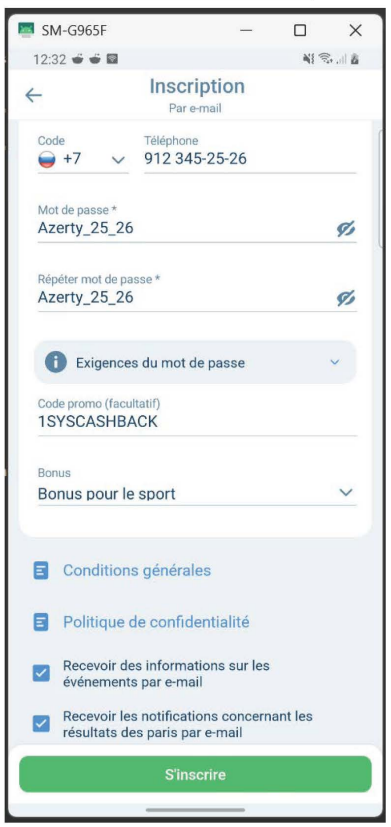

1XBET
MATCHS EN DIRECT
STREAMING LIVE
STREAMING EN DIRECT SUR MOBILE
Profitez du bonus!


CODE PROMO
1SYSCASHBACK


SM-G965F
12:16
1XBET
Se connecter
Inscription
Populaires
Sports
Esports
Casino
1xGames
Tout
Football
Tennis
Basket-ball
Hockey sur glace
Volleyball
World Win 26
World Flight 26
Brésil - Norvège
Tous les Jeux
Crash
Crystal
Burning Hot
Fruit Cocktail
TOP-EVENTS
WC2026
Populaire EN DIRECT
Sport
Tout
Corée du Sud. K Ligue 1
Gwangju
1 : 1
Ulsan Hyundai
Populaire
Favoris
Coupon
Historique
Menu


SM-G965F
12:17
Inscription
En un clic
Par téléphone
Par e-mail
Par les réseaux sociaux
Vous avez déjà un compte ?
Se connecter


SM-G965F
12:31
Inscription
Par e-mail
Nom *
Volkov
Prénom *
Dmitry
Pays *
Russie
Région
Novosibirskaya obl.
Ville/Commune
Novosibirsk
Date de naissance *
24.07.1996
Devise *
Rouble russe (RUB)
E-mail *
dmitriy-volkov@gmail.com
Code
+7
Téléphone
912 345-25-26
S'inscrire


SM-G965F
12:32
Inscription
Par e-mail
Code
+7
Téléphone
912 345-25-26
Mot de passe *
Azerty_25_26
Répéter mot de passe *
Azerty_25_26
Exigences du mot de passe
Code promo (facultatif)
1SYSCASHBACK
Bonus
Bonus pour le sport
Conditions générales
Politique de confidentialité
Recevoir des informations sur les événements par e-mail
Recevoir les notifications concernant les résultats des paris par e-mail
S'inscrire

Exemples

1. Pour tout entier $n \geqslant 2$ on a $n(n-1) \leqslant n^2$ donc $\frac{1}{n^2} \leqslant \frac{1}{n(n-1)}$. Comme la série de terme général $\frac{1}{n(n-1)}$ converge (voir p. 5), la série de terme général $\frac{1}{n^2}$ converge aussi. On peut par ailleurs montrer (voir l'exercice 1.15 p. 53) que sa somme vaut

$$\sum_{n=1}^{+\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

2. Pour tout $n \in \mathbb{N}$ avec $n \geqslant 2$ on a $0 < \ln(n) \leqslant n$ et par conséquent $\frac{1}{\ln(n)} \geqslant \frac{1}{n}$. Comme la série $\sum_{n \geqslant 1} \frac{1}{n}$ diverge, on en déduit que la série $\sum_{n \geqslant 2} \frac{1}{\ln(n)}$ diverge aussi.

On sera vigilant au fait que le résultat du théorème 1.2 n'est valable que pour des séries à termes positifs. Par exemple, pour tout $n \in \mathbb{N}^*$ on a $-\frac{1}{n} \leqslant \frac{(-1)^n}{n}$ et la série $\sum_n -\frac{1}{n}$ diverge alors que la série $\sum_n \frac{(-1)^n}{n}$ converge.

Remarque Si $\sum_n u_n$ est une série à termes positifs convergente alors $\lim_{n \to +\infty} u_n = 0$ et à partir d'un certain rang N, tous les termes de la suite $(u_n)_n$ sont compris entre 0 et 1. Comme pour tout $x \in [0,1]$ on a $x^2 \leqslant x$, on en déduit que pour tout $n \geqslant N$ on a $0 \leqslant u_n^2 \leqslant u_n$. Le théorème 1.2 permet d'en déduire que la série $\sum_n u_n^2$ converge. De même, on montre que si $\sum_n u_n$ est une série à termes positifs convergente, alors pour tout $p \in \mathbb{N}^*$ la série $\sum_n u_n^p$ converge.

On sera vigilant au fait que ce résultat n'est vrai que pour des séries à termes positifs. Par exemple, la série $\sum_n u_n$ où $u_n = (-1)^n/\sqrt{n}$ converge (voir p. 6) mais $u_n^2 = 1/n$ et la série $\sum_n 1/n$ diverge. ◇

Exercice 1.4 Soient $\sum_n u_n$ et $\sum_n v_n$ deux séries à termes positifs qui convergent.
1 - Montrer que pour tout $n \in \mathbb{N}$ on a $\sqrt{u_n \times v_n} \leqslant \frac{1}{2}(u_n + v_n)$, puis en utilisant le théorème 1.2, que la série $\sum_n \sqrt{u_n \times v_n}$ converge.
2 - Montrer que la série $\sum_n u_n \times v_n$ converge.

Exemple Soient $(u_n)_n$ une suite décroissante de réels strictement positifs et $p \in \mathbb{N}$ tel que $p \geqslant 2$. Montrons que les séries de terme général u_n et $v_n = p^n u_{p^n}$ sont de même nature. Ce résultat est connu sous le nom de *règle de la loupe.*

Tout d'abord, considérons la série de terme général $w_n = \sum_{k=p^n}^{p^{n+1}-1} u_k$. Le terme général W_n de la suite des sommes partielles associée à la série $\sum_n w_n$ est donné par

$$W_n = \sum_{\ell=0}^{n} w_\ell = \sum_{\ell=0}^{n} \sum_{k=p^\ell}^{p^{\ell+1}-1} u_k = \sum_{\ell=1}^{p^{n+1}-1} u_n = U_{p^{n+1}-1} - u_0 \qquad (4)$$

où $(U_n)_n$ est la suite des sommes partielles associée à la série $\sum_n u_n$. On déduit de la relation (4) que les suites de terme général W_n et $U_{p^{N+1}-1}$ sont de même nature et que la suite $(W_n)_n$ est, au terme u_0 près, une suite extraite[19] de la suite $(U_n)_n$. Cela implique

- d'une part que si la suite $(U_n)_n$ converge, la suite $(W_n)_n$ converge ;
- d'autre part que, si la suite $(U_n)_n$ diverge, alors, comme elle est strictement croissante, elle tend nécessairement vers $+\infty$ ce qui implique que la suite $(W_n)_n$ tend également vers $+\infty$ (et donc diverge).

En conclusion, les deux suites de sommes partielles $(U_n)_n$ et $(W_n)_n$ et par conséquent les deux séries $\sum_n u_n$ et $\sum_n w_n$ sont de même nature.

Montrons à présent que les deux séries $\sum_n v_n$ et $\sum_n w_n$ sont de même nature. Comme la suite $(u_n)_n$ est une suite décroissante de réels positifs, pour tout $k \in \{p^n, \dots, p^{n+1}-1\}$ on a : $u_{p^{n+1}} \leqslant u_k \leqslant u_{p^n}$. On en déduit que :

$$\sum_{k=p^n}^{p^{n+1}-1} u_{p^{n+1}} \leqslant w_n \leqslant \sum_{k=p^n}^{p^{n+1}-1} u_{p^n}.$$

Or,

$$\sum_{k=p^n}^{p^{n+1}-1} u_{p^{n+1}} = (p-1)p^n u_{p^{n+1}} = (p-1)p^{-1}v_{n+1}$$

$$\text{et} \quad \sum_{k=p^n}^{p^{n+1}-1} u_{p^n} = (p-1)p^n u_{p^n} = (p-1)v_n.$$

Le théorème 1.2, permet de déduire de la relation

$$(p-1)p^{-1}v_{n+1} \leqslant w_n \leqslant (p-1)v_n,$$

- d'une part que si la série $\sum_n w_n$ converge alors la série de terme général $(p-1)p^{-1}v_{n+1}$ converge (et donc $\sum_n v_n$ converge aussi) ;
- et d'autre part que si la série $\sum_n v_n$ converge alors la série $\sum_n w_n$ converge (car les deux séries $\sum_n v_n$ et $\sum_n (p-1)v_n$ sont de même nature).

Finalement, on peut conclure que les trois séries $\sum_n u_n$, $\sum_n w_n$ et $\sum_n v_n$ sont de même nature.

Un autre critère de comparaison très utile est fourni par la proposition suivante.

Proposition 1.5 *Soient $(u_n)_n$ et $(v_n)_n$ deux suites réelles à termes strictement positifs à partir d'un certain rang telles que la suite de terme général u_n/v_n admette pour limite $\ell \in \mathbb{R}^+ \cup \{+\infty\}$.*

1. *Si $\ell > 0$ alors les deux séries $\sum_n u_n$ et $\sum_n v_n$ sont de même nature.*
2. *Si $\ell = 0$ et si la série $\sum_n v_n$ converge alors la série $\sum_n u_n$ converge aussi.*
3. *Si $\ell = +\infty$ et si la série $\sum_n v_n$ diverge alors la série $\sum_n u_n$ diverge aussi.*

[19] En effet, comme $p > 1$, l'application $n \in \mathbb{N} \mapsto p^{n+1} - 1 \in \mathbb{N}$ est strictement croissante.

Démonstration $\trianglerighteq$ Si la suite $(w_n)_n$ de terme général $w_n = u_n/v_n$ converge, c'est nécessairement[20] vers un réel ℓ positif et[21] :

$$\exists N_2 \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad (n \geqslant N_2 \implies |w_n - \ell| \leqslant 1)\,.$$

On en déduit que pour tout entier $n \geqslant N_2$ on a $u_n/v_n \leqslant 1 + \ell$, c'est-à-dire que $u_n \leqslant (1+\ell)v_n$. La proposition 1.3 et le théorème 1.2 permettent de conclure que si la série $\sum_n v_n$ converge alors la série $\sum_n u_n$ converge aussi. La seconde assertion de la proposition est donc justifiée et pour montrer la première, il reste à établir que si la série $\sum_n u_n$ converge alors la série $\sum_n v_n$ converge aussi.

$\trianglerighteq$ Si la suite $(w_n)_n$ converge vers un réel ℓ strictement positif alors la suite $(\widetilde{w}_n)_n$ de terme général $\widetilde{w}_n = w_n^{-1}$ converge vers ℓ^{-1}. Reprenant le raisonnement précédent, on a

$$\exists N_3 \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad \left(n \geqslant N_3 \implies \left|w_n^{-1} - \ell^{-1}\right| \leqslant 1\right).$$

On en déduit que pour tout entier $n \geqslant N_3$ on a $v_n/u_n \leqslant 1 + \ell^{-1}$, c'est-à-dire que $v_n \leqslant (1+\ell^{-1})u_n$. Le théorème 1.2 permet de conclure que si la série $\sum_n u_n$ converge, alors la série $\sum_n v_n$ converge aussi.

$\trianglerighteq$ Si la suite de terme général u_n/v_n tend vers $+\infty$, alors la suite de terme général v_n/u_n tend vers 0. D'après ce qui vient d'être démontré, on en déduit que si la série $\sum_n u_n$ converge alors la série $\sum_n v_n$ converge aussi. La contraposée[22] de cette assertion est : si la série $\sum_n v_n$ diverge alors la série $\sum_n u_n$ diverge. □

Exemples

1. La série $\sum_{n\geqslant 1} \frac{1}{n^3}$ converge car $\frac{1}{n^3}/\frac{1}{n^2} = \frac{1}{n}$ tend vers 0 (quand n tend vers $+\infty$) et la série $\sum_{n\geqslant 1} \frac{1}{n^2}$ converge.

2. La série $\sum_{n\geqslant 1} \frac{1}{\sqrt{n}}$ diverge car $\frac{1}{\sqrt{n}}/\frac{1}{n} = \sqrt{n}$ tend vers $+\infty$ (quand n tend vers $+\infty$) et la série $\sum_{n\geqslant 1} \frac{1}{n}$ diverge.

COROLLAIRE 1.2 *Soient $(u_n)_n$ et $(v_n)_n$ deux suites réelles à termes strictement positifs telles que $u_n \underset{+\infty}{\sim} v_n$. Les deux séries $\sum_n u_n$ et $\sum_n v_n$ sont de même nature.*

[20] Voir la proposition 5.2 p. 169 du *Cours de première année*.

[21] Voir la définition 5.1 p. 168 du *Cours de première année* concernant la convergence d'une suite numérique. On prend ici $\varepsilon = 1$.

[22] Voir la proposition 1.2 p. 10 du *Cours de première année*.

Démonstration Si $u_n \underset{+\infty}{\sim} v_n$, on a $\lim_{n\to+\infty} u_n/v_n = 1$; il suffit d'appliquer la proposition 1.5 dans le cas $\ell = 1$. □

Si $(u_n)_n$ et $(v_n)_n$ sont deux suites réelles à termes strictement positifs telles que $u_n \underset{+\infty}{\sim} v_n$ alors les séries $\sum_n u_n$ et $\sum_n v_n$ sont de même nature mais leur somme est en général différente.

Exemples

1. On a $\frac{1}{n(n+1)} \underset{+\infty}{\sim} \frac{1}{n^2}$. Comme la série $\sum_{n\geqslant 1} \frac{1}{n^2}$ converge, la série $\sum_{n\geqslant 1} \frac{1}{n(n+1)}$ converge aussi. Les sommes de ces séries sont différentes puisque

$$\sum_{n\geqslant 1} \frac{1}{n^2} = \frac{\pi^2}{6} \quad \text{et} \quad \sum_{n\geqslant 1} \frac{1}{n(n+1)} = 1,$$

voir l'exercice 1.3 p. 8 et l'exercice 1.15 p. 53.

2. La série $\sum_{n\geqslant 1} \ln(1-\frac{1}{n})$ est une série à termes négatifs. On a $-\ln(1-\frac{1}{n}) \underset{+\infty}{\sim} \frac{1}{n}$ et $\sum_n \frac{1}{n}$ diverge, donc la série $\sum_{n\geqslant 1} \ln(1-\frac{1}{n})$ diverge aussi.

3. Le corollaire 1.2 est très utile lorsque le terme général de la série fait intervenir des factorielles car la formule de Stirling[23] donne un équivalent à $n!$ au voisinage de $+\infty$:

$$n! \underset{+\infty}{\sim} \sqrt{2\pi n}\left(\frac{n}{\mathrm{e}}\right)^n.$$

À titre d'illustration, déterminons la nature de la série $\sum_n n!/n^n$. Cette série est à termes positifs et d'après la formule de Stirling on a

$$u_n = \frac{n!}{n^n} \underset{+\infty}{\sim} \sqrt{2\pi}\sqrt{n}\mathrm{e}^{-n}.$$

Or[24], pour tout entier n on peut montrer par récurrence que $\sqrt{n}\mathrm{e}^{-n} \leqslant 1/2^n$. Comme la série $\sum_n 1/2^n$ est une série géométrique convergente, le théorème 1.2 permet de conclure que la série $\sum_n \sqrt{n}\mathrm{e}^{-n}$ converge aussi et donc que $\sum_n u_n$ converge.

Dans le corollaire 1.2, l'hypothèse imposant que les suites $(u_n)_n$ et $(v_n)_n$ soient à termes positifs est essentielle. A titre de contre-exemple, considérons les suites $(u_n)_n$ et $(v_n)_n$ dont les termes généraux sont

$$u_n = \frac{(-1)^n}{\sqrt{n}}, \qquad v_n = \ln\left(1 + \frac{(-1)^n}{\sqrt{n}}\right).$$

Ces suites ne sont pas des suites à termes de signe constant. On a $u_n \underset{+\infty}{\sim} v_n$ car $\ln(1+u) \underset{0}{\sim} u$ mais la série $\sum_n u_n$ converge (voir p. 43) alors que la série $\sum_n v_n$ diverge (ceci sera établi ultérieurement, voir p. 49).

---

[23] Cette formule est établie dans le *Cours de première année* p. 199.

[24] Nous verrons ultérieurement des critères qui permettent de conclure directement à la convergence de cette série, voir la proposition 1.13 p. 34.

En complément au corollaire 1.2, l'exercice suivant permet de montrer que, dans le cas de séries à termes positifs convergentes, si les termes généraux de deux séries sont équivalents au voisinage de $+\infty$, alors leurs restes d'ordre n sont également équivalents au voisinage de $+\infty$.

Exercice 1.5 Soient $(u_n)_n$ et $(v_n)_n$ deux suites réelles à termes strictement positifs à partir d'un certain rang telles que $u_n \underset{+\infty}{\sim} v_n$. On suppose que la série $\sum_n v_n$ converge et on considère les restes d'ordre n associés aux deux séries $\sum_n u_n$ et $\sum_n v_n$ définis par $R_n = \sum_{k=n+1}^{+\infty} u_k$ et $R'_n = \sum_{k=n+1}^{+\infty} v_k$.

1 - Montrer que : $\forall \varepsilon \in \mathbb{R}_+^* \ \exists N \in \mathbb{N} \ \forall n \in \mathbb{N} \quad (n \geqslant N \Rightarrow |u_n - v_n| \leqslant \varepsilon v_n)$.

2 - En déduire que :

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad (n \geqslant N \Longrightarrow |R'_n - R_n| \leqslant \varepsilon R'_n).$$

3 - Conclure que $R_n \underset{+\infty}{\sim} R'_n$.

COROLLAIRE 1.3 *Soient $(u_n)_n$ et $(v_n)_n$ deux suites réelles à termes strictement positifs telles que $u_n = o_{+\infty}(v_n)$.*

✗ *Si la série $\sum_n v_n$ converge alors la série $\sum_n u_n$ converge aussi.*

✗ *Si la série $\sum_n u_n$ diverge alors la série $\sum_n v_n$ diverge aussi.*

Démonstration ⊵ Si $u_n = o_{+\infty}(v_n)$ alors $\lim_{n\to+\infty} u_n/v_n = 0$. Il suffit d'appliquer la proposition 1.5 (dans le cas $\ell = 0$) pour établir que si la série $\sum_n v_n$ converge alors la série $\sum_n u_n$ converge aussi.

⊵ Si $u_n = o_{+\infty}(v_n)$ alors $\lim_{n\to+\infty} u_n/v_n = 0$ et puisque $(u_n)_n$ et $(v_n)_n$ sont deux suites réelles à termes positifs on a $\lim_{n\to+\infty} v_n/u_n = +\infty$. Il suffit d'appliquer la proposition 1.5 (dans le cas $\ell = +\infty$) pour établir que si la série $\sum_n u_n$ diverge alors la série $\sum_n v_n$ diverge. □

Exemples

1. La série $\displaystyle\sum_{n\geqslant 2} \frac{1}{n^2 \ln(n)}$ converge car $\displaystyle\frac{1}{n^2 \ln(n)} = o_{+\infty}\left(\frac{1}{n^2}\right)$ et la série $\displaystyle\sum_{n\geqslant 1} \frac{1}{n^2}$ converge.

2. La série $\displaystyle\sum_{n\geqslant 2} \frac{1}{\sqrt{n}\ln(n)}$ diverge car $\displaystyle\frac{1}{n} = o_{+\infty}\left(\frac{1}{\sqrt{n}\ln(n)}\right)$ et la série $\displaystyle\sum_{n\geqslant 1} \frac{1}{n}$ diverge.

Le lecteur attentif aura remarqué que les résultats utilisés pour l'étude de la nature des séries numériques sont très voisins dans leur forme de ceux énoncés dans le chapitre consacré aux intégrales généralisées (chap. 19 du *Cours de première année*). La proposition suivante permet d'en saisir la raison.

PROPOSITION 1.6 (Comparaison à une intégrale généralisée)
Soient n_0 un entier naturel et f une application de $[n_0, +\infty[$ dans $\mathbb{R}$, positive et décroissante. La série $\sum_{n \geqslant n_0} f(n)$ et l'intégrale généralisée $\int_{n_0}^{+\infty} f(t)\,\mathrm{d}t$ sont de même nature.

Démonstration Nous allons commencer par établir une inégalité (voir la relation (5) ci-dessous) qui va ensuite nous permettre d'établir l'équivalence annoncée.

$\trianglerighteq$ Comme[25] f est une application décroissante sur $[n_0, +\infty[$, elle est localement Riemann-intégrable sur $[n_0, +\infty[$. La décroissance de f implique par ailleurs que $f(k+1) \leqslant f(t) \leqslant f(k)$ pour tout $t \in [k, k+1]$ (où $k \in \mathbb{N}$, $k \geqslant n_0$) et par conséquent

$$f(k+1) \leqslant \int_k^{k+1} f(t)\,\mathrm{d}t \leqslant f(k).$$

En sommant ces inégalités pour $k \in \{n_0, \ldots, N-1\}$ (où $N \in \mathbb{N}$, $N > n_0$), il vient

$$\sum_{k=n_0}^{N-1} f(k+1) \leqslant \sum_{k=n_0}^{N-1} \int_k^{k+1} f(t)\,\mathrm{d}t \leqslant \sum_{k=n_0}^{N-1} f(k).$$

Puisque $\sum_{k=n_0}^{N-1} f(k+1) = \sum_{k=n_0+1}^{N} f(k)$ et $\sum_{k=n_0}^{N-1} \int_k^{k+1} f(t)\,\mathrm{d}t = \int_{n_0}^{N} f(t)\,\mathrm{d}t$, on obtient :

$$\sum_{k=n_0+1}^{N} f(k) \leqslant \int_{n_0}^{N} f(t)\,\mathrm{d}t \leqslant \sum_{k=n_0}^{N-1} f(k). \tag{5}$$

$\trianglerighteq$ L'intégrale généralisée $\int_{n_0}^{+\infty} f(t)\,\mathrm{d}t$ sera convergente si[26] l'intégrale indéfinie $F : x \in [n_0, +\infty[\mapsto \int_{n_0}^{x} f(t)\,\mathrm{d}t$ admet une limite quand x tend vers $+\infty$. Comme f est positive, l'application F est croissante. Pour montrer que F a une limite quand x tend vers $+\infty$, il suffit de montrer qu'elle est majorée. Pour tout réel x positif, on a $\mathrm{E}(x) \leqslant x < \mathrm{E}(x) + 1$ (E désignant la fonction partie entière). Puisque f est positive, on en déduit que pour tout $x \in [n_0, +\infty[$

$$\int_{n_0}^{\mathrm{E}(x)} f(t)\,\mathrm{d}t \leqslant F(x) \leqslant \int_{n_0}^{\mathrm{E}(x)+1} f(t)\,\mathrm{d}t$$

puis, compte tenu de la relation (5), que

$$\sum_{k=n_0+1}^{\mathrm{E}(x)} f(k) \leqslant F(x) \leqslant \sum_{k=n_0}^{\mathrm{E}(x)} f(k). \tag{6}$$

[25] Voir le *Cours de première année*, p. 917.
[26] Voir le *Cours de première année*, définition 19.2 p. 918.

Si la série $\sum_n f(n)$ converge et a pour somme le réel S alors, puisque f est positive, on a $\sum_{k=n_0}^{\mathrm{E}(x)} f(k) \leqslant S$. L'application F est donc majorée et par conséquent l'intégrale généralisée $\int_{n_0}^{+\infty} f(t)\,\mathrm{d}t$ converge.

$\trianglerighteq$ Inversement, supposons que l'intégrale généralisée $\int_{n_0}^{+\infty} f(t)\,\mathrm{d}t$ converge et notons $I = \int_{n_0}^{+\infty} f(t)\,\mathrm{d}t = \lim_{x\to+\infty} \int_{n_0}^{x} f(t)\,\mathrm{d}t$. De la relation (5), on déduit que :

$$\sum_{k=n_0}^{N} f(k) = f(n_0) + \sum_{k=n_0+1}^{N} f(k) \leqslant f(n_0) + F(N) = f(n_0) + \int_{n_0}^{N} f(t)\,\mathrm{d}t.$$

Comme f est positive, on a $\int_{n_0}^{N} f(t)\,\mathrm{d}t \leqslant I$. On en déduit que la suite des sommes partielles associée à la série $\sum_{n\geqslant n_0} f(n)$ est majorée par $f(n_0) + I$. Comme f est positive, cette suite est croissante ; par conséquent elle converge[27]. La série $\sum_{n\geqslant n_0} f(n)$ est donc bien une série convergente. □

Interprétation graphique

Pour tout entier $k \geqslant n_0$, $f(k)$ est l'aire du pavé $r_k = [k, k+1] \times [0, f(k)]$. L'aire cumulée des pavés r_k pour $k \geqslant n_0$ est finie si et seulement si l'aire sous la représentation graphique de la fonction f sur l'intervalle $[n_0, +\infty[$ est finie.

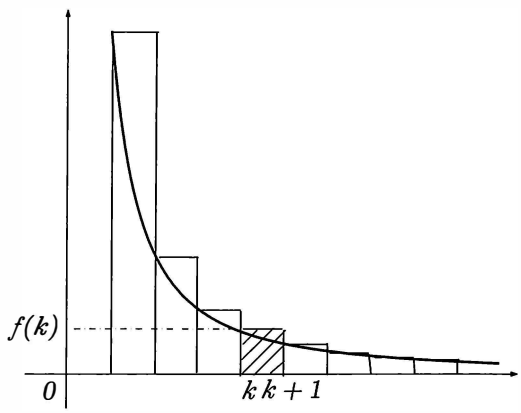


Fig. 1 Illustration de la comparaison série – intégrale généralisée.

Exemple Il est facile de vérifier que la fonction $f : x \mapsto x^{-1}\ln(x)^{-1}$ est positive et décroissante sur $]1, +\infty[$ et qu'elle admet pour primitive $x \mapsto \ln(\ln(x))$. L'intégrale généralisée $\int_2^{+\infty} t^{-1}\ln(t)^{-1}\,\mathrm{d}t$ diverge car

$$\int_2^x t^{-1}\ln(t)^{-1}\,\mathrm{d}t = \ln(\ln(x)) - \ln(\ln(2)) \quad \text{et} \quad \lim_{x\to+\infty} \ln(\ln(x)) = +\infty$$

Ceci implique que la série $\sum_{n\geqslant 2} \frac{1}{n\ln(n)}$ diverge.

[27] Voir la proposition 5.16 p. 183 du *Cours de première année*.

Remarque On utilise le critère de comparaison d'une série à une intégrale généralisée lorsque les autres critères permettant d'établir la nature d'une série à termes positifs ne permettent pas de conclure rapidement et qu'il est facile de déterminer la nature de l'intégrale généralisée par un calcul de primitive. Dans les autres cas ce critère est de peu d'intérêt dans la mesure où les critères permettant d'établir la nature de l'intégrale généralisée d'une fonction positive sont du même type que ceux permettant d'établir la nature d'une série à termes positifs. ◇

Bien que la série $\sum_{n\geqslant n_0} f(n)$ et l'intégrale généralisée $\int_{n_0}^{+\infty} f(t)\,\mathrm{d}t$ soient de même nature, si elles convergent, la somme de la série est en général différente de la valeur de l'intégrale généralisée. Ainsi, l'application $f : x \in [1,+\infty[\mapsto 1/x^2$ est positive et décroissante. Pour tout réel $x \geqslant 1$ on a

$$\int_1^x \frac{1}{t^2}\,\mathrm{d}t = \left[-\frac{1}{t}\right]_1^x = 1 - \frac{1}{x}.$$

On en déduit que $\int_1^{+\infty} f(t)\,\mathrm{d}t = \lim_{x\to+\infty} \int_1^x 1/t^2\,\mathrm{d}t = 1$. Par contre, la série $\sum_{n\geqslant 1} 1/n^2$ converge mais sa somme vaut $\pi^2/6$, voir l'exercice 1.15.

PROPOSITION 1.7 *Soient $n_0 \in \mathbb{N}$ et f une application de $[n_0,+\infty[$ dans $\mathbb{R}$, positive et décroissante, telle que la série $\sum_{n\geqslant n_0} f(n)$ converge.*
Pour $N > n_0$, le reste d'ordre N de la série $\sum_{n\geqslant n_0} f(n)$, qui est défini par $R_N = \sum_{k=N+1}^{+\infty} f(k)$, vérifie :

$$\int_{N+1}^{+\infty} f(t)\,\mathrm{d}t \leqslant R_N \leqslant \int_N^{+\infty} f(t)\,\mathrm{d}t.$$

Démonstration La décroissance de f implique que pour tout $k \in \mathbb{N}$ avec $k \geqslant n_0$, on a $f(k+1) \leqslant f(t) \leqslant f(k)$ pour tout $t \in [k, k+1]$; par conséquent

$$f(k+1) \leqslant \int_k^{k+1} f(t)\,\mathrm{d}t \leqslant f(k).$$

On en déduit que $\int_k^{k+1} f(t)\,\mathrm{d}t \leqslant f(k) \leqslant \int_{k-1}^{k} f(t)\,\mathrm{d}t$ et en sommant ces inégalités pour $k \geqslant N+1$, on obtient :

$$\int_{N+1}^{+\infty} f(t)\,\mathrm{d}t \leqslant \sum_{k=N+1}^{+\infty} f(k) \leqslant \int_N^{+\infty} f(t)\,\mathrm{d}t.$$

Le résultat est démontré. □

Il résulte de la proposition 1.7 qu'en approchant la somme de la série $\sum_{n\geqslant n_0} f(n)$ par $\sum_{n=n_0}^{N} f(n)$, on commet une erreur qui est comprise entre $\int_{N+1}^{+\infty} f(t)\,\mathrm{d}t$ et $\int_N^{+\infty} f(t)\,\mathrm{d}t$.

À titre d'illustration, considérons la série $\sum_{n\geqslant 2} \frac{1}{n^2 \ln(n)}$. L'application $f : x \in [2, +\infty[\mapsto x^{-2} \ln(x)^{-1}$ est une application continue, positive et décroissante. On pourra noter que MAPLE ne sait pas déterminer explicitement la somme de cette série en testant les instructions suivantes :

```
> f:=x-> (x^2*log(x))^(-1);
> sum(f(n),n=2..+infinity);
```

L'erreur commise en approchant la somme de la série par $S_N = \sum_{n=2}^{N} \frac{1}{n^2 \ln(n)}$ est majorée par $\int_N^{+\infty} f(t)\, \mathrm{d}t$. Utilisons MAPLE pour estimer cette erreur :

```
> EN:=int(f(x),x=N..+infinity):
> solve(EN=1E-3,N);
                              167.0027627
```

Ainsi, une valeur approchée à 10^{-3} près de la somme de la série est obtenue en calculant S_{168} :

```
> evalf(sum(f(n),n=2..168));
                              .6045252644
> evalf(int(f(x),x=2..+infinity));
                              .3786710430
```

On remarquera que la somme de la série $\sum_{n\geqslant 2} f(n)$ a une valeur différente de la valeur de l'intégrale généralisée $\int_2^{+\infty} f(t)\, \mathrm{d}t$.

Il existe de nombreux critères permettant d'établir la nature d'une série à termes positifs ; nous en avons donné les principaux. L'exercice suivant fournit un critère supplémentaire, d'usage moins fréquent que ceux donnés précédemment.

Exercice 1.6 Soient $(u_n)_n$ et $(v_n)_n$ deux suites réelles à termes positifs vérifiant

$$\exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad \left(n \geqslant N \implies \frac{u_{n+1}}{u_n} \leqslant \frac{v_{n+1}}{v_n} \right).$$

1 - a) Montrer que pour tout entier $n \geqslant N$ on a $u_n \leqslant \frac{u_N}{v_N} \times v_n$.

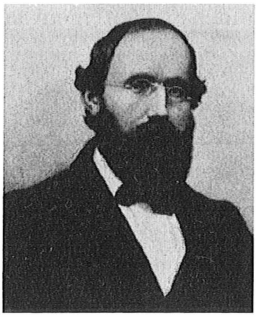
b) En déduire que si la série $\sum_n v_n$ converge alors la série $\sum_n u_n$ converge.

2 - Montrer que si la série $\sum_n u_n$ diverge alors la série $\sum_n v_n$ diverge.

1.2.3 Séries de référence

Nous venons de donner des critères de convergence d'une série par comparaison à une autre série dont la nature est connue. Nous allons à présent introduire des séries de référence, dont nous établirons la nature, et qui seront susceptibles d'être utilisées dans les critères de comparaison.

RIEMANN, Bernhard (1826, Hanovre - 1866, Selasca en Italie).



La thèse de doctorat de Bernhard Riemann, soutenue en 1851, est un travail important sur les fonctions de variables complexes. L'œuvre de Riemann est fondamentale dans de nombreux domaines des mathématiques : étude des surfaces, définition de l'intégrale, équations différentielles, théorie des nombres, séries, etc. Riemann était également profondément attaché aux relations entre les mathématiques et la physique. On lui doit des travaux sur les théories de la chaleur, de la lumière, de l'acoustique, etc. Riemann est mort précocement de tuberculose.

DÉFINITION 1.4 *On appelle* série de Riemann, *toute série réelle de terme général* $\dfrac{1}{n^\alpha}$ $(n \in \mathbb{N}^*)$ *où* α *est un réel.*

En utilisant la proposition 1.6 et les propriétés de convergence des intégrales généralisées de Riemann établies dans le cours de première année[28], on obtient le résultat suivant sur la convergence des séries de Riemann. Compte tenu de l'importance de cette proposition, nous en donnons une démonstration directe sans avoir recours aux résultats sur les intégrales généralisées de Riemann.

PROPOSITION 1.8 (Convergence des séries de Riemann)
La série de Riemann $\displaystyle\sum_{n\geqslant 1} \frac{1}{n^\alpha}$ *converge si* $\alpha > 1$. *Elle diverge sinon.*

Démonstration On procède par disjonction de cas pour déterminer la nature de la série de Riemann $\sum_{n\geqslant 1} 1/n^\alpha$ en fonction de la valeur de α.

$\trianglerighteq$ Si $\alpha < 0$ alors la suite de terme général $n^{-\alpha}$ tend vers $+\infty$. Comme elle ne converge pas vers 0, il y a divergence grossière de la série.

$\trianglerighteq$ Si $\alpha = 0$ alors la suite de terme général $n^{-\alpha}$ est la suite constante égale à 1. Comme elle ne converge pas vers 0, il y a là encore divergence grossière de la série.

$\trianglerighteq$ Si $\alpha = 1$, il s'agit de la série harmonique dont on a établi la divergence (voir l'exemple de la p. 10).

[28] Voir le *Cours de première année* p. 927.

$\trianglerighteq$ Si $\alpha \in]0,1[$ alors pour tout $n \in \mathbb{N}^*$ on a $n^{-\alpha} \geqslant n^{-1} \geqslant 0$. Comme la série $\sum_{n\geqslant 1} n^{-1}$ diverge, on déduit des critères de comparaison (voir le théorème 1.2) que la série $\sum_{n\geqslant 1} n^{-\alpha}$ diverge aussi.

$\trianglerighteq$ Intéressons-nous au cas où $\alpha > 1$. L'application $f : x \in [1,+\infty[\mapsto 1/x^\alpha$ est une application positive et décroissante, donc d'après la proposition 1.6 la série $\sum_{n\geqslant 1} 1/n^\alpha$ et l'intégrale généralisée $\int_1^{+\infty} f(t)\,\mathrm{d}t$ sont de même nature. Pour $x > 1$, on a

$$\int_1^x \frac{1}{t^\alpha}\,\mathrm{d}t = \frac{1}{\alpha-1}\Big(1 - \frac{1}{x^{\alpha-1}}\Big) \quad \text{et} \quad \int_1^{+\infty} \frac{1}{t^\alpha}\,\mathrm{d}t = \lim_{x\to+\infty} \int_1^x \frac{1}{t^\alpha}\,\mathrm{d}t = \frac{1}{\alpha-1}.$$

Comme l'intégrale généralisée $\int_1^{+\infty} 1/t^\alpha\,\mathrm{d}t$ converge, on en déduit que la série $\sum_{n\geqslant 1} 1/n^\alpha$ converge. □

Proposition 1.9 (Règle de Riemann)
Soit $\sum_n u_n$ une série à termes positifs.

✗ *S'il existe un réel $\alpha \in]1,+\infty[$ tel que la suite $(n^\alpha u_n)_n$ converge vers 0 alors la série $\sum_n u_n$ converge.*

✗ *S'il existe un réel $\alpha \in]-\infty,1]$ tel que la suite $(n^\alpha u_n)_n$ tende vers $+\infty$ alors la série $\sum_n u_n$ diverge.*

Démonstration $\trianglerighteq$ S'il existe un réel $\alpha \in]1,+\infty[$ tel que la suite $(n^\alpha u_n)_n$ converge vers 0 alors $u_n = \mathcal{O}_{+\infty}(1/n^\alpha)$. Si $\alpha \in]1,+\infty[$ alors la série de Riemann $\sum_n 1/n^\alpha$ converge et le corollaire 1.3 permet de conclure que la série $\sum_n u_n$ converge.

$\trianglerighteq$ S'il existe un réel $\alpha \in]-\infty,1]$ tel que la suite $(n^\alpha u_n)_n$ tende vers $+\infty$ alors $1/n^\alpha = \mathcal{O}_{+\infty}(u_n)$. Si $\alpha \in]-\infty,1]$ alors la série de Riemann $\sum_n 1/n^\alpha$ diverge et le corollaire 1.3 permet de conclure que la série $\sum_n u_n$ diverge. □

Exemple Déterminons la nature de la série de terme général $u_n = (\ln(n))^{-n}$. Pour tout $n \in \mathbb{N}^*$ on a

$$n^2 u_n = \exp(2\ln(n) - n\ln(\ln(n))).$$

Comme $\ln(n) = \mathcal{O}_{+\infty}(n)$, on a $2\ln(n) = \mathcal{O}_{+\infty}(n\ln(\ln(n)))$. On en déduit[29] que $2\ln(n) - n\ln(\ln(n)) \underset{+\infty}{\sim} -n\ln(\ln(n))$. Puisque $\lim_{n\to+\infty} -n\ln(\ln(n)) = -\infty$, cela implique que $\lim_{n\to+\infty} n^2 u_n = 0$ et par conséquent que la série $\sum_n u_n$ converge.

On se convaincra de l'intérêt de ce critère en notant que u_n n'a pas d'équivalent qui permette de conclure en ayant recours au corollaire 1.2.

[29] Voir le *Cours de première année* p. 694.

BERTRAND, Joseph (1822, Paris - 1900, Paris).



En 1839, année où il entre à l'École Polytechnique, Joseph Bertrand soutient une thèse de doctorat sur la théorie mathématique de l'électricité. Ingénieur au corps des Mines, il enseigne à l'École Polytechnique et à l'École Normale. En 1856, il succède à Charles Sturm à l'Académie des Sciences et ses talents littéraires lui ouvrent en 1874 les portes de l'Académie Française. En 1862, il est nommé titulaire de la chaire de physique et mathématique du Collège de France. Les travaux de J. Bertrand sont variés. Il étudia les applications des équations différentielles à la thermodynamique. Il travailla également sur les courbes et les surfaces puis sur la théorie des probabilités.

DÉFINITION 1.5 (Série de Bertrand)
On appelle série de Bertrand *toute série réelle de terme général* $\dfrac{1}{n^\alpha(\ln(n))^\beta}$ *($n \geqslant 2$) où α est un réel et β est un réel (non nul).*

PROPOSITION 1.10 (Convergence des séries de Bertrand)
La série de Bertrand $\displaystyle\sum_{n\geqslant 2} \frac{1}{n^\alpha(\ln(n))^\beta}$ *où α et β sont deux réels converge dans les cas suivants :*
- *si $\alpha > 1$;*
- *ou si $\alpha = 1$ et $\beta > 1$.*

Elle diverge dans les autres cas.

Démonstration On procède par disjonction de cas.

▷ Cas $\alpha > 1$. On a $\dfrac{1}{n^\alpha(\ln(n))^\beta} = \mathcal{O}_{+\infty}(n^{-\gamma})$ où $\gamma = \frac{1}{2}(1+\alpha)$ car

$$\lim_{n\to+\infty} \frac{\frac{1}{n^\alpha(\ln(n))^\beta}}{\frac{1}{n^\gamma}} = \lim_{n\to+\infty} \frac{1}{n^{\frac{\alpha-1}{2}}(\ln(n))^\beta} = 0.$$

Pour $\alpha > 1$ on a $\frac{1}{2}(1+\alpha) > 1$, donc la série de terme général $n^{-\gamma}$ est une série de Riemann convergente. Le corollaire 1.3 permet de conclure que la série de Bertrand converge si $\alpha > 1$.

▷ Cas $\alpha < 1$. On a $\displaystyle\lim_{n\to+\infty} \frac{\frac{1}{n^\alpha(\ln(n))^\beta}}{\frac{1}{n^\gamma}} = \lim_{n\to+\infty} \frac{1}{n^{\frac{\alpha-1}{2}}(\ln(n))^\beta} = +\infty$ et $\frac{1+\alpha}{2} < 1$ donc la série de terme général $n^{-\gamma}$ est une série de Riemann

divergente. On déduit de la proposition 1.5 que la série de Bertrand diverge si $\alpha < 1$.

$\trianglerighteq$ Cas $\alpha = 1$ et $\beta \leqslant 0$. Pour tout entier $n \geqslant 3$ on a $(\ln(n))^\beta \geqslant 1$ et

$$\frac{1}{n^\alpha(\ln(n))^\beta} = \frac{1}{n(\ln(n))^\beta} \geqslant \frac{1}{n}.$$

Comme la série harmonique $\sum_{n\geqslant 1} \frac{1}{n}$ diverge, on déduit du théorème 1.2 que la série de Bertrand diverge.

$\trianglerighteq$ Cas $\alpha = 1$ et $\beta > 0$. L'application $f : t \in [2, +\infty[\mapsto 1/(t\ln(t)^\beta)$ est positive et décroissante donc d'après la proposition 1.6, la série de Bertrand $\sum_{n\geqslant 2} \frac{1}{n(\ln(n))^\beta}$ et l'intégrale généralisée $\int_2^{+\infty} f(t)\,\mathrm{d}t$ sont de même nature. Considérons l'intégrale indéfinie $F : x \in [2, +\infty[\mapsto \int_2^x f(t)\,\mathrm{d}t$. En effectuant le changement de variable défini par $\phi : s \mapsto \mathrm{e}^s$, on obtient[30] pour tout $x \geqslant 2$,

$$\begin{aligned} F(x) &= \int_{\phi(\ln(2))}^{\phi(\ln(x))} f(t)\,\mathrm{d}t = \int_{\ln(2)}^{\ln(x)} f(\phi(s))\,\phi'(s)\,\mathrm{d}s \\ &= \int_{\ln(2)}^{\ln(x)} \frac{1}{\mathrm{e}^s \ln(\mathrm{e}^s)^\beta}\,\mathrm{e}^s\,\mathrm{d}s = \int_{\ln(2)}^{\ln(x)} s^{-\beta}\,\mathrm{d}s = G(\ln(x)) \end{aligned}$$

où $G : x \in [\ln(2), +\infty[\mapsto \int_{\ln(2)}^x s^{-\beta}\,\mathrm{d}s$. L'intégrale généralisée $\int_{\ln(2)}^{+\infty} s^{-\beta}\,\mathrm{d}s$ est une intégrale généralisée de Riemann[31] qui converge si et seulement si $\beta > 1$. On en déduit que G, et par conséquent que F, ont une limite finie en $+\infty$ si et seulement si $\beta > 1$. □

Remarque On peut également établir la proposition 1.10 en utilisant la proposition 1.6 et les propriétés de convergence des intégrales généralisées de Bertrand établies dans le cours de première année[32]. On pourra aussi consulter l'exemple de la p. 31 pour une troisième façon de démontrer ce résultat. ◇

Exercice 1.7 ***Constante d'Euler***[33]

Pour $n \in \mathbb{N}^*$, on note $\gamma_n = \sum_{k=1}^n 1/k - \ln(n)$ et $u_n = \gamma_{n+1} - \gamma_n$.

1 - a) Montrer que pour tout $x \in \mathbb{R}_+^*$ on a $\frac{x}{1+x} < \ln(1+x) < x$. En déduire que la suite $(\gamma_n)_n$ est décroissante et que la série $\sum_n u_n$ est une série à termes négatifs.

b) Étudier la nature de la série $\sum_n u_n$ et en déduire que la suite $(\gamma_n)_n$ converge. On note γ sa limite.

2 - On considère la suite $(v_n)_n$ de terme général $v_n = \gamma_n - 1/n$. Montrer que les suites $(v_n)_n$ et $(\gamma_n)_n$ sont adjacentes. En déduire une valeur approchée de γ à 10^{-3} près.

3 - Justifier que $\sum_{k=1}^n 1/k \underset{+\infty}{\sim} \ln(n)$.

---

[30] Voir le théorème 18.2 p. 863 du *Cours de première année.*

[31] Voir le *Cours de première année*, p. 927.

[32] Voir le *Cours de première année* p. 928.

1.3 Séries absolument convergentes

Lors de l'étude de la nature d'une série numérique $\sum_n u_n$ dont les termes ne sont pas de signe constant, la première étape consiste à considérer la série numérique de terme général $|u_n|$. Cette série est à termes positifs et les critères de la section 1.2 permettent d'en étudier la nature. Si la série $\sum_n |u_n|$ converge alors on pourra affirmer que la série $\sum_n u_n$ converge aussi (voir la proposition 1.11). Par contre, si la série $\sum_n |u_n|$ diverge, on ne pourra rien conclure sur la nature de la série $\sum_n u_n$ dont il faudra envisager l'étude en ayant recours à des critères spécifiques qui sont présentés dans la section 1.4.

1.3.1 Généralités

DÉFINITION 1.6 *La série numérique $\sum_n u_n$ est dite* absolument convergente *si la série $\sum_n |u_n|$ est une série convergente.*

Remarques

1. On dit aussi que la série numérique $\sum_n u_n$ *converge absolument* pour indiquer qu'il s'agit d'une série absolument convergente.

2. Si la suite $(u_n)_n$ est une suite complexe, $|u_n|$ désigne le module du nombre complexe u_n. Si la suite $(u_n)_n$ est une suite réelle, $|u_n|$ désigne la valeur absolue du nombre réel u_n. Dans les deux cas, la série $\sum_n |u_n|$ est une série à termes réels positifs.

3. La notion de convergence absolue faisant intervenir des séries à termes positifs, tous les résultats établis dans la section 1.2 peuvent (et doivent) être utilisés pour montrer qu'une série converge absolument. ◇

On note $\ell^1(\mathbb{K})$ l'ensemble des séries numériques à coefficients dans $\mathbb{K}$ absolument convergentes. Il s'agit d'un sous-espace vectoriel de l'espace vectoriel $\Sigma(\mathbb{K})$ des séries numériques. En effet,

- Il est non vide car la série nulle est absolument convergente.
- Si $\sum_n u_n$ et $\sum_n v_n$ sont deux séries numériques absolument convergente et α, β deux scalaires alors la série $\alpha \sum_n u_n + \beta \sum_n v_n$ dont le terme général est $\alpha u_n + \beta v_n$ et vérifie :

$$|\alpha u_n + \beta v_n| \leqslant |\alpha||u_n| + |\beta||v_n|.$$

  Or, si $\sum_n u_n$ et $\sum_n v_n$ convergent absolument, alors les séries $\sum_n |u_n|$ et $\sum_n |v_n|$ convergent donc la série de terme général $|\alpha||u_n| + |\beta||v_n|$ converge aussi. On en déduit que $\sum_n |\alpha u_n + \beta v_n|$ converge donc que $\sum_n \alpha u_n + \beta v_n$ qui est absolument convergente appartient à $\ell^1(\mathbb{K})$.

---

(33) La constante d'Euler, moins connue que e ou π, fait partie de ces nombres qui interviennent souvent et dans des situations variées en mathématiques. À ce jour, on sait très peu de choses sur cette constante : on ne sait pas si elle est irrationnelle, algébrique ou transcendante. En 1781, Leonhard Euler avait obtenu les 16 premières décimales de γ : 0, 5772156649015328...

Exercice 1.8
1 - On appelle *série trigonométrique* toute série dont le terme général est de la forme $\alpha_n e^{int}$ où $(\alpha_n)_n$ est une suite réelle ou complexe et t un réel. Donner une condition nécessaire et suffisante (portant sur la suite $(\alpha_n)_n$) pour qu'une série trigonométrique soit absolument convergente.
2 - On appelle *série alternée* une série réelle dont le terme général est de la forme $(-1)^n\varepsilon_n$ ou de la forme $(-1)^{n+1}\varepsilon_n$ où $(\varepsilon_n)_n$ désigne une suite de réels positifs. Donner une condition nécessaire et suffisante (portant sur la suite $(\varepsilon_n)_n$) pour qu'une série alternée soit absolument convergente.

PROPOSITION 1.11 *Toute série numérique absolument convergente est une série convergente.*

Démonstration Par définition, la série numérique $\sum_n u_n$ est absolument convergente si la série $\sum_n |u_n|$ converge. Mais si la série $\sum_n |u_n|$ converge, elle satisfait au critère de Cauchy[34] :

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N \in \mathbb{N} \quad \forall (n,m) \in \mathbb{N}^2 \quad \Big(n \geqslant m \geqslant N \Longrightarrow \Big|\sum_{k=m}^{n} |u_k|\Big| \leqslant \varepsilon\Big). \qquad (7)$$

Pour tout $(m,n) \in \mathbb{N}^2$ avec $n \geqslant m$, on obtient en utilisant l'inégalité triangulaire

$$\Big|\sum_{k=m}^{n} u_k\Big| \leqslant \sum_{k=m}^{n} |u_k| = \Big|\sum_{k=m}^{n} |u_k|\Big|.$$

On déduit alors de l'assertion (7) que

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N \in \mathbb{N} \quad \forall (n,m) \in \mathbb{N}^2 \quad \Big(n \geqslant m \geqslant N \Longrightarrow \Big|\sum_{k=m}^{n} u_k\Big| \leqslant \varepsilon\Big).$$

Ceci signifie que la série numérique $\sum_n u_n$ satisfait au critère de Cauchy[34]. Elle est donc convergente. □

Si la série numérique $\sum_n u_n$ est absolument convergente, alors la série $\sum_n u_n$ converge. Il n'y a toutefois pas de relation directe entre la somme ν de la série $\sum_n |u_n|$ et la somme μ de la série $\sum_n u_n$. En particulier, on n'a pas $|\mu| = \nu$, comme le prouve l'exemple qui suit (il est toutefois facile d'établir que $|\mu| \leqslant \nu$).

[34] Voir le théorème 1.1 p. 9.

Exemple La série de terme général $u_n = (-1)^n/n!$ est une série absolument convergente car $|u_n| = 1/n!$ et on a établi que la série de terme général $1/n!$ convergeait. On a, voir p. 5 et l'exercice 1.3,

$$\sum_{n=0}^{+\infty} \frac{1}{n!} = \mathrm{e} \quad \text{mais} \quad \sum_{n=0}^{+\infty} \frac{(-1)^n}{n!} = \frac{1}{\mathrm{e}}.$$

La proposition 1.11 n'admet pas de réciproque : une série numérique peut être convergente sans être absolument convergente. Par exemple, on a montré que la série $\sum_{n\geqslant 1}(-1)^n/n$ converge (vers $-\ln(2)$) mais que la série $\sum_{n\geqslant 1} 1/n$ diverge.

Remarque Si la série $\sum_n |u_n|$ diverge grossièrement (*i.e.* si la suite $(|u_n|)_n$ ne converge pas vers 0) alors la série $\sum_n u_n$ diverge également grossièrement. Ceci résulte du fait que la suite $(|u_n|)_n$ converge vers 0 si et seulement si[35] la suite $(u_n)_n$ converge vers 0. ◇

Exercice 1.9 Montrer que les séries suivantes convergent absolument.

$$1\text{ - }\sum_n \sin((-1)^n/n^2) \qquad 2\text{ - }\sum_n \frac{(-1)^n}{n^2+(-1)^n} \qquad 3\text{ - }\sum_n \frac{1+\mathrm{i}}{n^2+1}$$

1.3.2 Règles de Cauchy et de D'Alembert

Proposition 1.12 (Règle de Cauchy)
Soit $(u_n)_n$ une suite numérique telle que la suite de terme général $\sqrt[n]{|u_n|}$ converge ou tend vers $+\infty$. On note $\ell = \lim_{n\to+\infty} \sqrt[n]{|u_n|}$ ($\ell \in \mathbb{R}^+ \cup \{+\infty\}$).

- *Si $\ell < 1$ alors la série $\sum_n u_n$ converge absolument.*
- *Si $\ell > 1$ (en particulier si $\ell = +\infty$) alors la série $\sum_n u_n$ diverge grossièrement.*

Démonstration $\trianglerighteq$ Supposons que la suite $(v_n)_n$ de terme général $v_n = \sqrt[n]{|u_n|}$ converge vers un réel $\ell \in [0,1[$. Notons ε le réel strictement positif $\frac{1}{2}(1-\ell)$. Puisque la suite $(v_n)_n$ converge vers ℓ, on a[36] :

$$\exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad \left(n \geqslant N \implies \left|\sqrt[n]{|u_n|} - \ell\right| \leqslant \varepsilon\right).$$

Pour $n \geqslant N$, on a $-\varepsilon \leqslant \sqrt[n]{|u_n|} - \ell \leqslant \varepsilon$, c'est-à-dire

$$0 \leqslant \sqrt[n]{|u_n|} \leqslant \varepsilon + \ell = \tfrac{1}{2}(1+\ell).$$

[35] Voir la proposition 5.4 p. 171 du *Cours de première année*.
[36] Voir la définition 5.1 p. 168 du *Cours de première année*.

On en déduit que $|u_n| \leqslant \left(\frac{1}{2}(1+\ell)\right)^n$ pour tout $n \geqslant N$. Comme $\frac{1}{2}(1+\ell) < 1$, la série géométrique $\sum_n \left(\frac{1}{2}(1+\ell)\right)^n$ converge. On déduit des critères de comparaison (voir le théorème 1.2) que la série $\sum_n |u_n|$ converge. On a ainsi montré que si $\ell < 1$, alors la série $\sum_n u_n$ convergeait absolument.

$\trianglerighteq$ Supposons que la suite $(v_n)_n$ de terme général $v_n = \sqrt[n]{|u_n|}$ converge vers un réel $\ell > 1$. Notons ε le réel strictement positif $\frac{1}{2}(\ell - 1)$. Puisque la suite $(v_n)_n$ converge vers ℓ, on a[(36)] :

$$\exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad \left(n \geqslant N \implies \left|\sqrt[n]{|u_n|} - \ell\right| \leqslant \varepsilon\right).$$

On en déduit que pour tout $n \in \mathbb{N}$ avec $n \geqslant N$,

$$\tfrac{1}{2}(\ell+1) = \ell - \varepsilon \leqslant \sqrt[n]{|u_n|} \leqslant \ell + \varepsilon = \tfrac{1}{2}(3\ell - 1)$$

ce qui implique que $|u_n| \geqslant \left(\frac{1}{2}(1+\ell)\right)^n$. Comme $\frac{1}{2}(\ell+1) > 1$, la suite de terme général $|u_n|$ tend vers $+\infty$. et par conséquent la série $\sum_n |u_n|$ diverge grossièrement. On en conclut la série $\sum_n u_n$ diverge grossièrement.

$\trianglerighteq$ Le cas où la suite $(v_n)_n$ de terme général $v_n = \sqrt[n]{|u_n|}$ tend vers $+\infty$ se traite avec des arguments analogues, voir l'exercice 1.10. □

Exercice 1.10 Soit $(u_n)_n$ une suite numérique telle que la suite de terme général $\sqrt[n]{|u_n|}$ tend vers $+\infty$. Montrer que la série $\sum_n u_n$ diverge grossièrement.

Exemples

1. Considérons la série $\sum_n u_n$ de terme général $u_n = \mathrm{e}^{-n^2}$. Pour tout $n \in \mathbb{N}$ on a

$$\sqrt[n]{|u_n|} = \left(\mathrm{e}^{-n^2}\right)^{\frac{1}{n}} = \mathrm{e}^{-n} \quad \text{et} \quad \lim_{n\to+\infty} \mathrm{e}^{-n} = 0 < 1.$$

Le critère de Cauchy permet de conclure que la série $\sum_n \mathrm{e}^{-n^2}$ converge.

2. Considérons la série $\sum_n u_n$ de terme général $u_n = (\frac{1}{2} - \mathrm{i}\frac{2}{n})^n$. Pour tout $n \in \mathbb{N}$ on a

$$\sqrt[n]{|u_n|} = \left(\frac{1}{4} + \frac{4}{n^2}\right)^{\frac{1}{2}} \quad \text{et} \quad \lim_{n\to+\infty} \sqrt[n]{|u_n|} = \frac{1}{2} < 1.$$

Le critère de Cauchy permet de conclure que la série $\sum_n (\frac{1}{2} - \mathrm{i}\frac{2}{n})^n$ converge.

Si la suite de terme général $\sqrt[n]{|u_n|}$ converge vers 1 ou si cette suite n'a pas de limite, on ne peut pas conclure à la convergence ou à la divergence de la série $\sum_n u_n$. Par exemple, pour les séries de Riemann $\sum_{n\geqslant 1} n^{-\alpha}$ ($\alpha \in \mathbb{R}$) on a

$$\sqrt[n]{|u_n|} = \sqrt[n]{n^{-\alpha}} = n^{-\frac{\alpha}{n}} = \mathrm{e}^{-\frac{\alpha}{n}\ln(n)}$$

et pour tout réel α, cette quantité tend vers 1 quand n tend vers l'infini. Or, on sait que si $\alpha > 1$ la série converge et que si $\alpha \leqslant 1$ elle diverge.

Le critère de Cauchy permet de déterminer la nature des séries de Bertrand $\sum_{n\geqslant 2}\frac{1}{n^\alpha(\ln(n))^\beta}$ en fonction des valeurs de α et β. Nous avons montré à la p. 15 que si $(u_n)_n$ désignait une suite décroissante de réels strictement positifs et $p\in\mathbb{N}$ avec $p\geqslant 2$ alors les séries de terme général u_n et $v_n=a^n u_{p^n}$ étaient de même nature.

Posons $u_n=n^{-\alpha}\ln(n)^{-\beta}$ pour $n\geqslant 2$. Pour $\alpha>0$ ou, pour $\alpha=0$ et $\beta>0$, la suite $(u_n)_n$ est bien une suite décroissante de réels strictement positifs[37]. Dans tous les autres cas, comme $(u_n)_n$ ne converge pas vers 0, il y a divergence grossière de la série $\sum_n u_n$. Pour tout $n\in\mathbb{N}$, on a

$$v_n=\frac{p^n}{p^{n\alpha}\ln(p^n)^\beta}=\left(\frac{p}{p^\alpha}\right)^n\frac{1}{n^\beta\ln(p)^\beta}.$$

Étudions la nature de la série $\sum_n v_n$ en ayant recours au critère de Cauchy :

$$\sqrt[n]{|v_n|}=p^{1-\alpha}\,(n\ln(p))^{-\frac{\beta}{n}}=p^{1-\alpha}\exp\left(-\frac{\beta}{n}\ln(n\ln(p))\right).$$

On en déduit que $\ell=\lim_{n\to+\infty}\sqrt[n]{|v_n|}=p^{1-\alpha}$. Cette limite est strictement inférieure à 1 si et seulement si $\alpha>1$ et elle est strictement supérieure à 1 si et seulement si $\alpha<1$. On en conclut que la série de Bertrand $\sum_{n\geqslant 2}\frac{1}{n^\alpha(\ln(n))^\beta}$ converge si $\alpha>1$ et diverge si $\alpha<1$ et cela quelle que soit la valeur de β.

Exercice 1.11 En utilisant le critère de Cauchy, déterminer la nature de la série $\sum_n u_n$ de terme général :

1 - $u_n=\left(\frac{n}{2n+1}\right)^{\frac{1}{2}n}$ 2 - $u_n=\frac{(n+\mathrm{i})^n}{n^{2n}}$ 3 - $u_n=\left(\mathrm{sh}\left(\frac{1}{n}\right)\right)^{n^2}$

Majoration du reste d'une série dont la convergence absolue a été établie par la règle de Cauchy

Soit $\sum_n u_n$ une série numérique convergente de somme U telle que $\ell=\lim_{n\to+\infty}\sqrt[n]{|u_n|}<1$. Pour tout $r\in]\ell,1[$, il existe un entier naturel N tel que pour tout entier $k\geqslant N$ on ait[38] $\sqrt[k]{|u_k|}\leqslant r$, *i.e.* $|u_k|\leqslant r^k$. On en déduit que pour tout $n\geqslant N$ on a :

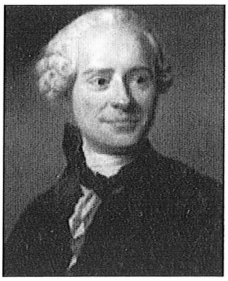
$$|R_n|=\left|U-\sum_{k=0}^{n}u_k\right|=\left|\sum_{k=n+1}^{+\infty}u_k\right|\leqslant\sum_{k=n+1}^{+\infty}|u_k|\leqslant\sum_{k=n+1}^{+\infty}r^k=\frac{r^{n+1}}{1-r},\qquad(8)$$

la dernière relation étant obtenue en remarquant que l'on a le reste d'ordre n d'une série géométrique de raison r. La relation (8) fournit une majoration du reste R_n de la série numérique (pour $n\geqslant N$).

[37] Pour l'établir, on pourra étudier l'application $x\in[2,+\infty[\mapsto x^{-\alpha}\ln(x)^{-\beta}$ et montrer qu'il s'agit d'une application strictement positive et décroissante.

[38] Voir la définition 5.1 p. 168 du *Cours de première année* ; on prend $\varepsilon=r-\ell$.

D'ALEMBERT, Jean Le Rond (1717, Paris - 1783, Paris).



Fils naturel du chevalier Destouches et de Mme de Tencin qui l'abandonna dès sa naissance sur les marches de l'église Saint-Jean-Le-Rond à Paris dont on lui donna le nom. Géomètre et mathématicien, il entra à l'Académie des Sciences à 23 ans. Ami de Diderot (avec qui il codirigea l'Encyclopédie) et de Voltaire, D'Alembert fut l'un des mathématiciens et physiciens les plus renommés du 18^{e} siècle. Il fut élu le 28 novembre 1754 à l'Académie Française[39].

Exemple Considérons la série de terme général $u_n = \left(\mathrm{sh}\left(\frac{1}{n}\right)\right)^{n^2}$ dont la convergence a été établie à l'exercice 1.11. Pour cette série on a $\ell = \lim_{n\to+\infty} \sqrt[n]{|u_n|} = 0$ et pour tout $p \in \mathbb{N}$ avec $p \geqslant 2$ on a $\sqrt[p]{|u_p|} \in]0,1[$. Prenons $r = \sqrt[p]{|u_p|}$ où $p \geqslant 2$. Comme la suite de terme général $\sqrt[n]{|u_n|} = \mathrm{sh}\left(\frac{1}{n}\right)^n$ est décroissante, pour tout $k \geqslant p$ on a

$$\sqrt[k]{|u_k|} = \mathrm{sh}(\tfrac{1}{k})^k \leqslant \mathrm{sh}(\tfrac{1}{p})^p = r.$$

La relation (8) fournit pour cet exemple la majoration suivante du reste d'ordre p pour $p \geqslant 2$:

$$|R_p| \leqslant \frac{\mathrm{sh}(\frac{1}{p})^{p(p+1)}}{1 - \mathrm{sh}(\frac{1}{p})^p}.$$

Utilisons MAPLE pour obtenir une valeur approchée de la somme de la série.

```
> u := n->(sinh(1/n))^(n^2):
> R := n->(sinh(1/n)^(n*(n+1))*sinh(1/n)^n/(sinh(1/n)^n-1)):
> evalf(R(4));
                              -.4578050998 E-14
> Digits:=16: seq(evalf(sum(u(n),n=1..N)),N=2..4);
      1.248935337621633, 1.248995320170124, 1.248995320445086
```

Le reste d'ordre 4 vaut environ 10^{-14}. La convergence de la série est donc extrêmement rapide et $U_4 = 1.248995320445086$ constitue une très bonne approximation de la somme de la série.

[39] http ://www.academie-francaise.fr/immortels/

Proposition 1.13 (Règle de D'Alembert)
Soit $(u_n)_n$ une suite numérique dont les termes sont tous non nuls à partir d'un certain rang. On suppose que la suite de terme général $|u_{n+1}/u_n|$ converge ou tend vers $+\infty$. On note

$$\ell = \lim_{n\to+\infty} \left|\frac{u_{n+1}}{u_n}\right| \qquad (\ell \in \mathbb{R}^+ \cup \{+\infty\}).$$

- *Si $\ell < 1$ alors la série $\sum_n u_n$ converge absolument.*
- *Si $\ell > 1$ (en particulier si $\ell = +\infty$) alors la série $\sum_n u_n$ diverge grossièrement.*

Démonstration $\trianglerighteq$ Supposons que la suite $(v_n)_n$ de terme général $v_n = |u_{n+1}/u_n|$ converge vers un réel $\ell \in [0,1[$. Notons ε le réel strictement positif égal à $\frac{1}{2}(1-\ell)$. Puisque la suite $(v_n)_n$ converge vers ℓ, on a[(40)] :

$$\exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad \left(n \geqslant N \implies \left|\left|\frac{u_{n+1}}{u_n}\right| - \ell\right| \leqslant \varepsilon\right).$$

On en déduit que pour tout $n \in \mathbb{N}$ avec $n \geqslant N$,

$$0 \leqslant \left|\frac{u_{n+1}}{u_n}\right| \leqslant \varepsilon + \ell = \tfrac{1}{2}(1+\ell).$$

On a donc $|u_{n+1}| \leqslant \frac{1}{2}(1+\ell)|u_n|$ pour tout $n \geqslant N$. On montre par récurrence que ceci implique que

$$|u_{n+1}| \leqslant \left(\tfrac{1}{2}(1+\ell)\right)^{n-N} |u_N| = K\left(\tfrac{1}{2}(1+\ell)\right)^n$$

où $K = \left(\frac{1}{2}(1+\ell)\right)^N u_N$. La série de terme général $\left(\frac{1}{2}(1+\ell)\right)^n$ converge car il s'agit d'une série géométrique de raison $r = \frac{1}{2}(1+\ell) < 1$. On déduit des critères de comparaison pour les séries à termes positifs (voir le théorème 1.2) que la série $\sum_n |u_n|$ converge. On a donc montré que pour $\ell < 1$ la série $\sum_n u_n$ convergeait absolument.

$\trianglerighteq$ Supposons que la suite $(v_n)_n$ de terme général $v_n = |u_{n+1}/u_n|$ converge vers un réel $\ell > 1$. Notons ε le réel strictement positif égal à $\frac{1}{2}(\ell - 1)$. Puisque la suite $(v_n)_n$ converge vers ℓ, on a :

$$\exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad \left(n \geqslant N \implies \left|\left|\frac{u_{n+1}}{u_n}\right| - \ell\right| \leqslant \varepsilon\right).$$

On en déduit que pour tout entier $n \geqslant N$, on a

$$\tfrac{1}{2}(\ell+1) = \ell - \varepsilon \leqslant \left|\frac{u_{n+1}}{u_n}\right| \leqslant \ell + \varepsilon = \tfrac{1}{2}(3\ell - 1).$$

[(40)] Voir la définition 5.1 p. 168 du *Cours de première année*.

Comme $\frac{1}{2}(\ell+1) > 1$, la suite de terme général $|u_n|$ est strictement croissante à partir du rang N. Puisqu'elle est à termes positifs, elle ne tend pas vers 0. Cela signifie que la série $\sum_n |u_n|$ diverge grossièrement et par conséquent que la série $\sum_n u_n$ diverge aussi grossièrement.

$\trianglerighteq$ Le cas où la suite $(v_n)_n$ de terme général $v_n = |u_{n+1}/u_n|$ tend vers $+\infty$ se traite avec des arguments analogues, voir l'exercice 1.12. □

Exercice 1.12 Soit $(u_n)_n$ une suite numérique dont les termes sont tous non nuls à partir d'un certain rang. On suppose que la suite de terme général $|u_{n+1}/u_n|$ tend vers $+\infty$. Montrer que la série $\sum_n u_n$ diverge grossièrement.

Exemples

1. Pour tout $z \in \mathbb{C}^*$ fixé, la série numérique $\sum_n u_n$ de terme général $u_n = z^n/n!$ converge absolument car

$$\left|\frac{u_{n+1}}{u_n}\right| = \frac{|z|}{n+1}$$

tend vers 0 quand n tend vers $+\infty$. (Pour $z = 0$, on a la série nulle.)

2. Considérons la série $\sum_n u_n$ de terme général $u_n = n/a^n$ où a désigne un réel non nul fixé. Pour tout $n \in \mathbb{N}^*$, on a

$$\left|\frac{u_{n+1}}{u_n}\right| = \frac{1}{|a|}\,\frac{n+1}{n} \quad \text{d'où} \quad \lim_{n\to+\infty}\left|\frac{u_{n+1}}{u_n}\right| = \frac{1}{a}$$

Le théorème de D'Alembert indique que la série $\sum_n u_n$ converge absolument si $|a| > 1$ et diverge si $|a| < 1$. On peut ajouter que dans le cas où $|a| = 1$ elle diverge grossièrement.

Si la suite de terme général $|u_{n+1}/u_n|$ converge vers 1 ou si cette suite n'a pas de limite, on ne peut pas conclure à la convergence ou à la divergence de la série $\sum_n u_n$. Par exemple, pour les séries de Riemann $\sum_{n\geqslant 1} n^{-\alpha}$ ($\alpha \in \mathbb{R}$) on a

$$\left|\frac{u_{n+1}}{u_n}\right| = \frac{(n+1)^{-\alpha}}{n^{-\alpha}} = \left(1+\frac{1}{n}\right)^{-\alpha}$$

et pour tout $\alpha \in \mathbb{R}$, cette quantité tend vers 1 quand n tend vers l'infini. Or, on sait que si $\alpha > 1$ alors la série converge et que si $\alpha \leqslant 1$ alors elle diverge.

Exercice 1.13 En utilisant la règle de D'Alembert, déterminer la nature de la série $\sum_n u_n$ de terme général :

1 - $u_n = \dfrac{n!}{n^n}$ 2 - $u_n = \dfrac{(n!)^2}{n^2}$ 3 - $u_n = \dfrac{2n}{2^n}$

Remarque On peut démontrer qu'étant donné une suite numérique $(u_n)_n$, si la suite de terme général $|u_{n+1}/u_n|$ tend vers $\ell \in \mathbb{R}^+ \cup \{+\infty\}$ alors la suite de

terme général $\sqrt[n]{|u_n|}$ tend aussi vers ℓ. La règle de Cauchy s'applique donc dans tous les cas où la règle de D'Alembert s'applique. Toutefois selon l'expression du terme général de la série, la règle de D'Alembert peut être plus simple à appliquer que la règle de Cauchy (ce qui justifie l'énoncé des deux règles). Par ailleurs, il résulte de cette propriété que si les limites correspondant aux règles de Cauchy et de D'Alembert existent, elles sont nécessairement égales. En particulier, si l'on a obtenu pour limite 1 avec l'une des deux règles (et que l'on ne peut donc pas conclure), il est inutile d'essayer d'utiliser la seconde règle : elle ne permettra pas, elle non plus, de conclure. ◇

Exemple Intéressons-nous à la série de terme général $u_n = n^{(-1)^n}/2^n$. Pour tout $n \in \mathbb{N}^*$, on a

$$\left|\frac{u_{n+1}}{u_n}\right| = \frac{1}{2}\frac{(n+1)^{(-1)^{n+1}}}{n^{(-1)^n}} = \begin{cases} \dfrac{1}{2n(n+1)} & \text{si } n \text{ est pair} \\ \frac{1}{2}n(n+1) & \text{si } n \text{ est impair} \end{cases}$$

La suite de terme général $|u_{n+1}/u_n|$ n'a pas de limite. La règle de D'Alembert ne s'applique donc pas à cette série. Intéressons-nous à présent à la règle de Cauchy. Pour tout $n \in \mathbb{N}^*$ on a

$$\sqrt[n]{u_n} = \begin{cases} \frac{1}{2}\sqrt[n]{n} & \text{si } n \text{ est pair} \\ \frac{1}{2}\sqrt[n]{1/n} & \text{si } n \text{ est impair} \end{cases}$$

Comme $\lim_{n\to+\infty} \sqrt[n]{n} = 1$ et $\lim_{n\to+\infty} \sqrt[n]{1/n} = 1$, la suite de terme général $\sqrt[n]{u_n}$ converge[41] vers $\frac{1}{2}$. La règle de Cauchy permet de conclure que la série $\sum_n u_n$ converge absolument.

Majoration du reste d'une série dont la convergence absolue a été établie par la règle de D'Alembert

Soit $\sum_n u_n$ une série numérique à termes tous non nuls, convergente. Supposons que $\ell = \lim_{n\to+\infty} \left|\frac{u_{n+1}}{u_n}\right| < 1$. Dans ce cas, pour tout $r \in]\ell, 1[$, il existe $N \in \mathbb{N}$ tel que pour tout entier $k \geqslant N$ on ait[42] $\left|\frac{u_{k+1}}{u_k}\right| \leqslant r$, *i.e.* $|u_{k+1}| \leqslant r\,|u_k|$. On en déduit pour tout $n \in \mathbb{N}$ avec $n \geqslant N$ et pour tout $k \in \mathbb{N}$ avec $k \geqslant n$ la relation $|u_k| \leqslant r^{k-n}|u_n|$. On obtient ainsi pour tout $n \in \mathbb{N}$ avec $n \geqslant N$, la majoration

---

[41] Voir la proposition 5.19 p. 189 du *Cours de première année* : une condition nécessaire et suffisante pour qu'une suite converge vers le réel ℓ est que la suite extraite correspondant aux termes d'indices pairs et la suite extraite correspondant au termes d'indices impairs convergent toutes les deux vers ℓ.

[42] Voir la définition 5.1 p. 168 du *Cours de première année* ; on prend $\varepsilon = r - \ell$.

suivante du reste d'ordre n de la série :

$$|R_n| = \left|\sum_{k=n+1}^{+\infty} u_k\right| \leqslant \sum_{k=n+1}^{+\infty} |u_k|$$
$$\leqslant \sum_{k=n+1}^{+\infty} r^{k-n}\,|u_n| = r^{-n}\,|u_n| \sum_{k=n+1}^{+\infty} r^k = |u_n|\,\frac{r}{1-r}, \qquad (9)$$

la dernière relation étant obtenue en remarquant que la somme correspond au reste d'ordre n d'une série géométrique de raison r.

Exemple Considérons la série de terme général $u_n = n!/n^n$ dont la convergence a été établie à l'exercice 1.13. On a $\ell = \lim_{n\to+\infty} \left|\frac{u_{n+1}}{u_n}\right| = 0$ pour cette série. Pour tout $p \in N^*$, on a $\left|\frac{u_{p+1}}{u_p}\right| = \left(\frac{p}{p+1}\right)^p \in]0,1[$. On peut donc prendre $r = \left(\frac{p}{p+1}\right)^p$. Comme la suite de terme général $v_n = \left(\frac{n}{n+1}\right)^n$ est décroissante, on a pour tout $k \in \mathbb{N}$ avec $k \geqslant p$:

$$\left|\frac{u_{k+1}}{u_k}\right| = \left(\frac{k}{k+1}\right)^k \leqslant \left(\frac{p}{p+1}\right)^p = r.$$

La relation (9) fournit alors pour cet exemple la majoration suivante du reste d'ordre p pour tout $p \in \mathbb{N}$:

$$|R_p| \leqslant |u_p|\,\frac{r}{1-r} = \frac{p!}{(p+1)^p - p^p}.$$

En utilisant MAPLE, on établit que pour que U_N soit une valeur approchée de la somme de la série avec une erreur inférieure à 10^{-6}, il faut choisir $N \geqslant 16$. On trouve alors pour valeur approchée de la somme de la série : 2.879853.

```
> u := n->n!/n^n:
> R := n->n!/((n+1)^n-n^n):
> fsolve(R(p)=1E-6,p=10);
                              15.62180712
> evalf(sum(u(n),n=0..16));
                              2.879853171
```

1.3.3 Produit de Cauchy de deux séries

Nous avons vu que l'ensemble $\Sigma(\mathbb{K})$ des séries numériques à valeurs dans $\mathbb{K}$ muni des lois $+$ (somme de deux séries) et $\cdot$ (produit d'une série par un scalaire) avait une structure d'espace vectoriel.

On peut également définir une loi produit $\times$ sur $\Sigma(\mathbb{K})$ de sorte que $\Sigma(\mathbb{K})$ muni des lois $+$ et $\times$ ait une structure d'anneau commutatif[43].

[43] Voir la définition 2.36 p. 64 du *Cours de première année*.

Rappelons que si la série $\sum_n u_n$ n'est définie qu'à partir d'un rang $n_0 \in \mathbb{N}^*$, on peut convenir de prolonger la série en une série dont les termes sont définis pour tout $n \in \mathbb{N}$ en posant $u_n = 0$ pour tout entier naturel n avec $n < n_0$ (voir p. 2). Cette convention ne modifie ni la nature de la série, ni sa somme.

DÉFINITION 1.7 (Produit de Cauchy de deux séries)
Soient $\sum_n u_n$ et $\sum_n v_n$ deux séries numériques. On appelle produit (de Cauchy) *des séries $\sum_n u_n$ et $\sum_n v_n$, la série $\sum_n w_n$ dont le terme général est :*

$$w_n = \sum_{k=0}^{n} u_k \, v_{n-k}.$$

On vérifie que $(\Sigma(\mathbb{K}), +, \times)$ a bien une structure d'anneau commutatif.

- La loi $+$ est associative et commutative ; elle possède pour élément neutre la série nulle et toute série $\sum_n u_n$ possède un symétrique pour cette loi qui est la série de terme général $-u_n$.
- La loi $\times$ est associative, commutative et distributive sur la loi $+$; elle possède pour élément neutre la série $\sum_n e_n$ où $e_0 = 1$ et pour tout $n \in \mathbb{N}^*$, $e_n = 0$.

Toutefois, $(\Sigma(\mathbb{K}), +, \times)$ n'est pas un corps car les séries dont le premier terme est nul n'ont pas de symétrique pour la loi $\times$ (on a $w_0 = u_0 v_0$ donc si $u_0 = 0$, il n'existe pas de scalaire v_0 tel que $u_0 v_0 = e_0 = 1$).

Remarque Ce produit de Cauchy entre séries numériques est du même type que le produit entre polynômes[44] ou entre matrices[45]. Dans $\mathbb{K}[X]$, le polynôme produit des polynômes $P = \sum_{k=0}^{p} a_k X^k$ et $Q = \sum_{k=0}^{q} b_k X^k$ est le polynôme $R = \sum_{k=0}^{p+q} c_k X^k$ où

$$\forall n \in \{0, \dots, p+q\} \qquad c_n = \sum_{k=0}^{n} a_k \, b_{n-k}.$$

Dans $\mathcal{M}_p(\mathbb{K})$, la matrice produit des matrices $A = (a_{i,j})_{i,j}$ et $B = (b_{i,j})_{i,j}$ est la matrice $C = (c_{i,j})_{i,j}$ où

$$\forall (i,j) \in \{0, \dots, p\}^2 \qquad c_{i,j} = \sum_{k=0}^{p} a_{i,k} \, b_{k,j}.$$

◇

Le produit de Cauchy de deux séries convergentes n'est pas nécessairement une série convergente comme le prouve l'exemple suivant.

[44] Voir la définition 6.7 p. 221 du *Cours de première année*.
[45] Voir la définition 10.8 p. 398 du *Cours de première année*.

Exemple Nous avons établi p. 6 que la série de terme général $u_n = (-1)^n/\sqrt{n}$ était convergente. Déterminons la série $\sum_n w_n$ produit de $\sum_{n\geqslant 1} u_n$ par elle-même. On pose $u_0 = 0$ de sorte que pour tout $n \in \mathbb{N}^*$:

$$w_n = \sum_{k=0}^{n} u_k\, u_{n-k} = (-1)^n \sum_{k=1}^{n-1} \frac{1}{\sqrt{k(n-k)}}.$$

L'étude de la fonction $x \mapsto (x(n-x))^{-\frac{1}{2}}$ montre que les n termes de la somme $S_n = \sum_{k=1}^{n-1} \frac{1}{\sqrt{k(n-k)}}$. sont tous supérieurs à $\frac{1}{2n}$. On a donc $S_n \geqslant n \times \frac{1}{2n} = \frac{1}{2}$. On en déduit que la suite $(S_n)_n$ ne converge pas vers 0 et par conséquent que la série de terme général w_n diverge grossièrement.

Le résultat suivant, que nous admettrons, donne une condition suffisante pour que le produit de deux séries convergentes soit une série convergente.

PROPOSITION 1.14 *Si $\sum_n u_n$ et $\sum_n v_n$ sont deux séries numériques absolument convergentes alors la série $\sum_n w_n$ produit de $\sum_n u_n$ et de $\sum_n v_n$ est absolument convergente. De plus,*

$$\sum_{n=0}^{+\infty} w_n = \left(\sum_{n=0}^{+\infty} u_n\right)\left(\sum_{n=0}^{+\infty} v_n\right).$$

Exemples

1. La série $\sum_n 1/n!$ est une série absolument convergente dont la somme est e. Le produit de cette série avec elle-même a pour terme général :

$$w_n = \sum_{k=0}^{n} \frac{1}{k!(n-k)!} = \frac{1}{n!} \sum_{k=0}^{n} C_n^k = \frac{2^n}{n!}.$$

On déduit de la proposition 1.14 que la série $\sum_n \frac{2^n}{n!}$ est absolument convergente et de somme e^2.

2. Pour tout $z \in \mathbb{C}$ avec $|z| < 1$ la série géométrique de raison z est absolument convergente de somme $1/(1-z)$. Le produit de cette série avec elle-même a pour terme général

$$w_n = \sum_{k=0}^{n} z^k\, z^{n-k} = (n+1)z^n.$$

On déduit de la proposition 1.14 que la série $\sum_n (n+1)z^n$ est absolument convergente et de somme $1/(1-z)^2$.

Remarque On a également le résultat suivant : si $\sum_n u_n$ est une série numérique convergente et $\sum_n v_n$ est une série numérique absolument convergente alors la série $\sum_n w_n$ produit de $\sum_n u_n$ et $\sum_n v_n$ est convergente. ◇

Calcul du produit de Cauchy de deux séries avec MAPLE

Il existe une bibliothèque de MAPLE dédiée au calcul avec les séries. Il s'agit de la bibliothèque `powseries`. Pour illustrer l'utilisation de cette bibliothèque, vérifions le résultat du calcul du produit de la série $\sum_n u_n$ de terme général $u_n = 1/n!$ par elle-même effectué dans l'exemple précédent. La commande `powcreate` est utilisée pour définir le terme général de la série ; la commande `multiply` pour effectuer le produit de Cauchy de deux séries.

```
> with(powseries):
> powcreate(u(n)=1/n!,u(0)=1):
> w:=multiply(u,u):
> seq(w(n),n=0..10);
```

$$1, 2, 2, \frac{4}{3}, \frac{2}{3}, \frac{4}{15}, \frac{4}{45}, \frac{8}{315}, \frac{2}{315}, \frac{4}{2835}, \frac{4}{14175}$$

```
> seq(2^n/n!,n=0..10);
```

$$1, 2, 2, \frac{4}{3}, \frac{2}{3}, \frac{4}{15}, \frac{4}{45}, \frac{8}{315}, \frac{2}{315}, \frac{4}{2835}, \frac{4}{14175}$$

Notre résultat semble conforme à celui obtenu par MAPLE.

1.4 Séries semi-convergentes

Comme cela a été indiqué à la section 1.3, pour étudier la nature d'une série numérique $\sum_n u_n$ dont les termes ne sont pas de signe constant, il convient dans un premier temps de s'intéresser à la série numérique $\sum_n |u_n|$. Cette série est à termes positifs et les critères pour les séries à termes positifs permettent d'en étudier la nature. Nous avons montré que si la série $\sum_n |u_n|$ convergeait alors la série $\sum_n u_n$ convergeait aussi mais que si la série $\sum_n |u_n|$ divergeait aucune conclusion n'était possible sur la nature de la série $\sum_n u_n$. Dans cette section nous allons présenter des résultats permettant d'étudier la nature des séries non absolument convergentes.

1.4.1 Définition

DÉFINITION 1.8 *Une série est dite* semi-convergente *si elle est convergente sans être absolument convergente.*

Exemple La série harmonique alternée $\sum_{n\geqslant 1}(-1)^n/n$ est une série qui est semi-convergente. Elle converge et sa somme vaut $-\ln(2)$ mais elle ne converge pas absolument car $|(-1)^n/n| = 1/n$ et la série harmonique $\sum_{n\geqslant 1} 1/n$ diverge.

Deux types de séries semi-convergentes sont fréquemment utilisées dans la pratique : les séries alternées et les séries trigonométriques. Nous allons étudier ces deux types de séries et donner des critères permettant d'établir leur convergence. Les résultats pour la semi-convergence de ces séries sont des cas particuliers d'un résultat plus général connu sous le nom de *règle d'Abel.*

1.4.2 Séries alternées

DÉFINITION 1.9 *On appelle* série alternée *une série numérique dont le terme général est de la forme* $(-1)^n\alpha_n$ *ou* $(-1)^{n+1}\alpha_n$, *où* $(\alpha_n)_n$ *désigne une suite de réels positifs.*

Une série alternée $\sum_n u_n$ est donc une série dont les termes u_n changent de signe une fois sur deux :

1. ou bien la sous-suite $(u_{2n})_n$ est positive et la sous-suite $(u_{2n+1})_n$ est négative ;
2. ou bien la sous-suite $(u_{2n})_n$ est négative et la sous-suite $(u_{2n+1})_n$ est positive.

Remarques

1. Si la suite $(\alpha_n)_n$ ne tend pas vers 0, les séries alternées $\sum_n(-1)^n\alpha_n$ et $\sum_n(-1)^{n+1}\alpha_n$ divergent grossièrement.

2. On a $\sum_n(-1)^{n+1}\alpha_n = -\sum_n(-1)^n\alpha_n$; d'après la proposition 1.3 les deux séries $\sum_n(-1)^n\alpha_n$ et $\sum_n(-1)^{n+1}\alpha_n$ sont de même nature (et dans le cas où elles convergent leur somme est opposée). Nous nous contenterons dans ce qui suit de l'étude des propriétés de la série $\sum_n(-1)^n\alpha_n$ avec $(\alpha_n)_n$ positive. ◇

Exemples

1. La série $\sum_{n\geqslant 1}(-1)^n/n$ est une série alternée, ce qui justifie son nom de *série harmonique alternée.*

2. La série $\sum_{n\geqslant 1} u_n$ de terme général $u_n = 1-\sqrt{1+\frac{(-1)^n}{n}}$ est une série alternée. Pour tout $k \in \mathbb{N}^*$, on a

$$\begin{aligned} u_{2k} &= 1-\sqrt{1+\frac{1}{2k}} = 1-\sqrt{\frac{2k+1}{2k}} < 0, \\ u_{2k+1} &= 1-\sqrt{1-\frac{1}{2k+1}} = 1-\sqrt{\frac{2k}{2k+1}} > 0. \end{aligned}$$

On a donc

$$u_n = (-1)^{n+1}\alpha_n \quad \text{où} \quad \begin{cases} \alpha_n = \sqrt{\frac{n+1}{n}} - 1 & \text{si } n \text{ est pair} \\ \alpha_n = 1-\sqrt{\frac{n-1}{n}} & \text{si } n \text{ est impair} \end{cases}$$

autrement dit, $u_n = (-1)^{n+1}\alpha_n$ avec $\alpha_n = (-1)^n\left(\sqrt{1+\frac{(-1)^n}{n}}-1\right)$.

THÉORÈME 1.3 (Critère spécial pour les séries alternées)
Si la suite $(\alpha_n)_n$ est une suite de réels positifs, décroissante et convergeant vers 0 alors la série alternée $\sum_n (-1)^n\alpha_n$ converge et a pour somme un réel U qui vérifie pour tout $n \in \mathbb{N}$

$$U_{2n+1} \leqslant U \leqslant U_{2n}$$

où $(U_n)_n$ désigne la suite des sommes partielles associée à la série alternée.

Démonstration Soit $\sum_n u_n$ une série de terme général $u_n = (-1)^n\alpha_n$ où $(\alpha_n)_n$ désigne une suite de réels positifs, décroissante et convergeant vers 0. Considérons la suite $(U_n)_n$ des sommes partielles associée à cette série alternée. Par définition, la série alternée $\sum_n u_n$ converge si la suite $(U_n)_n$ converge. Une condition nécessaire et suffisante[46] pour que la suite $(U_n)_n$ converge est que la suite extraite $(A_n)_n$ correspondant à ses termes d'indice pair et que la suite extraite $(B_n)_n$ correspondant à ses termes d'indice impair convergent toutes les deux et admettent la même limite. Nous allons montrer que les suites $(A_n)_n$ et $(B_n)_n$ sont deux suites adjacentes[47].

Pour tout $n \in \mathbb{N}$ on a :

$$\begin{aligned} A_{n+1} - A_n &= U_{2n+2} - U_{2n} = \sum_{k=0}^{2n+2} u_k - \sum_{k=0}^{2n} u_k = u_{2n+2} + u_{2n+1} \\ &= (-1)^{2n+2}\alpha_{2n+2} + (-1)^{2n+1}\alpha_{2n+1} = \alpha_{2n+2} - \alpha_{2n+1}. \end{aligned}$$

La suite $(\alpha_n)_n$ étant une suite décroissante de réels positifs, on a $A_{n+1} - A_n \leqslant 0$ ce qui montre que la suite $(A_n)_n$ est décroissante. Par ailleurs, pour tout $n \in \mathbb{N}$ on a :

$$\begin{aligned} B_{n+1} - B_n &= U_{2n+3} - U_{2n+1} = \sum_{k=0}^{2n+3} u_k - \sum_{k=0}^{2n+1} u_k = u_{2n+3} + u_{2n+2} \\ &= (-1)^{2n+3}\alpha_{2n+3} + (-1)^{2n+2}\alpha_{2n+2} = \alpha_{2n+2} - \alpha_{2n+3}. \end{aligned}$$

La suite $(\alpha_n)_n$ étant une suite décroissante de réels positifs, on a $B_{n+1} - B_n \geqslant 0$ ce qui montre que la suite $(B_n)_n$ est croissante.

Pour montrer que les suites $(A_n)_n$ et $(B_n)_n$ sont deux suites adjacentes, il reste à vérifier que $\lim_{n\to+\infty}(A_n - B_n) = 0$. Pour tout $n \in \mathbb{N}$ on a :

$$A_n - B_n = U_{2n} - U_{2n+1} = \sum_{k=0}^{2n} u_k - \sum_{k=0}^{2n+1} u_k = -u_{2n+1} = \alpha_{2n+1}.$$

[46] Voir la proposition 5.9 p. 189 du *Cours de première année*.
[47] Voir la définition 5.6 p. 185 du *Cours de première année*.

La suite $(\alpha_n)_n$ convergeant vers 0, on a bien $\lim_{n\to+\infty}(A_n - B_n) = 0$ et les suites $(A_n)_n$ et $(B_n)_n$ sont adjacentes. Elles convergent donc et ont même limite[48]. Leur limite commune U vérifie[48] :

$$\forall n \in \mathbb{N} \qquad B_n = U_{2n+1} \leqslant U \leqslant U_{2n} = A_n.$$

Le théorème est démontré. □

Exemples

1. La série $\sum_{n\geqslant 1}(-1)^n/\sqrt{n}$ est une série alternée convergente. En effet, la suite $(\alpha_n)_n$ de terme général $\alpha_n = 1/\sqrt{n}$ est une suite à termes positifs, tendant vers 0 et strictement décroissante puisque pour tout entier $n \geqslant 1$,

$$\alpha_{n+1} - \alpha_n = \frac{1}{\sqrt{n+1}} - \frac{1}{\sqrt{n}} = -\frac{1}{\sqrt{n+1}\sqrt{n}(\sqrt{n+1}+\sqrt{n})} < 0.$$

2. Nous avons vu que la série $\sum_{n\geqslant 1} u_n$ de terme général $u_n = 1 - \sqrt{1 + \frac{(-1)^n}{n}}$ était une série alternée et que pour tout $n \in \mathbb{N}^*$ on avait $u_n = (-1)^{n+1}\alpha_n$ avec $\alpha_n = (-1)^n\left(\sqrt{1+\frac{(-1)^n}{n}} - 1\right) > 0$. La suite $(\alpha_n)_n$ converge clairement vers 0. On établira à titre d'exercice qu'elle est bien décroissante. Il s'agit donc d'une série alternée convergente.

3. Considérons la série $\sum_n u_n$ de terme général $u_n = (-1)^{n+1}\alpha_n$ où $\alpha_n = \ln\left(\frac{(4n+4)(4n+5)}{(4n+2)(4n+3)}\right)$. On a

$$\alpha_n = \ln\left(1 + \frac{3}{2n+1} - \frac{2}{4n+3}\right) \underset{+\infty}{\sim} \frac{3}{2n+1} - \frac{2}{4n+3} \underset{+\infty}{\sim} \frac{1}{n}.$$

La suite $(\alpha_n)_n$ converge donc vers 0 et est positive à partir d'un certain rang. Pour tout $n \in \mathbb{N}$, on a

$$\begin{aligned}\alpha_{n+1} - \alpha_n &= \ln\left(\frac{(4n+8)(4n+9)(4n+2)(4n+3)}{(4n+6)(4n+7)(4n+5)(4n+4)}\right) \\ &= \ln\left(1 - \frac{32n^2+88n+51}{(2n+3)(4n+7)(n+1)(4n+5)}\right).\end{aligned}$$

Comme $32n^2 + 88n + 51 > 0$ pour tout $n \in \mathbb{N}$ et que la fonction logarithme est négative sur $]0,1[$, on en déduit que $\alpha_{n+1} - \alpha_n < 0$. La suite $(\alpha_n)_n$ est donc décroissante et on peut en conclure que la série $\sum_n u_n$ converge.

Remarque Il est établi à l'exercice 1.8 qu'une condition nécessaire et suffisante pour que les séries alternées $\sum_n(-1)^n\alpha_n$ et $\sum_n(-1)^{n+1}\alpha_n$ convergent absolument est que la série $\sum_n \alpha_n$ converge. Si la série $\sum_n \alpha_n$ converge il est

[48] Voir le théorème 5.2 p. 185 du *Cours de première année*.

donc inutile et fortement déconseillé d'utiliser le critère spécial pour les séries alternées. ◇

La condition de décroissance de la suite $(\alpha_n)_n$ est essentielle. Par exemple, on ne peut pas appliquer le critère spécial pour les séries alternées à la série alternée de terme général $u_n = \frac{(-1)^n}{\sqrt{n}+(-1)^n}$.

La suite $(\alpha_n)_n$ de terme général $\alpha_n = \frac{1}{\sqrt{n}+(-1)^n}$ est bien une suite à termes positifs, tendant vers 0 mais elle n'est pas décroissante. Pour le vérifier, il suffit de remarquer que si n désigne un entier pair supérieur à 2, on a

$$\frac{\alpha_{n+1}}{\alpha_n} = \frac{\sqrt{n}+1}{\sqrt{n+1}-1} = \frac{(\sqrt{n}+1)(\sqrt{n+1}+1)}{n} > 1$$

car $\sqrt{n}+1 > \sqrt{n}$ et $\sqrt{n+1}+1 > \sqrt{n}$. On remarquera que la suite $(\alpha_n)_n$ possède une sous-suite croissante, donc cette suite ne peut pas être décroissante. Il ne faudrait toutefois pas penser pour autant que la suite $(\alpha_n)_n$ est croissante. On peut montrer qu'elle n'est pas monotone en vérifiant que $\frac{\alpha_{n+1}}{\alpha_n} < 1$ si n est un entier impair. On ne peut donc pas appliquer le critère spécial pour les séries alternées à la série alternée de terme général $u_n = \frac{(-1)^n}{\sqrt{n}+(-1)^n}$. On peut toutefois montrer par un autre procédé que cette série alternée diverge.

Pour cela, considérons la série de terme général $v_n = \frac{1}{n+(-1)^n\sqrt{n}}$. Il s'agit d'une série à termes positifs (car pour tout $n \in \mathbb{N}^*$ on a $\sqrt{n} \leqslant n$) qui diverge car $v_n \underset{+\infty}{\sim} \frac{1}{n}$. Or, pour tout entier $n \geqslant 2$ on a

$$v_n = \frac{(-1)^n}{\sqrt{n}} - \frac{(-1)^n}{\sqrt{n}+(-1)^n} = \frac{(-1)^n}{\sqrt{n}} - u_n.$$

Comme on a montré précédemment que la série alternée $\sum_{n\geqslant 1} \frac{(-1)^n}{\sqrt{n}}$ convergeait, la série $\sum_{n\geqslant 2} u_n$ ne peut converger sans quoi d'après la proposition 1.3 la série $\sum_{n\geqslant 2} v_n$ convergerait.

Exercice 1.14 Déterminer la nature de la série $\sum_n u_n$ de terme général :

1 - $u_n = (-1)^n \sin\left(\frac{1}{n}\right)$ 2 - $u_n = (-1)^n \frac{2^n}{n!}$ 3 - $u_n = \frac{(-1)^n}{\ln(n)}$

Proposition 1.15 (Majoration du reste d'une série alternée)
Si la suite $(\alpha_n)_n$ est une suite de réels positifs, décroissante et convergeant vers 0, alors la série alternée $\sum_n (-1)^n \alpha_n$ converge et le reste à l'ordre n défini par $R_n = \sum_{k=n+1}^{+\infty} (-1)^n \alpha_n$ vérifie : $|R_n| \leqslant \alpha_{n+1}$.

Démonstration Reprenons le cadre de la démonstration du théorème 1.3. Pour tout $n \in \mathbb{N}$ on a :

$$|R_n| = |U - U_n| \leqslant |U_{n+1} - U_n|$$

car

- si n est pair on a $U_{n+1} \leqslant U \leqslant U_n$ donc $0 \leqslant U - U_{n+1} \leqslant U_n - U_{n+1}$;
- si n est impair on a $U_n \leqslant U \leqslant U_{n+1}$ donc $U_n - U_{n+1} \leqslant U - U_{n+1} \leqslant 0$.

Pour conclure, il suffit de remarquer que

$$|U_{n+1} - U_n| = |(-1)^{n+1}\alpha_{n+1}| = \alpha_{n+1}$$

car $(\alpha_n)_n$ est une suite de réels positifs. □

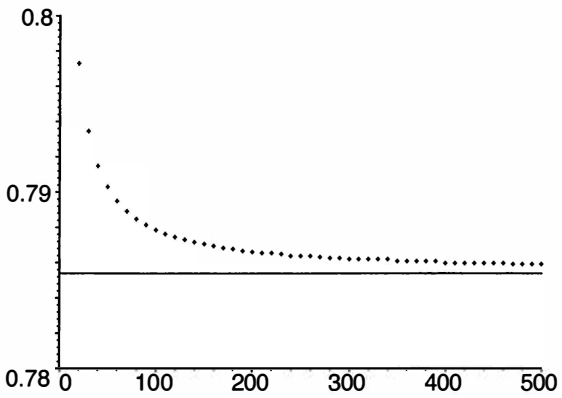
Illustration avec MAPLE

La série $\sum_n \frac{(-1)^n}{2n+1}$ est une série alternée qui, d'après le théorème 1.3, converge. On peut obtenir une valeur approchée de sa somme S en calculant $S_N = \sum_{n=0}^{N} \frac{(-1)^n}{2n+1}$. L'erreur commise $|S_N - S|$ est d'après la proposition 1.15 majorée par $\frac{1}{2N+3}$. Ainsi, si l'on souhaite obtenir une valeur approchée à 10^{-3} près, il suffit de prendre $N \geqslant 499$.

Nous montrerons ultérieurement dans le cours en ayant recours aux propriétés des séries de Fourier que la somme de cette série est $\frac{\pi}{4}$, voir p. 326.

On peut, à l'aide de MAPLE, explorer la qualité de l'approximation.

```
> SN:= N->sum((-1)^n/(2*n+1),n=0..N):
> E:=10*seq(i,i=1..50):
> F:=seq([N,evalf(SN(N))],N=E):
> D1:=plot([F],style=point): D2:=plot(Pi/4,x=0..500,y=0.78..0.80):
> with(plots):display(D1,D2);
```



On peut remarquer que la majoration de l'erreur donnée par la proposition 1.15 est sur cet exemple est largement excessive. On observe qu'une l'erreur devient inférieure à 10^{-3} dès $N = 250$ et que pour $N = 499$ l'erreur est en fait de l'ordre de $5\,10^{-4}$.

On comprendra les limites de l'approximation de la somme de série alternée selon cette approche en considérant la série $\sum_{n\geqslant 1}(-1)^n/\sqrt{n}$. Pour obtenir une valeur approchée de sa somme à 10^{-6} près en calculant $S_N = \sum_{n=1}^{N}(-1)^n/\sqrt{n}$, la valeur de N doit être d'environ 10^{12} (mille milliards).

1.4.3 Règle d'Abel

Le critère spécial pour les séries alternées correspond à un cas particulier d'un résultat plus général connu sous le nom de *règle d'Abel.*

THÉORÈME 1.4 (Règle d'Abel)
Soient $(u_n)_n$ une suite numérique et $(\alpha_n)_n$ une suite de réels positifs. Si,
1. la suite $(\alpha_n)_n$ est décroissante et converge vers 0,
2. $\exists M \in \mathbb{R}^+ \quad \forall n \in \mathbb{N} \quad |\sum_{k=0}^{n} u_k| \leqslant M$,
alors la série $\sum_n \alpha_n u_n$ converge.

Démonstration Pour montrer que la série de terme général $v_n = \alpha_n u_n$ converge, nous allons montrer qu'elle vérifie le critère de Cauchy[49]. Pour tout $n \in \mathbb{N}$, on note $U_n = \sum_{k=0}^{n} u_k$; on a donc $u_n = U_n - U_{n-1}$ pour $n \geqslant 1$ et $U_0 = u_0$. De plus, d'après les hypothèses du théorème, on a $|U_n| \leqslant M$ pour tout $n \in \mathbb{N}$. Il s'agit de montrer que

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N \in \mathbb{N} \quad \forall (m,n) \in \mathbb{N}^2 \quad \Big(n \geqslant m \geqslant N \implies \Big|\sum_{k=m}^{n} v_k\Big| \leqslant \varepsilon\Big).$$

Soient ε un réel strictement positif et m et n deux entiers tels que $n \geqslant m$. Posant $R(m,n) = \sum_{k=m}^{n} v_k$, on obtient

$$\begin{aligned} R(m,n) &= \sum_{k=m}^{n} \alpha_k u_k = \sum_{k=m}^{n} \alpha_k (U_k - U_{k-1}) = \sum_{k=m}^{n} \alpha_k U_k - \sum_{k=m}^{n} \alpha_k U_{k-1} \\ &= \sum_{k=m}^{n} \alpha_k U_k - \sum_{k=m-1}^{n-1} \alpha_{k+1} U_k = \alpha_n U_n - \alpha_m U_{m-1} + \sum_{k=m}^{n-1} (\alpha_k - \alpha_{k+1})\, U_k. \end{aligned}$$

Comme la suite $(\alpha_n)_n$ est positive et décroissante, on a $\alpha_k - \alpha_{k+1} \geqslant 0$ pour

[49] Voir le théorème 1.1 p. 9.

ABEL, Niels Henrik (1802, Stavanger - 1829, Froland).



Abel a démontré, à l'âge de 19 ans, l'impossibilité de résoudre par radicaux les équations polynomiales de degré 5, ce que son contemporain Evariste Galois généralisera à tout entier strictement supérieur à 4. Suite à ces premiers travaux, il a introduit les fonctions elliptiques et plus généralement les intégrales abéliennes. Récemment les intégrales abéliennes ont joué un rôle fondamental dans la démonstration par Andrew Wiles du théorème de Fermat résolvant par la négative un problème posé voici plus de trois cents ans : est-il possible de trouver trois nombres entiers non nuls x, y et z, tels que $x^n + y^n = z^n$ pour une puissance n supérieure à 2 ?
Le Prix Abel[50] est une récompense créé en 2002 par le gouvernement norvégien dans le but de combler l'absence d'un Prix Nobel en mathématiques. (On considérait toutefois en général que la médaille Fields, d'origine canadienne, en était l'équivalent.)

tout $k \in \{m, \dots, n-1\}$. On en déduit que

$$\begin{aligned} |R(m,n)| &\leqslant \alpha_n |U_n| + \alpha_m |U_{m-1}| + \sum_{k=m}^{n-1} (\alpha_k - \alpha_{k+1}) |U_k| \\ &\leqslant M\Big(\alpha_n + \alpha_m + \sum_{k=m}^{n-1} (\alpha_k - \alpha_{k+1})\Big) = 2M\alpha_m. \end{aligned} \tag{10}$$

Si $M = 0$ on a $R(m,n) = 0$ et le résultat est trivial. On suppose donc que $M > 0$. La suite $(\alpha_n)_n$ étant une suite à termes positifs qui tend vers 0, on a :

$$\forall \widetilde{\varepsilon} \in \mathbb{R}_+^* \quad \exists N \in \mathbb{N} \quad \forall m \in \mathbb{N} \quad (m \geqslant N \implies \alpha_m \leqslant \widetilde{\varepsilon}).$$

En particulier, pour $\widetilde{\varepsilon} = \frac{\varepsilon}{2M}$ on obtient :

$$\exists N \in \mathbb{N} \quad \forall m \in \mathbb{N} \quad \Big(m \geqslant N \implies \alpha_m \leqslant \frac{\varepsilon}{2M}\Big),$$

ce qui, compte tenu de la relation (10), implique que pour tous entiers m et n tels que $n \geqslant m \geqslant N$, on a $|R(m,n)| \leqslant 2M\alpha_m \leqslant \varepsilon$. La série de terme général $v_n = \alpha_n u_n$ vérifie bien le critère de Cauchy. Par conséquent elle converge. □

1.4.4 Séries trigonométriques

DÉFINITION 1.10 *On appelle* série trigonométrique *toute série dont le terme général est de la forme* $\alpha_n e^{int}$ *où* $(\alpha_n)_n$ *est une suite réelle et* t *un réel.*

[50] Site officiel : `http ://www.abelprisen.no`

On remarquera que si la série $\sum_n \alpha_n$ converge absolument alors la série trigonométrique $\sum_n \alpha_n e^{int}$ converge elle aussi absolument pour tout $t \in \mathbb{R}$ car $|\alpha_n e^{int}| = |\alpha_n|$. Le théorème suivant donne une condition suffisante sur la suite $(\alpha_n)_n$ permettant de s'assurer de la convergence de la série trigonométrique $\sum_n \alpha_n e^{int}$. D'un point de vue pratique, il conviendra d'utiliser ce théorème dans le cas où la convergence de la série trigonométrique ne pourra pas être établie par la convergence absolue de la série $\sum_n \alpha_n$.

THÉORÈME 1.5 *Si $(\alpha_n)_n$ est une suite réelle à termes positifs, décroissante et convergeant vers 0, alors pour tout $t \in \mathbb{R} \setminus 2\pi\mathbb{Z}$ la série trigonométrique $\sum_n \alpha_n e^{int}$ est convergente.*

Démonstration On utilise la règle d'Abel. La suite $(\alpha_n)_n$ étant par hypothèse une suite à termes positifs, décroissante et convergeant vers 0, il faut montrer que pour tout $t \in \mathbb{R} \setminus 2\pi\mathbb{Z}$ la suite de terme général $u_n = e^{int}$ vérifie :

$$\exists M \in \mathbb{R}^+ \quad \forall n \in \mathbb{N} \quad \left|\sum_{k=0}^{n} u_k\right| \leqslant M. \tag{11}$$

La suite $(u_n)_n$ est une suite géométrique de raison $r = e^{it}$. Pour tout $t \in \mathbb{R} \setminus 2\pi\mathbb{Z}$ on a $r \neq 1$ et par conséquent

$$\left|\sum_{k=0}^{n} u_k\right| = \left|\frac{1 - e^{i(n+1)t}}{1 - e^{it}}\right| = \frac{\sqrt{1 - \cos((n+1)t)}}{\sqrt{1 - \cos(t)}} \leqslant \frac{\sqrt{2}}{\sqrt{1 - \cos(t)}}.$$

L'assertion (11) est établie avec $M = \sqrt{\frac{2}{1-\cos(t)}}$. □

Exemple Considérons la série $\sum_n u_n$ de terme général $u_n = \frac{\sin(n)}{\sqrt{n}}$. Pour tout $n \in \mathbb{N}^*$, on a $u_n = \mathcal{I}\mathrm{m}(\alpha_n e^{in})$ où $\alpha_n = 1/\sqrt{n}$. La suite $(\alpha_n)_n$ est une suite réelle à termes positifs, décroissante et convergeant vers 0 et $t = 1 \notin 2\pi\mathbb{Z}$. Le théorème 1.5 indique que la série $\sum_n \alpha_n e^{in}$ converge. On en déduit que la série $\sum_{n \geqslant 1} \frac{\sin(n)}{\sqrt{n}}$ converge aussi.

1.4.5 Séries semi-convergentes ne relevant pas de la règle d'Abel

Lorsque la série $\sum_n u_n$ n'est pas absolument convergente et que les critères liés à la règle d'Abel ne s'appliquent pas, on peut essayer de déterminer un développement asymptotique du terme général u_n pour n au voisinage de $+\infty$. Si on peut déterminer la nature des séries dont le terme général intervient dans le développement asymptotique[(51)], on pourra conclure en ayant recours à la proposition 1.3 et au corollaire 1.3. Nous allons illustrer cette méthode nommée « méthode par éclatement de termes » sur deux exemples.

[(51)] Voir le *Cours de première année*, section 17.4.3 p. 797 et section 17.6 p. 809.

Exemples

1. Déterminons la nature de la série de terme général $u_n = \frac{1}{1+(-1)^n n}$ avec $n \geqslant 2$. Remarquons tout d'abord qu'il ne s'agit pas d'une série à termes positifs puisque si n est pair, u_n est positif alors que si n est impair, u_n est négatif. De plus cette série ne converge pas absolument car $|u_n| \underset{+\infty}{\sim} \frac{1}{n}$ et la série harmonique $\sum_{n\geqslant 1} \frac{1}{n}$ diverge. Pour tout entier $n \geqslant 2$, on obtient en utilisant le développement limité $\frac{1}{1+u} = 1 - u + o_0(u)$,

$$u_n = \frac{1}{1+(-1)^n n} = \frac{(-1)^n}{n} \frac{1}{1+\frac{(-1)^n}{n}} = \frac{(-1)^n}{n} - \frac{1}{n^2} + R_n$$

où $R_n = o_{+\infty}\left(\frac{1}{n^2}\right)$. Comme la série $\sum_{n\geqslant 1} \frac{1}{n^2}$ converge, d'après le corollaire 1.3, le « reste » R_n est le terme général d'une série absolument convergente. De plus, la série harmonique alternée $\sum_{n\geqslant 1} \frac{(-1)^n}{n}$ convergeant, on peut en déduire que la série $\sum_{n\geqslant 2} u_n$ converge en tant que somme de trois séries convergentes, voir la proposition 1.3.

2. Déterminons la nature de la série de terme général $u_n = \ln\left(1 + \frac{(-1)^n}{\sqrt{n}}\right)$. Remarquons tout d'abord qu'il ne s'agit pas d'une série à termes positifs : si n est pair, u_n est positif alors que si n est impair, u_n est négatif. En utilisant le développement limité $\ln(1+x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 + o_0(x^3)$, on obtient :

$$u_n = \ln\left(1 + \frac{(-1)^n}{\sqrt{n}}\right) = \frac{(-1)^n}{\sqrt{n}} - \frac{1}{2n} + \frac{(-1)^{3n}}{3n\sqrt{n}} + R_n$$

où le reste R_n vérifie $R_n = o_{+\infty}\left(n^{-3/2}\right)$. Nous avons vu que la série $\sum_n \frac{(-1)^n}{\sqrt{n}}$ était simplement convergente et que la série $\sum_n \frac{1}{2n}$ était divergente. Par ailleurs, la série $\sum_n \frac{(-1)^{3n}}{3n\sqrt{n}}$ est absolument convergente car

$$\left|\frac{(-1)^{3n}}{3n\sqrt{n}}\right| = \frac{1}{3}\frac{1}{n^{3/2}}$$

et $\sum_n \frac{1}{n^{3/2}}$ est une série de Riemann convergente. Enfin, pour conclure remarquons que $|R_n| = o_{+\infty}\left(n^{-3/2}\right)$ et par conséquent, d'après le corollaire 1.3, la série $\sum_n R_n$ est absolument convergente. La série $\sum_n u_n$ est donc divergente : elle se décompose en la somme de trois séries convergentes et d'une série divergente.

1.5 Groupement de termes et changement de l'ordre des termes d'une série

Intéressons-nous à présent à la question suivante : est-il possible, sans changer la nature ni la somme d'une série, de regrouper et sommer des termes consécutifs de la série ou de permuter l'ordre des termes de la série ? On peut se convaincre

sur des exemples que ces manipulations sur les termes d'une série peuvent modifier la nature de la série et que des hypothèses sur la série sont nécessaires pour que la nature et la somme de la série restent inchangées.

Considérons la série $\sum_n u_n$ de terme général $u_n = (-1)^{E(\frac{n}{4})} \ln(n)$ qui diverge grossièrement et effectuons un regroupement de quatre termes consécutifs de la série :

$$\underbrace{\ln(2)\ln(3) - \ln(4) - \ln(5)}_{=v_0}, \quad \underbrace{-\ln(6) - \ln(7)\ln(8)\ln(9)}_{=v_1},$$
$$\underbrace{\ln(9)\ln(10) - \ln(11) - \ln(12)}_{=v_2}, \ldots$$

On définit ainsi une série $\sum_n v_n$ de terme général

$$v_n = u_{4n+2} + u_{4n+3} + u_{4n+4} + u_{4n+5} = (-1)^{n+1} \ln\left(\frac{(4n+4)(4n+5)}{(4n+2)(4n+3)}\right)$$

dont on a établi la convergence p. 43.

Considérons par ailleurs la série $\sum_n u_n$ de terme général $u_n = \frac{(-1)^n}{\sqrt{n}}$ qui est une série convergente. Si l'on permute l'ordre des termes de sorte d'avoir tous les termes d'indice pair puis tous les termes d'indice impair. La série ainsi construite est divergente car $\sum_{n=1}^{+\infty} u_{2n} = \sum_{n=1}^{+\infty} \frac{1}{\sqrt{2n}} = +\infty$ et $\sum_{n=1}^{+\infty} u_{2n-1} = \sum_{n=1}^{+\infty} \frac{-1}{\sqrt{2n-1}} = -\infty$.

Les résultats qui suivent vont fournir des conditions permettant de regrouper des termes consécutifs ou permuter l'ordre des termes d'une série sans en changer ni la nature ni la somme.

PROPOSITION 1.16 (Groupement de termes)
Soient $\sum_n u_n$ une série numérique, φ une application de $\mathbb{N}$ dans $\mathbb{N}$ strictement croissante et $\sum_n v_n$ la série numérique définie par :

$$v_0 = \sum_{k=0}^{\varphi(0)} u_k, \quad \text{et} \quad v_n = \sum_{k=\varphi(n-1)+1}^{\varphi(n)} u_k \quad \text{pour tout } n \geqslant 1.$$

- *Si la série $\sum_n u_n$ converge alors la série $\sum_n v_n$ converge et ces deux séries ont même somme.*
- *Si la série $\sum_n u_n$ est à termes positifs et si la série $\sum_n v_n$ converge alors la série $\sum_n u_n$ converge et ces deux séries ont même somme.*

Démonstration $\trianglerighteq$ Soient $(U_n)_n$ la suite des sommes partielles associée à la série $\sum_n u_n$ et $(V_n)_n$ celle associée à la série $\sum_n v_n$. Pour tout $n \in \mathbb{N}$, on a

$$V_n = \sum_{k=0}^{n} v_k = \sum_{k=0}^{\varphi(n)} u_k = U_{\varphi(n)};$$

la suite $(V_n)_n$ est donc une suite extraite de la suite $(U_n)_n$. La série $\sum_n u_n$ converge et admet pour somme μ si et seulement si la suite $(U_n)_n$ converge vers μ. Dans ce cas, la suite $(V_n)_n$ extraite de la suite $(U_n)_n$ converge[52] aussi vers μ. La série $\sum_n v_n$ est donc une série convergente de somme μ.

$\trianglerighteq$ Supposons que la série $\sum_n u_n$ est à termes positifs et que la série $\sum_n v_n$ converge et admet pour somme μ. Comme $u_n \geqslant 0$ pour tout $n \in \mathbb{N}$, les suites $(U_n)_n$ et $(V_n)_n$ sont croissantes. On a donc $V_n \leqslant \mu$ pour tout $n \in \mathbb{N}$ et par ailleurs, puisque φ est une application de $\mathbb{N}$ dans $\mathbb{N}$ strictement croissante : $\varphi(n) \geqslant n$. Cela implique que :

$$\forall n \in \mathbb{N} \qquad U_n \leqslant U_{\varphi(n)} = V_n \leqslant \mu.$$

La suite $(U_n)_n$ est par conséquent une suite croissante et majorée ; il s'agit donc d'une suite convergente[53]. Comme la suite $(V_n)_n$ qui est extraite de la suite $(U_n)_n$ converge vers μ, la suite $(U_n)_n$ admet pour limite μ. □

Exemple En considérant l'application $\phi : n \in \mathbb{N} \mapsto 2n+1$, la série $\sum_n v_n$ définie à la proposition 1.16 correspond à des groupements de 2 termes de la série $\sum_n u_n$: pour tout $n \in \mathbb{N}$, $v_n = u_{2n} + u_{2n+1}$. Plus généralement, pour $\eta \in \mathbb{N}^*$ fixé, l'application $\phi : n \in \mathbb{N} \mapsto \eta n + (\eta - 1)$ permet d'effectuer des groupements de η termes.

En regroupant convenablement les termes de certaines séries divergentes, on peut obtenir une série numérique $\sum_n v_n$ qui est convergente.

Plus précisément, la démonstration de la proposition 1.16 indique que si $\sum_n u_n$ est une série convergente alors quel que soit le regroupement de termes qui est opéré, la série obtenue est convergente (et de même somme que $\sum_n u_n$). Par contre, une série obtenue par regroupement de termes peut converger alors que la série initiale divergeait ; cela ne peut avoir lieu que pour des séries qui ne sont pas à termes positifs (ou plus généralement, qui ne sont pas de signe constant à partir d'un certain rang).

Le programme MAPLE suivant permet d'explorer le regroupement de termes. La variable T représente le nombre de terme regroupés, u définit la série considérée et v désigne la série issue de u par le regroupement de termes.

```
> T:=3: phi:= n-> T*n+1;
> u:= n-> (-1)^n/n:  n0:=1:
> v:= n-> simplify(sum(u(k),k=phi(n-1)..phi(n)-1),assume=integer):
> sum(u(n),n=n0..+infinity);evalf(%);
```

$$-\ln(2)$$
$$-.6931471806$$

[52] Voir la proposition 5.18 p. 187 du *Cours de première année*.

[53] Voir la proposition 5.16 p. 183 du *Cours de première année*.

```
> sum(v(n),n=n0..+infinity):evalf(%);
```

$$-.6931471821 - .288055614110^{-9}\, I$$

On notera que l'expression formelle de la somme de la série après regroupement de terme obtenue par MAPLE est inexploitable. La valeur numérique de cette somme évaluée par MAPLE diffère légèrement de la valeur de la somme de la série $\sum_n u_n$ compte tenu d'erreurs d'arrondis dans l'évaluation numérique de l'expression formelle.

L'instruction suivante :

```
> seq([seq(k,k=phi(n-1)..phi(n)-1)],n=n0..5);
```

permet par ailleurs de visualiser les groupements de termes qui sont effectués.

```
[1, 2, 3], [4, 5, 6], [7, 8, 9], [10, 11, 12], [13, 14, 15]
```

PROPOSITION 1.17 (Changement de l'ordre des termes)
Soient $\sum_n u_n$ une série absolument convergente et σ une bijection de $\mathbb{N}$ dans $\mathbb{N}$ (on parle aussi de permutation de $\mathbb{N}$). La série $\sum_n u_{\sigma(n)}$ est une série absolument convergente qui a même somme que la série $\sum_n u_n$.

Démonstration Montrons que les séries $\sum_n u_{\sigma(n)}$ et $\sum_n u_n$ sont de même nature ; nous admettrons que ces séries ont même somme[(54)].

Soient $(S_n)_n$ la suite des sommes partielles associée à la série $\sum_n |u_n|$ et $(T_n)_n$ celle associée à la série $\sum_n |u_{\sigma(n)}|$. Les suites $(S_n)_n$ et $(T_n)_n$ sont des suites croissantes et à termes positifs. Pour tout $n \in \mathbb{N}$, on a

$$T_n = \sum_{k=0}^{n} |u_{\sigma(k)}| \leqslant \sum_{k=0}^{m_n} |u_k| = S_{m_n},$$

où $m_n = \sup_{k\in\{0,\ldots,n\}} \sigma(k)$. Puisque la série $\sum_n u_n$ est absolument convergente, la série $\sum_n |u_n|$ est convergente et par conséquent la suite $(S_n)_n$ converge vers un réel positif S. Cette suite $(S_n)_n$ étant croissante, elle est majorée par S. On en déduit que pour tout $n \in \mathbb{N}$ on a $T_n \leqslant S$. La suite $(T_n)_n$ est donc une suite croissante et majorée ; il s'agit par conséquent d'une suite convergente[(55)]. On en conclut que la série $\sum_n |u_{\sigma(n)}|$ est une série convergente et par conséquent que la série $\sum_n u_{\sigma(n)}$ est une série absolument convergente. □

Le programme MAPLE suivant permet d'explorer l'effet d'un ré-arrangement des termes d'une série alternée ($\sum_n u_n$ avec $u_n = (-1)^n \alpha_n$ où $(\alpha_n)_n$ est une suite de réels positifs) sur la somme de cette série. On réarrange les termes de la manière suivante : un terme négatif de la série est suivi des T termes positifs suivants.

---

(54) Pour une preuve de cette propriété, on pourra consulter J. LELONG-FERRAND, J.M. ARNAUDIÈS, *Cours de Mathématiques*, tome 2, p. 281.

(55) Voir la proposition 5.16 p. 183 du *Cours de première année*.

```
> alpha:= n-> [à définir]: u:= n-> (-1)^n*alpha(n):
> T:=3:
> sum(u(n),n=1..+infinity);
> v:= n -> -alpha(2*n+1) + sum(alpha(2*(T*n+1)+2*k),k=0..T-1):
> sum(v(n),n=0..+infinity);
```

Si on considère la série alternée $\sum_{n\geqslant 1}\frac{(-1)^n}{n^2}$, on peut observer que la valeur de la somme reste identique quelle que soit la valeur de T (pour laquelle MAPLE parvient à calculer la somme...).

Par contre, si on prend la série harmonique alternée $\sum_{n\geqslant 1}\frac{(-1)^n}{n}$ et $T=2$, les premiers termes de la série réarrangée sont :

```
> v:= n -> -alpha(2*n+1),seq(alpha(2*(T*n+1)+2*k),k=0..T-1):
> seq(v(n),n=0..5);
```

$$-1,\frac{1}{2},\frac{1}{4},-\frac{1}{3},\frac{1}{6},\frac{1}{8},-\frac{1}{5},\frac{1}{10},\frac{1}{12},-\frac{1}{7},\frac{1}{14},\frac{1}{16},-\frac{1}{9},\frac{1}{18},\frac{1}{20},-\frac{1}{11},\frac{1}{22},\frac{1}{24}$$

et la somme de la série $\sum_n v_n$ réarrangée est $-\frac{1}{2}\ln(2)$.
En choisissant $T=3$ on obtient pour somme $-\ln(2)+\frac{1}{2}\ln(3)$ et en choisissant $T=4$, la somme est nulle.

En changeant l'ordre des termes d'une série convergente qui n'est pas absolument convergente, on change en général la valeur de la somme de cette série. Le cas de la série harmonique alternée est intéressante à ce point : on peut montrer que quel que soit le réel que l'on se fixe, il existe un réarrangement de ses termes pour lequel la somme de la série réarrangée est égal à ce réel.

1.6 Exercices de synthèse

Exercice 1.15 ***Calcul de*** $\zeta(2)=\sum_{n=1}^{+\infty}1/n^2$

1 - On considère le polynôme de $\mathbb{R}[X]$ de degré n ($n\in\mathbb{N}^*$) défini par[56] :

$$P_n=\sum_{k=0}^{n}(-1)^k C_{2n+1}^{2k+1}\,X^{n-k}.$$

a) En utilisant la formule de Moivre[57], montrer que pour tout $\theta\in]0,\pi/2[$ on a :

$$\sin((2n+1)\theta)=\sum_{k=0}^{n}C_{2n+1}^{2k+1}(-1)^k\sin^{2k+1}(\theta)\cos^{2(n-k)}(\theta).$$

En déduire que : $\sin((2n+1)\theta)=P_n(\operatorname{cotan}^2(\theta))\times\sin^{2n+1}(\theta)$.

[56] On convient que $(-1)^0=1$.

[57] Voir le *Cours de première année* p. 138.

b) Pour $k \in \{1, \ldots, n\}$, on note $\theta_k = \frac{k\pi}{2n+1}$.
Montrer que pour tout $k \in \{1, \ldots, n\}$ le réel $x_k = \operatorname{cotan}^2(\theta_k)$ est racine de P_n.

c) Montrer que le polynôme P_n admet n racines réelles simples. En déduire une décomposition du polynôme P_n en produit de facteurs irréductibles dans $\mathbb{R}[X]$.

2 - a) Montrer que :

$$\sum_{k=1}^{n} \operatorname{cotan}^2(\theta_k) = \frac{n}{3}(2n-1) \quad \text{et que} \quad \sum_{k=1}^{n} \sin^{-2}(\theta_k) = \frac{2}{3}n(n+1).$$

b) Montrer que pour tout $x \in]0, \pi/2[$ on a : $\dfrac{1}{\tan^2(x)} < \dfrac{1}{x^2} < \dfrac{1}{\sin^2(x)}$.

c) En déduire que :

$$\frac{\pi^2 n(2n-1)}{3(2n+1)^2} < \sum_{k=1}^{n} \frac{1}{k^2} < \frac{2\pi^2 n(n+1)}{3(2n+1)^2}.$$

3 - Conclure que la série $\sum_n 1/n^2$ est convergente et donner la valeur de $\zeta(2) = \sum_{n=1}^{+\infty} 1/n^2$.

Exercice 1.16 ***Nature de la série*** $\sum_n \sin(\pi e\, n!)$

1 - Soient a un réel non nul et $(b_n)_n$ une suite bornée. Déterminer, en le justifiant, la nature de la série de terme général $u_n = (-1)^n \sin\left(\dfrac{a}{n} + \dfrac{b_n}{n(n+1)}\right)$.

2 - a) Montrer que pour tout $n \in \mathbb{N}$ avec $n \geqslant 2$ le réel $p = \sum_{k=0}^{n-2} n!/k!$ est un entier naturel pair.

b) En utilisant la formule de Taylor-Lagrange, montrer que pour tout $n \in \mathbb{N}$ avec $n \geqslant 2$, il existe $c_n \in]0, 1[$ et $q_n \in \mathbb{N}$ tel que :

$$e n! = 2q_n + (n+1) + \frac{1}{n+1} + \frac{\exp(c_n)}{(n+1)(n+2)}.$$

c) Montrer que la série $\sum_n \sin(\pi e\, n!)$ est une série numérique semi-convergente.

1.7 Solutions des exercices

Solution de l'exercice 1.1

La série $\sum_n v_n$ converge si et seulement si la suite $(V_n)_n$ des sommes partielles qui lui est associée converge. Or, pour tout $n \in \mathbb{N}$ on a :

$$V_n = \sum_{k=0}^{n} v_k = \sum_{k=0}^{n} (u_{k+1} - u_k) = \sum_{k=0}^{n} u_{k+1} - \sum_{k=0}^{n} u_k = u_{n+1} - u_0. \qquad (12)$$

On en déduit que la suite $(V_n)_n$ converge si et seulement si la suite $(u_n)_n$ converge. Ainsi, la série $\sum_n v_n$ converge si et seulement si la suite $(u_n)_n$ converge. De plus, la relation (12) indique que la somme V de la série $\sum_n v_n$ et la limite ℓ de la suite $(u_n)_n$ sont liées par la relation $V = \ell + u_0$.

Solution de l'exercice 1.2

1 - a) La suite extraite correspondant aux termes d'indices pairs a pour terme général $u_n = S_{2n} = \sum_{k=1}^{2n} (-1)^k \frac{1}{k}$, et la suite extraite correspondant aux termes d'indices impairs a pour terme général $v_n = S_{2n+1} = \sum_{k=1}^{2n+1} (-1)^k \frac{1}{k}$. Pour tout $n \in \mathbb{N}$, on a :

$$u_{n+1} - u_n = (-1)^{2n+1} \frac{1}{2n+1} + (-1)^{2n+2} \frac{1}{2n+2} = -\frac{1}{2n+1} + \frac{1}{2n+2} < 0,$$

$$v_{n+1} - v_n = (-1)^{2n+2} \frac{1}{2n+2} + (-1)^{2n+3} \frac{1}{2n+3} = \frac{1}{2n+2} - \frac{1}{2n+3} > 0.$$

On en déduit que la suite $(u_n)_n$ est décroissante et que la suite $(v_n)_n$ est croissante. Par ailleurs, pour tout $n \in \mathbb{N}$, on a :

$$v_n - u_n = S_{2n+1} - S_{2n} = -\frac{1}{2n+1},$$

donc $\lim_{n \to +\infty} (v_n - u_n) = 0$. Les deux suites $(u_n)_n$ et $(v_n)_n$ sont adjacentes[58].

b) Puisque les deux suites $(u_n)_n$ et $(v_n)_n$ sont adjacentes, elles convergent[59] vers une même limite ℓ. La suite $(S_n)_n$ admet une suite extraite correspondant aux termes d'indices pairs et une suite extraite correspondant aux termes d'indices impairs qui convergent vers une même limite. On en déduit[60] que la suite $(S_n)_n$ converge. Par définition, ceci signifie que la série $\sum_{n \geqslant 1} (-1)^n \frac{1}{n}$ converge.

2 - a) On a $f = g \circ h$ où $h : x \in [0,1] \mapsto x + 1 \in [1,2]$ et $g : x \in [1,2] \mapsto \ln(x)$ sont deux applications de classe $\mathcal{C}^\infty$. On en déduit que f est une application de classe $\mathcal{C}^\infty$. Il est aisé de vérifier par récurrence que, pour tout $n \in \mathbb{N}^*$ et pour tout $x \in [0,1]$, on a

$$f^{(n)}(x) = \frac{(-1)^{n+1}(n-1)!}{(1+x)^n}.$$

---

[58] Voir la définition 5.6 p. 185 du *Cours de première année*.

[59] Voir le théorème 5.2 p. 185 du *Cours de première année*.

[60] Voir la proposition 5.19 p. 189 du *Cours de première année*.

b) Puisque f est de classe $\mathcal{C}^\infty$, on peut appliquer la formule de Taylor-Lagrange[61] entre 0 et 1 à l'ordre n, pour tout $n \in \mathbb{N}$. Il existe un réel $c \in]0,1[$ tel que :

$$\begin{aligned} f(1) &= \sum_{k=1}^{n} \frac{f^{(k)}(0)}{k!} + \frac{f^{(n+1)}(c)}{(n+1)!} = \sum_{k=1}^{n} \frac{(-1)^{k+1}(k-1)!}{k!} + \frac{(-1)^n}{(n+1)(1+c)^{n+1}} \\ &= -S_n + \frac{(-1)^n}{(n+1)(1+c)^{n+1}}. \end{aligned}$$

On en déduit que pour tout $n \in \mathbb{N}^*$,

$$|f(1) + S_n| = \left| \frac{1}{(n+1)(1+c)^{n+1}} \right| \leqslant \frac{1}{n+1}$$

car $1 + c > 1$. Il résulte du théorème d'encadrement[62] que la suite $(S_n)_n$ converge vers $-f(1)$, c'est-à-dire vers $-\ln(2)$. La somme de la série $\sum_{n \geqslant 1} (-1)^n \frac{1}{n}$ est donc $\sum_{n=1}^{+\infty} (-1)^n \frac{1}{n} = -\ln(2)$.

Solution de l'exercice 1.3

1 - On a la décomposition en éléments simples[63] suivante dans $\mathbb{R}(X)$:

$$\frac{1}{X(X+1)} = \frac{1}{X} - \frac{1}{X+1}.$$

Le terme général U_n de la suite des sommes partielles associé à la série $\sum_n \frac{1}{n(n+1)}$ s'écrit donc :

$$U_n = \sum_{k=1}^{n} u_k = \sum_{k=1}^{n} \left(\frac{1}{k} - \frac{1}{k+1} \right) = \sum_{k=1}^{n} \frac{1}{k} - \sum_{k=2}^{n+1} \frac{1}{k} = 1 - \frac{1}{n+1}.$$

La suite des sommes partielles $(U_n)_n$ converge vers 1 et par conséquent la série $\sum_n \frac{1}{n(n+1)}$ est convergente, de somme 1.

2 - La formule de Taylor-Lagrange[64] appliquée à l'ordre n à la fonction exponentielle entre -1 et 0 donne :

$$\exists c_n \in]-1,0[\qquad \mathrm{e}^{-1} = \sum_{k=0}^{n} \frac{(-1)^k}{k!} + \frac{(-1)^{n+1}}{(n+1)!} \exp(c_n).$$

Le terme général U_n de la suite des sommes partielles associée à la série $\sum_n (-1)^n/n!$ vérifie donc :

$$U_n = \sum_{k=0}^{n} \frac{(-1)^k}{k!} = \mathrm{e}^{-1} - \frac{(-1)^{n+1}}{(n+1)!} \exp(c_n).$$

---

[61] Voir le théorème 16.4 p. 750 du *Cours de première année*.

[62] Voir le théorème 5.1 p. 179 du *Cours de première année*.

[63] Voir le chap. 9 du *Cours de première année*.

[64] Voir le *Cours de première année*, théorème 16.4 p. 750.

Puisque $c_n \in]-1,0[$, on a $\mathrm{e}^{-1} < \exp(c_n) < 1$ et

$$|U_n - \mathrm{e}^{-1}| = \frac{1}{(n+1)!}\,\exp(c_n) < \frac{1}{(n+1)!}.$$

Par le théorème d'encadrement, on en déduit que la suite des sommes partielles $(U_n)_n$ converge vers e^{-1} et par conséquent que la série $\sum_n (-1)^n/n!$ converge et admet pour somme $1/\mathrm{e}$.

3 - Le terme général U_n de la suite des sommes partielles associée à la série $\sum_n \frac{\mathrm{e}^{-n}}{n}$ est :

$$\begin{aligned} U_n &= \sum_{k=1}^{n} \frac{\mathrm{e}^{-k}}{k} = -\sum_{k=1}^{n} \int_1^{+\infty} \mathrm{e}^{-kx}\,\mathrm{d}x = -\int_1^{+\infty} \sum_{k=1}^{n} \mathrm{e}^{-kx}\,\mathrm{d}x \\ &= -\int_1^{+\infty} \mathrm{e}^{-x}\,\frac{1-\mathrm{e}^{-nx}}{1-\mathrm{e}^{-x}}\,\mathrm{d}x = \int_1^{+\infty} \frac{\mathrm{e}^{-nx}-1}{\mathrm{e}^{x}-1}\,\mathrm{d}x. \end{aligned}$$

En effectuant le changement de variable $t = \mathrm{e}^{-x}$, on obtient :

$$U_n = \int_{\frac{1}{\mathrm{e}}}^{0} \frac{t^n - 1}{t-1}\,\mathrm{d}t = \int_0^{\frac{1}{\mathrm{e}}} \frac{1}{t-1}\,\mathrm{d}t - \int_0^{\frac{1}{\mathrm{e}}} \frac{t^n}{t-1}\,\mathrm{d}t.$$

On a

$$\int_0^{\frac{1}{\mathrm{e}}} \frac{1}{1-t}\,\mathrm{d}t = -\Big[\ln(1-t)\Big]_0^{\frac{1}{\mathrm{e}}} = -\ln(1-\tfrac{1}{\mathrm{e}}) = \ln\left(\frac{\mathrm{e}}{\mathrm{e}-1}\right).$$

et

$$R_n = \left|\int_0^{\frac{1}{\mathrm{e}}} \frac{t^n}{1-t}\,\mathrm{d}t\right| \leqslant \frac{1}{1-\frac{1}{\mathrm{e}}} \int_0^{\frac{1}{\mathrm{e}}} t^n\,\mathrm{d}t = \frac{\mathrm{e}}{(n+1)(\mathrm{e}-1)\,\mathrm{e}^{n+1}}.$$

Ainsi, $U_n = \ln\left(\frac{\mathrm{e}}{\mathrm{e}-1}\right) + R_n$ et le théorème d'encadrement indique que la suite $(R_n)_n$ converge vers 0. Par conséquent la suite $(U_n)_n$ converge vers $-\ln(1-\frac{1}{\mathrm{e}})$ et la série $\sum_{n\geqslant 1} \frac{\mathrm{e}^{-n}}{n}$ est une série convergente de somme $-\ln(1-\frac{1}{\mathrm{e}})$.

On pourra remarquer que MAPLE fournit l'expression de S_n à l'aide d'une fonction spéciale : la fonction Lerch Phi.

```
> sum( exp(-n)/n,n=1..N) ;
```

$$-\ln\left(1-e^{-1}\right) - e^{-N}\left(LerchPhi\left(e^{-1},1,N\right) - N^{-1}\right)$$

SOLUTION DE L'EXERCICE 1.4

1 - Pour tout $n \in \mathbb{N}$, on a $u_n + v_n - 2\sqrt{u_n \times v_n} = \left(\sqrt{u_n} - \sqrt{v_n}\right)^2 \geqslant 0$, d'où le résultat. Si les séries $\sum_n u_n$ et $\sum_n v_n$ convergent alors, d'après la proposition 1.3, la série de terme général $\frac{1}{2}(u_n + v_n)$ converge aussi. On déduit du théorème 1.2 que la série $\sum_n \sqrt{u_n \times v_n}$ converge.

2 - Comme $\sum_n u_n$ est une série à termes positifs convergente, $\lim_{n\to+\infty} u_n = 0$ et à partir d'un certain rang N, tous les termes de la suite $(u_n)_n$ sont compris entre 0 et 1. On en déduit que pour $n \geqslant N$ on a : $0 \leqslant u_n \times v_n \leqslant v_n$. Comme la série à termes positifs $\sum_n v_n$ converge, le théorème 1.2 permet de conclure que la série $\sum_n u_n \times v_n$ converge.

Solution de l'exercice 1.5

Comme $\sum_n u_n$ et $\sum_n v_n$ sont deux séries à termes positifs avec $u_n \underset{+\infty}{\sim} v_n$, la convergence de la série $\sum_n v_n$ implique la convergence de la série $\sum_n u_n$. Par conséquent les suites $(R_n)_n$ et $(R'_n)_n$ sont des suites à termes positifs, décroissantes, et convergeant vers 0.

1 - Puisque $u_n \underset{+\infty}{\sim} v_n$, on a $\lim_{n\to+\infty}(u_n - v_n)/v_n = \lim_{n\to+\infty}\left(\frac{u_n}{v_n} - 1\right) = 0$. En utilisant la définition de la limite d'une suite[65], on obtient

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N_2 \in \mathbb{N} \quad \forall k \in \mathbb{N} \quad \left(k \geqslant N \implies \left|\frac{u_k - v_k}{v_k}\right| \leqslant \varepsilon\right).$$

Comme la suite $(v_n)_n$ est à termes strictement positifs (à partir d'un certain rang N_1), on a pour $k \geqslant N = \max(N_1, N_2)$,

$$\left|\frac{u_k - v_k}{v_k}\right| \leqslant \varepsilon \implies |u_k - v_k| \leqslant \varepsilon\, v_k.$$

2 - Soit ε un réel strictement positif. Pour tout $(n,p) \in \mathbb{N}^2$ avec $p > n \geqslant N$ on a

$$|R'_n - R_n| = \left|\sum_{k=n+1}^{p} (v_k - u_k)\right| \leqslant \sum_{k=n+1}^{p} |v_k - u_k| \leqslant \sum_{k=n+1}^{p} \varepsilon\, v_k = \varepsilon \sum_{k=n+1}^{p} v_k.$$

Considérant la limite quand p tend vers $+\infty$, on obtient pour tout $n \geqslant N$

$$|R'_n - R_n| \leqslant \varepsilon \sum_{k=n+1}^{+\infty} v_k = \varepsilon R'_n.$$

3 - On a montré l'assertion suivante :

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad (n \geqslant N \implies |R'_n - R_n| \leqslant \varepsilon\, R'_n).$$

Comme $|R'_n - R_n| \leqslant \varepsilon R'_n$ équivaut à $|R_n/R'_n - 1| \leqslant \varepsilon$, cette assertion indique que $\lim_{n\to+\infty} R_n/R'_n = 1$. On a donc bien $R_n \underset{+\infty}{\sim} R'_n$.

Solution de l'exercice 1.6

1 - a) Pour tout entier $k \geqslant N$, on a d'après les hypothèses

$$0 \leqslant \frac{u_{k+1}}{u_{k+1}} \leqslant \frac{v_{k+1}}{v_k}.$$

[65] Voir la définition 5.1 p. 168 du *Cours de première année*.

On en déduit que pour tout entier $n \geqslant N$

$$\frac{u_n}{u_N} = \prod_{k=n}^{N-1} \frac{u_{k+1}}{u_k} \leqslant \prod_{k=n}^{N-1} \frac{v_{k+1}}{v_k} = \frac{v_n}{v_N}.$$

On a donc bien $u_n \leqslant \dfrac{u_N}{v_N} \times v_n$ pour tout entier $n \geqslant N$.

b) En utilisant le théorème 1.2, on peut affirmer que si la série $\sum_n v_n$ converge alors la série $\sum_n u_n$ converge aussi.

2 - Toujours d'après le théorème 1.2, la relation $u_n \leqslant \dfrac{u_N}{v_N} \times v_n$ pour tout entier $n \geqslant N$ permet d'affirmer que si la série $\sum_n v_n$ diverge alors la série $\sum_n u_n$ diverge aussi.

Solution de l'exercice 1.7

1 - a) Appliquons le théorème des accroissements finis à l'application $f : x \in \mathbb{R}^+ \mapsto \ln(1+x)$ entre 0 et x pour $x \in \mathbb{R}_+^*$ fixé. Il existe un réel $c \in]0, x[$ tel que $\ln(1+x) = \frac{x}{1+c}$. Comme $x > 0$ et $0 < c < x$ on a $\frac{x}{1+x} < \frac{x}{1+c} < x$ et par conséquent : $\frac{x}{1+x} < \ln(1+x) < x$.
Considérons la relation précédente avec des réels $x = \frac{1}{n}$ où $n \in \mathbb{N}^*$; on obtient,

$$\frac{1}{n+1} = \frac{\frac{1}{n}}{1+\frac{1}{n}} < \ln\left(1+\frac{1}{n}\right) < \frac{1}{n}.$$

Pour tout $n \in \mathbb{N}^*$, on a donc $u_n = \gamma_{n+1} - \gamma_n = \frac{1}{n+1} - \ln\left(1+\frac{1}{n}\right) < 0$.
On en conclut que la série $\sum_n u_n$ est à termes négatifs et que la suite $(\gamma_n)_n$ est décroissante.

b) Comme $\ln(1+u) = u - \frac{1}{2}u^2 + o_0(u^2)$, pour tout $n \in \mathbb{N}^*$, on a

$$u_n = \gamma_{n+1} - \gamma_n = \frac{1}{n+1} - \ln\left(1+\frac{1}{n}\right) = \frac{1}{n+1} - \frac{1}{n} + \frac{1}{2n^2} + o_{+\infty}(n^{-2}).$$

Puisque $\frac{1}{1+u} = 1 + o_0(1)$, on obtient

$$\frac{1}{n+1} - \frac{1}{n} = -\frac{1}{n^2}\,\frac{1}{1+\frac{1}{n}} = -\frac{1}{n^2} + o_{+\infty}(n^{-2})$$

et finalement,

$$u_n = -\frac{1}{n^2} + o_{+\infty}(n^{-2}) \underset{+\infty}{\sim} -\frac{1}{n^2}.$$

La série $\sum_n u_n$ est à termes négatifs et la série $\sum_n \frac{1}{n^2}$ converge. Les critères de comparaison pour les séries positives permettent de conclure que la série $\sum_n u_n$ converge. On note μ sa somme.
Comme la série $\sum_n u_n$ converge et a pour somme μ, sa suite des sommes partielles $(U_n)_n$ converge aussi vers μ. Or pour tout $n \in \mathbb{N}^*$ on a

$$U_n = \sum_{k=1}^{n} u_k = \sum_{k=1}^{n} (\gamma_{k+1} - \gamma_k) = \gamma_{n+1} - \gamma_1 = \gamma_{n+1} - 1.$$

On en déduit que $\lim_{n\to+\infty} \gamma_{n+1} = \mu + 1$ et par conséquent que la suite $(\gamma_n)_n$ converge vers $\gamma = \mu + 1$.

2 - Pour tout $n \in \mathbb{N}^*$, on a

$$\begin{aligned} \gamma_{n+1} - \gamma_n &= u_n = \frac{1}{n+1} - \ln\left(1 + \frac{1}{n}\right), \\ \text{et } v_{n+1} - v_n &= \gamma_{n+1} - \frac{1}{n+1} - \gamma_n + \frac{1}{n} = \frac{1}{n} - \ln\left(1 + \frac{1}{n}\right). \end{aligned}$$

On déduit de la question précédente que la suite $(\gamma_n)_n$ est décroissante et que la suite $(v_n)_n$ est croissante. De plus, $\lim_{n\to+\infty}(\gamma_n - v_n) = \lim_{n\to+\infty} \frac{1}{n} = 0$. Les suites $(\gamma_n)_n$ et $(v_n)_n$ sont donc adjacentes : elles convergent et ont même limite. Comme $(\gamma_n)_n$ converge vers γ, la suite $(v_n)_n$ converge aussi vers γ. Des propriétés des suites adjacentes, on déduit que pour tout $n \in \mathbb{N}^*$ on a

$$v_n < \gamma < \gamma_n.$$

Une valeur approchée de γ est donc $\frac{1}{2}(\gamma_n + v_n) = \gamma_n - \frac{1}{2n}$ et l'erreur commise est inférieure à $\frac{1}{2}(\gamma_n - v_n) = \frac{1}{2n}$. Pour obtenir une valeur approchée de γ à 10^{-3} près, il suffit de prendre $n = 500$ et calculer $\gamma_{500} - \frac{1}{1000}$.

On peut effectuer ce calcul d'une valeur approchée de γ avec MAPLE :

```
> evalf(sum( 1/k,k=1..500) - ln(500) - 1/1000);
                              .577215332
```

On observera que l'erreur commise est nettement inférieure à 10^{-3} (ce qui s'explique par le fait que la quantité $\frac{1}{2n}$ est un majorant de l'erreur) :

```
> evalf(gamma);
                              .5772156649
```

3 - Pour tout $n \in \mathbb{N}^*$ on a $\sum_{k=1}^{n} \frac{1}{k} = \gamma_n + \ln(n)$. Comme la suite $(\gamma_n)_n$ converge vers γ et que $\lim_{n\to+\infty} \ln(n) = +\infty$, on a $\gamma_n = \mathcal{O}_{+\infty}(\ln(n))$ et $\sum_{k=1}^{n} \frac{1}{k} \underset{+\infty}{\sim} \ln(n)$.

Solution de l'exercice 1.8

1 - Pour tout $t \in \mathbb{R}$ et pour tout $n \in \mathbb{N}$ on a

$$\left|\alpha_n e^{int}\right| = |\alpha_n|.$$

Une condition nécessaire et suffisante pour que la série trigonométrique $\sum_n \alpha_n e^{int}$ converge absolument est donc que la série $\sum_n |\alpha_n|$ converge, autrement dit que la série $\sum_n \alpha_n$ converge absolument.

2 - Pour tout $n \in \mathbb{N}$ on a

$$|(-1)^n \varepsilon_n| = \left|(-1)^{n+1} \varepsilon_n\right| = \varepsilon_n.$$

On en déduit qu'une condition nécessaire et suffisante pour que les séries alternées $\sum_n (-1)^n \varepsilon_n$ et $\sum_n (-1)^{n+1} \varepsilon_n$ convergent est que la série $\sum_n \varepsilon_n$ converge.

Solution de l'exercice 1.9

1 - On montre, en utilisant par exemple le théorème des accroissements finis, que pour tout $x \in \mathbb{R}$ on a $-x \leqslant \sin(x) \leqslant x$. On a donc $|\sin(x)| \leqslant |x|$ pour tout réel x. On en déduit que pour tout $n \in \mathbb{N}^*$, $|\sin((-1)^n/n^2)| \leqslant 1/n^2$. Comme la série $\sum_n \frac{1}{n^2}$ converge, la série $\sum_n \sin((-1)^n/n^2)$ converge absolument.

2 - Pour tout $n \in \mathbb{N}$ avec $n \geqslant 2$, on a

$$\left|\frac{(-1)^n}{n^2+(-1)^n}\right| = \frac{1}{n^2+(-1)^n} \leqslant \frac{1}{n^2-1}.$$

Or $\frac{1}{n^2-1} \underset{+\infty}{\sim} \frac{1}{n^2}$ et la série à termes positifs $\sum_n \frac{1}{n^2}$ converge. On en déduit que $\sum_n \frac{1}{n^2-1}$ converge puis que $\sum_n \frac{(-1)^n}{n^2+(-1)^n}$ converge absolument.

3 - Pour tout $n \in \mathbb{N}$ on a

$$\left|\frac{(1+\mathrm{i})}{n^2+1}\right| = \frac{|1+\mathrm{i}|}{n^2+1} = \frac{1}{n^2+1}.$$

Or $\frac{1}{n^2+1} \underset{+\infty}{\sim} \frac{1}{n^2}$ et la série à termes positifs $\sum_n \frac{1}{n^2}$ converge. On en déduit que $\sum_n \frac{1}{n^2+1}$ converge et que $\sum_n \frac{(1+\mathrm{i})}{n^2+1}$ converge absolument.

Solution de l'exercice 1.10

Supposons que la suite $(v_n)_n$ de terme général $v_n = \sqrt[n]{|u_n|}$ tende vers $+\infty$. On a alors[66] :

$$\exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad \left(n \geqslant N \implies \sqrt[n]{|u_n|} \geqslant 1\right).$$

On en déduit que pour tout entier $n \geqslant N$, on a $|u_n| \geqslant 1$. Il est alors clair que la suite de terme général $|u_n|$ ne peut pas converger vers 0. Or[67], la suite de terme général u_n converge vers 0 si et seulement si la suite de terme général $|u_n|$ converge vers 0. On peut donc affirmer que la suite de terme général u_n ne converge pas vers 0 et par conséquent que la série $\sum_n u_n$ diverge grossièrement.

Solution de l'exercice 1.11

1 - On a $\sqrt[n]{|u_n|} = \left(\frac{n}{2n+1}\right)^{\frac{1}{2}}$; cette quantité tend vers $1/\sqrt{2}$ quand n tend vers l'infini. Par conséquent la série converge.

---

[66] Voir la définition 5.2 p. 172 du *Cours de première année*. On prend ici $\kappa = 1$.

[67] Voir la proposition 5.4 p. 171 du *Cours de première année*.

2 - On a $\sqrt[n]{|u_n|} = \left|\frac{n+\mathrm{i}}{n^2}\right| = \frac{1}{n^2}\sqrt{n^2+1}$; cette quantité tend vers 0 quand n tend vers l'infini. Par conséquent la série converge.

3 - On a $\sqrt[n]{|u_n|} = \mathrm{sh}\left(\frac{1}{n}\right)^n = \mathrm{e}^{n\ln(\mathrm{sh}\left(\frac{1}{n}\right))}$. Or $\mathrm{sh}(x) \underset{0}{\sim} x$ donc $\mathrm{sh}\left(\frac{1}{n}\right) \underset{+\infty}{\sim} \frac{1}{n}$ et $n\ln(\mathrm{sh}\left(\frac{1}{n}\right)) \underset{+\infty}{\sim} -n\ln(n)$. On en déduit que $\sqrt[n]{|u_n|}$ tend vers 0 quand n tend vers l'infini. Par conséquent la série converge.

Solution de l'exercice 1.12

Supposons que la suite $(v_n)_n$ de terme général $v_n = |u_{n+1}/u_n|$ tende vers $+\infty$. On a alors[68] :

$$\exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad \left(n \geqslant N \implies \left|\frac{u_{n+1}}{u_n}\right| \geqslant 1\right).$$

On en déduit que pour tout entier $n \geqslant N$ on a $|u_{n+1} \geqslant |u_n|$. Donc, à partir du rang N la suite de terme général $|u_n|$ est croissante. Comme elle est positive (et non nulle), elle ne tend pas vers 0. Or[69], la suite de terme général u_n converge vers 0 si et seulement si la suite de terme général $|u_n|$ converge vers 0. On peut donc affirmer que la suite de terme général u_n ne converge pas vers 0 et par conséquent que la série $\sum_n u_n$ diverge grossièrement.

Solution de l'exercice 1.13

1 - Pour tout $n \in \mathbb{N}^*$, on a

$$\left|\frac{u_{n+1}}{u_n}\right| = \left(\frac{n}{n+1}\right)^n = \left(1+\frac{1}{n}\right)^{-n} = \mathrm{e}^{-n\ln(1+\frac{1}{n})}.$$

Puisque $\ln(1+\frac{1}{n}) \underset{+\infty}{\sim} \frac{1}{n}$, on a $\lim_{n\to+\infty}\left|\frac{u_{n+1}}{u_n}\right| = 1/\mathrm{e} < 1$. Par conséquent la série converge.

2 - Pour tout $n \in \mathbb{N}^*$, on a

$$\left|\frac{u_{n+1}}{u_n}\right| = \left(\frac{n}{n+1}\right)^2 \left(\frac{(n+1)!}{n!}\right)^2 = n^2.$$

Comme cette quantité tend vers $+\infty$, la série diverge grossièrement.

3 - Pour tout $n \in \mathbb{N}$, on a

$$\left|\frac{u_{n+1}}{u_n}\right| = \frac{n+1}{2n}.$$

Comme cette quantité tend vers $1/2$, la série converge absolument.

---

[68] Voir la définition 5.2 p. 172 du *Cours de première année*. On prend ici $\kappa = 1$.

[69] Voir la proposition 5.4 p. 171 du *Cours de première année*.

Solution de l'exercice 1.14

1 - La fonction sinus est positive et strictement croissante sur $[0, \frac{\pi}{2}]$ et pour tout $n \in \mathbb{N}^*$ on a

$$0 < \frac{1}{n+1} < \frac{1}{n} \leqslant 1 < \frac{\pi}{2}.$$

On en déduit que $0 < \sin(\frac{1}{n+1}) < \sin(\frac{1}{n})$ et donc que la suite $(\alpha_n)_n$ de terme général $\alpha_n = \sin(1/n)$ est positive et décroissante. Comme de plus $\lim_{n\to+\infty} \alpha_n = 0$, le critère spécial pour les séries alternées permet de conclure que la série $\sum_n (-1)^n \sin(1/n)$ converge. Elle ne converge pas absolument car : $|(-1)^n \sin(1/n)| = \sin(1/n) \underset{+\infty}{\sim} 1/n$ et la série $\sum_n 1/n$ diverge.

2 - Le critère de D'Alembert permet d'établir que la série $\sum_n (-1)^n \frac{2^n}{n!}$ converge absolument : pour tout $n \in \mathbb{N}^*$ on a

$$\left|\frac{u_{n+1}}{u_n}\right| = \frac{2}{n+1}$$

et cette quantité tend vers 0 quand n tend vers $+\infty$. Il n'y a donc pas lieu d'utiliser le critère spécial pour les séries alternées.

3 - La fonction logarithme est positive et croissante sur $[1, +\infty[$ donc pour tout $n \in \mathbb{N}^*$ on a $\ln(n+1) \geqslant \ln(n)$. On en déduit que la suite $(\alpha_n)_n$ de terme général $\alpha_n = \frac{1}{\ln(n)}$ est positive et décroissante. Comme de plus $\lim_{n\to+\infty} \alpha_n = 0$, le critère spécial pour les séries alternées permet de conclure que la série $\sum_n \frac{(-1)^n}{\ln(n)}$ converge. Elle ne converge pas absolument car la série de Bertrand $\sum_n \frac{1}{\ln(n)}$ diverge.

Solution de l'exercice 1.15

1 - a) Pour $\theta \in]0, \pi/2[$, on obtient en utilisant la formule de Moivre,

$$\begin{aligned}
\sin((2n+1)\theta) &= \mathrm{Im}\Big\{\big(\cos(\theta) + \mathrm{i}\sin(\theta)\big)^{2n+1}\Big\} \\
&= \mathrm{Im}\Big\{\sum_{k=0}^{2n+1} C_{2n+1}^k \cos^{2n+1-k}(\theta)\ \mathrm{i}^k \sin^k(\theta)\Big\} \\
&= \mathrm{Im}\Big\{\sum_{\ell=0}^{n} C_{2n+1}^{2\ell} \cos^{2n+1-2\ell}(\theta)\ (-1)^\ell \sin^{2\ell}(\theta) \\
&\qquad + \mathrm{i}\sum_{\ell=0}^{n} C_{2n+1}^{2\ell+1} \cos^{2n-2\ell}(\theta)\ (-1)^\ell \sin^{2\ell+1}(\theta)\Big\} \\
&= \sum_{\ell=0}^{n} (-1)^\ell C_{2n+1}^{2\ell+1} \cos^{2(n-\ell)}(\theta)\ \sin^{2\ell+1}(\theta),
\end{aligned}$$

en décomposant la somme selon que k prend une valeur paire ($k = 2\ell$ avec $\ell \in \{0, \dots, n\}$) ou impaire ($k = 2\ell + 1$ avec $\ell \in \{0, \dots, n\}$). On en déduit que

$$\begin{aligned}\sin((2n+1)\theta) &= \sin^{2n+1}(\theta) \times \sum_{\ell=0}^{n} (-1)^\ell C_{2n+1}^{2\ell+1} \frac{\cos^{2(n-\ell)}(\theta)}{\sin^{2(n-\ell)}(\theta)} \\ &= \sin^{2n+1}(\theta) \times \sum_{\ell=0}^{n} (-1)^\ell C_{2n+1}^{2\ell+1} \left(\operatorname{cotan}^2(\theta)\right)^{n-\ell} \\ &= \sin^{2n+1}(\theta) \times P_n(\operatorname{cotan}^2(\theta)).\end{aligned}$$

b) Pour tout $k \in \{1, \dots, n\}$, le réel $\theta_k = \frac{k\pi}{2n+1}$ appartient clairement à l'intervalle $]0, \pi/2[$ puisque $\theta_k = \frac{k}{n+1/2}\frac{\pi}{2}$. Il résulte alors de la question précédente que pour $k \in \{1, \dots, n\}$,

$$\sin((2n+1)\theta_k) = \sin^{2n+1}(\theta_k) \times P_n(\operatorname{cotan}^2(\theta_k)).$$

Or, $\sin((2n+1)\theta_k) = \sin(k\pi) = 0$ et $\sin(\theta_k) \neq 0$ car la fonction sinus ne s'annule pas sur l'intervalle $]0, \pi/2[$. On a donc $P_n(\operatorname{cotan}^2(\theta_k)) = 0$, autrement dit le réel $x_k = \operatorname{cotan}^2(\theta_k)$ est une racine de P_n pour tout $k \in \{1, \dots, n\}$.

c) Puisque la fonction cotangente est une injection de $]0, \pi/2[$ dans $]0, +\infty[$ (elle est en effet continue et strictement décroissante sur $]0, \pi/2[$) et que $\theta_k \in]0, \pi/2[$, les réels x_k pour $k \in \{1, \dots, n\}$ sont tous distincts. Le polynôme P_n de $\mathbb{R}[X]$ étant de degré n, il possède au plus n racines réelles. Les réels $x_1, \dots, x_n$ sont donc les n racines réelles distinctes du polynôme P_n. La décomposition en produit de facteurs irréductibles du polynôme P_n dans $\mathbb{R}[X]$ est alors

$$P_n = a_n(X - x_1) \times (X - x_2) \times \cdots \times (X - x_n)$$

où a_n désigne le coefficient du monôme de degré n dans l'expression de P_n. On a $a_n = (-1)^0 C_{2n+1}^1 = 2n + 1$ et par conséquent,

$$P_n = (2n+1) \prod_{k=1}^{n} \left(X - \operatorname{cotan}^2\left(\frac{k\pi}{2n+1}\right)\right)$$

2 - a) En utilisant la relation donnant la somme des racines d'un polynôme[70], on obtient :

$$\sum_{k=1}^{n} x_k = \sum_{k=0}^{n} \operatorname{cotan}^2(\theta_k) = -\frac{a_{n-1}}{a_n} = \frac{C_{2n+1}^3}{C_{2n+1}^1} = \frac{n(2n-1)}{3}.$$

où a_n (resp. a_{n-1}) désigne le coefficient du monôme de degré n (resp. $n-1$) dans l'expression de P_n. Par ailleurs, pour tout $\theta \in]0, \pi/2[$ on a

$$\frac{1}{\sin^2(\theta)} = \frac{\sin^2(\theta) + \cos^2(\theta)}{\sin^2(\theta)} = 1 + \operatorname{cotan}^2(\theta).$$

[70] Voir le *Cours de première année* p. 247.

On en déduit que

$$\sum_{k=1}^{n} \frac{1}{\sin^2(\theta)} = n + \sum_{k=1}^{n} \operatorname{cotan}^2(\theta_k) = n + \frac{n(2n-1)}{3} = \frac{2n(n+1)}{3}.$$

b) Remarquons que pour $x \in]0, \pi/2[$ on a l'équivalence suivante liée au fait que $\tan(x) > 0$, $\sin(x) > 0$ et que la fonction carré est strictement croissante :

$$\frac{1}{\tan^2(x)} < \frac{1}{x^2} < \frac{1}{\sin^2(x)} \iff \sin(x) < x < \tan(x).$$

Pour montrer que pour $x \in]0, \pi/2[$ on a $\sin(x) < x$ et $\tan(x) > x$, il suffit d'appliquer le théorème des accroissements finis aux fonctions sinus et tangente entre 0 et x réel fixé dans l'intervalle $]0, \pi/2[$.

c) Appliquons la relation précédente en prenant $x = \theta_k = \frac{k\pi}{2n+1}$ pour $k \in \{1, \ldots, n\}$. Il a déjà été établi que $\theta_k \in]0, \pi/2[$. On a alors pour tout $k \in \{1, \ldots, n\}$,

$$\operatorname{cotan}^2(\theta_k) < \frac{(2n+1)^2}{\pi^2} \frac{1}{k^2} < \sin^{-2}(\theta_k).$$

En sommant ces relations pour $k \in \{1, \ldots, n\}$, et en utilisant les résultats de la question 4.a, on obtient :

$$\frac{1}{3}n(n+1) = \sum_{k=1}^{n} \operatorname{cotan}^2(\theta_k) < \frac{(2n+1)^2}{\pi^2} \sum_{k=1}^{n} \frac{1}{k^2} < \sum_{k=1}^{n} \sin^{-2}(\theta_k) = \frac{2}{3}n(n+1).$$

On en déduit que

$$\frac{\pi^2 n(2n-1)}{3(2n+1)^2} < \sum_{k=1}^{n} \frac{1}{k^2} < \frac{2\pi^2 n(n+1)}{3(2n+1)^2}.$$

3 - Considérons les suites $(u_n)_n$ et $(v_n)_n$ de terme général

$$u_n = \frac{\pi^2 n(2n-1)}{3(2n+1)^2} = \frac{\pi^2(2-1/n)}{3(2+1/n)^2} \quad \text{et} \quad v_n = \frac{2\pi^2 n(n+1)}{3(2n+1)^2} = \frac{2\pi^2(1+1/n)}{3(2+1/n)^2}.$$

Ces deux suites convergent vers $\pi^2/6$. D'après le théorème d'encadrement, on en conclut que :

$$\zeta(2) = \lim_{n \to +\infty} \sum_{k=1}^{n} \frac{1}{k^2} = \frac{\pi^2}{6}.$$

Solution de l'exercice 1.16

1 - Comme la suite $(b_n)_n$ est bornée, la suite de terme général $v_n = \frac{a}{n} + \frac{b_n}{n(n+1)}$ converge vers 0. En utilisant le développement limité à l'ordre 1 de sinus en 0 ($\sin(x) = x + o_0(x)$), on obtient :

$$u_n = (-1)^n \frac{a}{n} + (-1)^n \frac{b_n}{n(n+1)} + R_n \quad \text{où} \quad R_n = o_{+\infty}(n^{-2}).$$

La série de terme général $(-1)^n \frac{1}{n}$ est la série harmonique alternée qui est semi-convergente. La série de terme général $(-1)^n \frac{b_n}{n(n+1)}$ est absolument convergente car en désignant par M la borne supérieure des valeurs prises par la suite $(b_n)_n$, on a :

$$\left|(-1)^n \frac{b_n}{n(n+1)}\right| \leqslant \frac{M}{n^2}$$

et la série $\sum_{n\geqslant 1} \frac{1}{n^2}$ converge. Enfin, il est clair que la série de terme général R_n est absolument convergente d'après les critères de comparaison pour les séries à termes positifs. On en conclut que la série $\sum_n u_n$ est simplement convergente.

2 - a) Pour tout $n \in \mathbb{N}$ avec $n \geqslant 2$ on a

$$p = \sum_{k=0}^{n-2} \frac{n!}{k!} = n(n-1) \sum_{k=0}^{n-2} \frac{(n-2)!}{k!}.$$

Pour tout $k \in \{0, \ldots, n-2\}$ le réel $\frac{(n-2)!}{k!}$ est un entier naturel puisque

$$\frac{(n-2)!}{k!} = (n-2) \times \cdots \times (k+1).$$

On en déduit que $S = \sum_{k=0}^{n-2} \frac{(n-2)!}{k!}$ est un entier naturel. Comme le produit de deux entiers naturels successifs est un entier pair, $n(n-1)$ est pair et l'entier $p = n(n-1)S$ est pair.

b) Appliquons la formule de Taylor à l'ordre $(n+1)$ à la fonction exponentielle entre 0 et 1 : il existe $c_n \in]0,1[$ tel que

$$\mathrm{e} - 1 = \mathrm{e}^1 - \mathrm{e}^0 = \sum_{k=0}^{n+1} \frac{1}{k!} + \frac{\mathrm{e}^{c_n}}{(n+2)!}.$$

On en déduit que :

$$n!\mathrm{e} = n! + \sum_{k=0}^{n+1} \frac{n!}{k!} + \frac{\mathrm{e}^{c_n}}{(n+1)(n+2)} = n! + \underbrace{\sum_{k=0}^{n-2} \frac{n!}{k!}}_{=p} + n + 1 + \frac{1}{n+1} + \frac{\mathrm{e}^{c_n}}{(n+1)(n+2)}.$$

Pour $n \geqslant 2$, les entiers $n!$ et p sont pairs : il existe $q \in \mathbb{N}$ tel que $n! + p = 2q$. On a donc

$$n!\mathrm{e} = 2q + (n+1) + \frac{1}{n+1} + \frac{\mathrm{e}^{c_n}}{(n+1)(n+2)}$$

et compte tenu de la périodicité de la fonction sinus

$$\begin{aligned} \sin(n!\mathrm{e}\pi) &= \sin\left(\frac{\pi \mathrm{e}^{c_n}}{(n+1)(n+2)} + (n+1)\pi + \frac{\pi}{n+1} + 2q\pi\right) \\ &= (-1)^{n+1} \sin\left(\frac{\pi}{n+1} + \frac{\pi \mathrm{e}^{c_n}}{n+1}\right). \end{aligned}$$

Pour tout $n \in \mathbb{N}$ on a $|b_n| = |\pi \mathrm{e}^{c_n - 1}| \leqslant \mathrm{e}\pi$. D'après la question 1 (où on prend $a = \pi$ et $b_n = \pi \mathrm{e}^{c_n}$) on peut affirmer que la série $\sum_n \sin(\mathrm{e}\pi n!)$ est semi-convergente.

CHAPITRE 2

Suites de fonctions

Nous allons aborder dans ce chapitre l'étude des suites de fonctions. De même qu'une suite numérique peut être vue comme une famille de scalaires indicée par $\mathbb{N}$, une suite de fonctions est définie comme une famille de fonctions indicée par $\mathbb{N}$. Bien que nous ne considérions dans ce chapitre que des suites de fonctions réelles à valeurs réelles, les différentes notions introduites s'étendent au cas des fonctions complexes de la variable complexe. Toutefois, à ce stade du cours, les fonctions de la variable complexe n'ont pas été étudiées et nous disposons de très peu d'exemples de telles fonctions. C'est la raison pour laquelle nous nous restreignons dans un premier temps au cas de fonctions d'une variable réelle, reléguant à un chapitre ultérieur l'étude des suites de fonctions complexes de la variable complexe.

2.1 L'espace vectoriel des applications bornées

On désigne par X un sous-ensemble non vide de $\mathbb{R}$. Rappelons[1] qu'une application f de X dans $\mathbb{R}$ est dite *bornée* sur X si elle est à la fois majorée sur X et minorée sur X, en d'autres termes si :

$$\exists (M,m) \in \mathbb{R}^2 \quad \forall x \in X \quad m \leqslant f(x) \leqslant M.$$

De manière équivalente, f est bornée sur X si :

$$\exists K \in \mathbb{R}_+^* \quad \forall x \in X \quad |f(x)| \leqslant K.$$

On rappelle également[2] qu'une application f continue sur un intervalle fermé et borné $[a,b]$ est une application bornée qui atteint ses bornes :

$$\exists (x_1,x_2) \in [a,b]^2 \quad f(x_1) = \inf_{x\in[a,b]} f(x) \quad f(x_2) = \sup_{x\in[a,b]} f(x).$$

On désigne par $\mathcal{B}(X,\mathbb{R})$ l'ensemble des applications de X dans $\mathbb{R}$ qui sont bornées sur X. Il s'agit d'un sous-espace vectoriel de l'espace vectoriel des applications de X dans $\mathbb{R}$. Pour $f \in \mathcal{B}(X,\mathbb{R})$, on note

$$\|f\|_{\infty,X} \quad \text{ou} \quad \sup_{x\in X} |f(x)|$$

[1] Voir le *Cours de première année* p. 565.

[2] Voir le théorème 13.4 p. 596 du *Cours de première année*.

la borne supérieure de l'ensemble $\{|f(x)| \ ; \ x \in X\}$. On a les propriétés suivantes.

Proposition 2.1 *Soient* $(f,g) \in \mathcal{B}(X,\mathbb{R})^2$ *et* $\lambda \in \mathbb{R}$ *; on a,*

1. $\|f\|_{\infty,X} = 0 \Longleftrightarrow f = 0$;
2. $\|\lambda \cdot f\|_{\infty,X} = |\lambda| \ \|f\|_{\infty,X}$;
3. $\|f+g\|_{\infty,X} \leqslant \|f\|_{\infty,X} + \|g\|_{\infty,X}$;
4. $\|f \times g\|_{\infty,X} \leqslant \|f\|_{\infty,X} \times \|g\|_{\infty,X}$.

Démonstration Ces propriétés se démontrent en revenant à la définition de la borne supérieure et en utilisant les propriétés de la valeur absolue. Montrons la quatrième de ces relations. On a

$$\|f \times g\|_{\infty,X} = \sup_{x\in X}\big(|f \times g|(x)\big) = \sup\{|f(x) \times g(x)| \ \ ; \ x \in X\}.$$

Or, pour tout $x \in X$ on a$|f(x) \times g(x)| = |f(x)| \times |g(x)|$ et

$$0 \leqslant |f(x)| \leqslant \sup_{t\in X} |f(t)|, \qquad 0 \leqslant |g(x)| \leqslant \sup_{t\in X} |g(t)|.$$

On en déduit que pour tout $x \in X$,

$$|f(x) \times g(x)| \leqslant \sup_{t\in X} |f(t)| \times \sup_{t\in X} |g(t)|,$$

et par conséquent que $\sup_{t\in X} |f(t)| \times \sup_{t\in X} |g(t)|$ est un majorant de l'ensemble $\{|f(x) \times g(x)| \ ; \ x \in X\}$. Comme la borne supérieure de cet ensemble est, par définition, le plus petit des majorants, on a nécessairement :

$$\sup_{x\in X} (|f \times g|(x)) \leqslant \sup_{x\in X} |f(x)| \times \sup_{x\in X} |g(x)|,$$

autrement dit $\|f \times g\|_{\infty,X} \leqslant \|f\|_{\infty,X} \times \|g\|_{\infty,X}$. □

Remarque Les trois premières assertions de la proposition 2.1 permettent de dire que l'application

$$N : f \in \mathcal{B}(X,\mathbb{R}) \longmapsto \|f\|_{\infty,X} = \sup_{x\in X} |f(x)|$$

est une « norme » sur l'espace vectoriel $\mathcal{B}(X,\mathbb{R})$. La notion d'espace vectoriel normé sera abordée en détail au chap. 7. Comme autres exemples déjà connus d'espaces vectoriels normés, on peut citer l'espace vectoriel $\mathbb{R}$ des réels muni de l'application valeur absolue et l'espace vectoriel $\mathbb{C}$ des complexes avec pour norme l'application qui à un nombre complexe associe son module. On remarquera que pour ces deux cas on a $|x \times y| = |x| \times |y|$, que x et y désignent des réels et $|\cdot|$ la valeur absolue ou que x et y désignent des complexes et $|\cdot|$ le module. ◇

Exemple Considérons l'application $f : x \in \mathbb{R} \mapsto (x^2-1)\mathrm{e}^{-x^2}$ et déterminons la valeur de $\sup_{x\in[-2,2]} |f(x)|$. Remarquons que comme f est continue sur $\mathbb{R}$, elle est continue sur l'intervalle fermé borné $[-2,2]$, donc elle est bornée et atteint ses bornes sur $[-2,2]$. La quantité $\sup_{x\in[-2,2]} |f(x)|$ est par conséquent bien définie. Par ailleurs, f étant paire, on a $\sup_{x\in[-2,2]} |f(x)| = \sup_{x\in[0,2]} |f(x)|$. Pour déterminer la valeur de cette borne supérieure, étudions les variations de f sur $[0,2]$. L'application f est dérivable sur $\mathbb{R}$ avec $f'(x) = 2x(2-x^2)\mathrm{e}^{-x^2}$ pour tout $x \in \mathbb{R}$. L'application f admet le tableau de variation suivant :

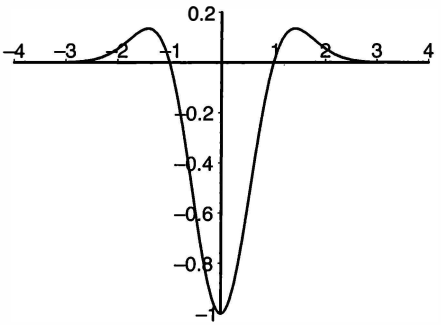
x	0		$\sqrt{2}$		$+\infty$
$f'(x)$	1	$+$	0	$-$	0
$f(x)$	-1	↗	e^{-2}	↘	0

On déduit du tableau de variation que

$$\sup_{x\in[-2,2]} |f(x)| = \sup_{x\in[0,2]} |f(x)| = \max(|-1|, \mathrm{e}^{-2}, f(2)) = 1.$$

On peut visualiser la représentation graphique de f avec MAPLE.

```
> plot((x^2-1)*exp{-x^2},x=-4..4);
```



On remarquera que l'on a aussi $\sup_{x\in]-2,2[} |f(x)| = 1$ et $\sup_{x\in\mathbb{R}} |f(x)| = 1$. Par ailleurs, $\sup_{x\in[1,2]} |f(x)| = \mathrm{e}^{-2}$.

Tirons profit de cet exemple pour remarquer que $\sup_{x\in[-2,2]} |(x^2-1)| = 3$ et $\sup_{x\in[-2,2]} |\mathrm{e}^{-x^2}| = 1$ et par conséquent que

$$\sup_{x\in[-2,2]} |(x^2-1)\mathrm{e}^{-x^2}| < \sup_{x\in[-2,2]} |(x^2-1)| \times \sup_{x\in[-2,2]} |\mathrm{e}^{-x^2}|.$$

Exercice 2.1 Calculer $\sup_{x\in]0,2[} |\mathrm{sh}(x^2-x-1)|$.

2.2 Convergence simple et convergence uniforme

2.2.1 Généralités

On désigne par X un sous-ensemble non vide de $\mathbb{R}$. On appelle *suite d'applications de X dans $\mathbb{R}$*, une suite $(f_n)_n$ d'éléments de l'espace vectoriel $\mathcal{A}(X,\mathbb{R})$ des applications de X dans $\mathbb{R}$. En d'autres termes, une suite d'applications $(f_n)_n$ est une famille dénombrable d'applications de X dans $\mathbb{R}$ indicée par $\mathbb{N}$. On parle de la suite d'applications $(f_n)_n$ de X dans $\mathbb{R}$ de terme général $f_n : x \in X \mapsto f_n(x)$.

On note $\mathcal{S}(X,\mathbb{R})$ l'ensemble des suites d'applications de X dans $\mathbb{R}$.

Remarque On confond très souvent du point de vue terminologique *suite de fonctions* avec *suite d'applications*. Nous utiliserons le terme « suite d'applications » lorsque les termes de la suite seront définis sur un même ensemble qui sera précisé. Le terme « suite de fonctions » sera utilisé comme terme générique. Ainsi, on parlera de la suite de fonctions de terme général $f_n : x \mapsto \ln(x - \frac{1}{n})$ et de la suite d'applications de terme général $f_n : x \in]1,+\infty[\mapsto \ln(x - \frac{1}{n})$. ◇

Il est important que toutes les fonctions f_n soient définies sur un même ensemble X. On ne peut pas, par exemple, considérer que les fonctions $f_n : x \mapsto \ln(x-n)$ constituent une suite d'applications car f_n est définie sur $]n,+\infty[$ et $\bigcap_{n\in\mathbb{N}}]n,+\infty[= \varnothing$.

Étant donné deux suites $(f_n)_n$ et $(g_n)_n$ d'applications de X dans $\mathbb{R}$, on définit la somme $(f_n)_n + (g_n)_n$ de ces deux suites comme étant la suite d'applications de terme général $x \in X \mapsto f_n(x) + g_n(x)$.

On munit également l'ensemble $\mathcal{S}(X,\mathbb{R})$ d'une loi de composition externe : pour tout $\lambda \in \mathbb{R}$ et pour tout $(f_n)_n \in \mathcal{S}(X,\mathbb{R})$, la suite $\lambda \cdot (f_n)_n$ est la suite de terme général $x \in X \mapsto \lambda\, f_n(x)$.

L'ensemble $\mathcal{S}(X,\mathbb{R})$ muni de ces deux lois est un $\mathbb{R}$-espace vectoriel.

On définit enfin le produit des deux suites $(f_n)_n$ et $(g_n)_n$ d'applications de X dans $\mathbb{R}$ comme étant la suite d'applications de terme général $x \in X \mapsto f_n(x)\, g_n(x)$.

2.2.2 Convergence simple

Si $(f_n)_n$ est une suite d'applications de X dans $\mathbb{R}$, alors pour $x \in X$ fixé, la suite de terme général $f_n(x)$ est une suite numérique réelle. Si pour tout $x \in X$ fixé la suite numérique de terme général $f_n(x)$ converge, alors on dit que la suite d'application $(f_n)_n$ converge simplement sur X.

DÉFINITION 2.1 (Convergence simple)

✗ *On dit que la suite d'applications $(f_n)_n$ de X dans $\mathbb{R}$* converge simplement *sur X vers l'application f de X dans $\mathbb{R}$ si*

$$\boxed{\forall x \in X} \quad \forall \varepsilon \in \mathbb{R}_+^* \quad \exists N_{x,\varepsilon} \in \mathbb{N} \quad \forall n \in \mathbb{N}$$
$$\left(n \geqslant N_{x,\varepsilon} \implies |f_n(x) - f(x)| \leqslant \varepsilon\right), \qquad (1)$$

c'est-à-dire si pour tout $x \in X$, la suite numérique de terme général $u_n = f_n(x)$ converge vers $f(x)$. L'application f est alors qualifiée de limite de la suite d'applications $(f_n)_n$.

✗ *On dit que la suite d'applications $(f_n)_n$ de X dans $\mathbb{R}$* converge simplement sur X *s'il existe une application f de X dans $\mathbb{R}$ telle que la suite d'applications $(f_n)_n$ converge simplement vers f.*

Remarques

1. On peut montrer que l'on obtient une définition équivalente si on remplace dans l'assertion quantifiée (1) chacune des inégalités larges par une inégalité stricte.

2. Pour des applications f_n ayant pour ensemble de définition X, l'expression « la suite d'applications $(f_n)_n$ converge simplement sur X vers l'application f de X dans $\mathbb{R}$ » contient des redondances. Celles-ci sont usuelles.

3. Il y a unicité de l'application vers laquelle une suite d'applications converge simplement. Ceci résulte du fait qu'une suite d'applications $(f_n)_n$ de X dans $\mathbb{R}$ converge simplement sur X si pour tout $x \in X$ la suite numérique de terme général $u_n = f_n(x)$ converge et de l'unicité de la limite d'une suite numérique[3]. ◇

Exemples

1. La suite d'applications $(f_n)_n$ de terme général $f_n : x \in [0, +\infty[\mapsto \mathrm{e}^{-nx}$ converge simplement vers l'application

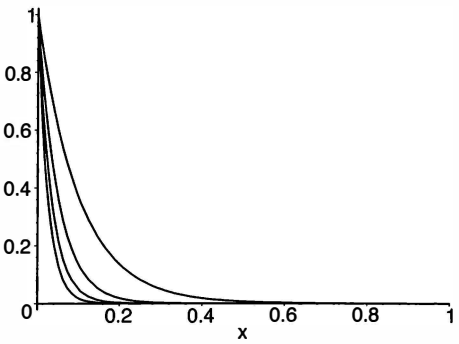
$$f : x \in [0, +\infty[\longmapsto \begin{cases} 0 & \text{si } x > 0 \\ 1 & \text{si } x = 0 \end{cases}.$$

Sur cet exemple on peut faire la constatation suivante : bien que toutes les applications f_n soient continues sur $[0, +\infty[$, la limite f de la suite $(f_n)_n$ n'est pas une application continue sur $[0, +\infty[$.

On obtient les représentations graphiques des termes f_{10}, f_{20}, f_{30} et f_{40} de la suite d'applications $(f_n)_n$ avec MAPLE de la manière suivante.

```
> fn := (n,x) -> exp(-n*x):
> plot([seq(fn(10*i,x),i=1..4)],x=0..1);
```

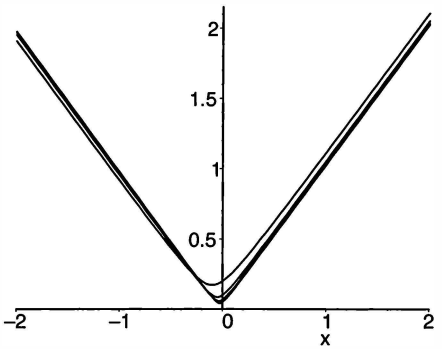
[3] Voir la proposition 5.1 p. 169 du *Cours de première année*.



2. La suite d'applications $(f_n)_n$ de terme général $f_n : x \in \mathbb{R} \mapsto \sqrt{x^2 + \frac{x}{n} + \frac{1}{n^2}}$ converge simplement sur $\mathbb{R}$ vers l'application valeur absolue. En effet pour tout réel x, la suite numérique de terme général $u_n = \sqrt{x^2 + \frac{x}{n} + \frac{1}{n^2}}$ tend vers $\sqrt{x^2}$ (*i.e.* $|x|$) quand n tend vers $+\infty$. Sur cet exemple on peut faire la constatation suivante : bien que toutes les applications f_n soient dérivables en 0, la limite f de la suite $(f_n)_n$ n'est pas une application dérivable en 0.

On obtient les représentations graphiques des termes f_5, f_{10}, f_{15} et f_{20} de la suite d'applications $(f_n)_n$ avec MAPLE de la manière suivante.

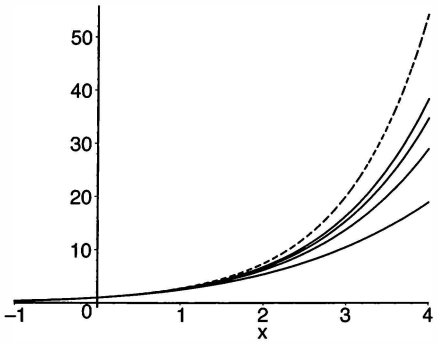
```
> fn := (n,x) -> sqrt(x^2+x/n+1/n^2):
> plot([seq(fn(5*i,x),i=1..4)],x=-2..2);
```



3. Considérons la suite $(f_n)_n$ de terme général $f_n : x \in \mathbb{R} \mapsto \left(1 + \frac{x}{n}\right)^n$. Pour un réel x fixé considérons la suite numérique réelle de terme général $f_n(x)$. Pour tout $n \in \mathbb{N}$ avec $n > -x$ on a $f_n(x) = \exp(n \ln(1 + x/n))$ et $n \ln(1 + x/n) \underset{n \to +\infty}{\sim} x$. On en déduit que la suite numérique réelle de terme général $f_n(x)$ converge vers e^x et par conséquent que la suite d'applications $(f_n)_n$ converge simplement sur $\mathbb{R}$ vers l'application exponentielle.

On obtient les représentations graphiques des termes f_5, f_{10}, f_{15} et f_{20} de la suite d'applications $(f_n)_n$ avec MAPLE de la manière suivante.

```
> fn := (n,x) -> (1+x/n)^n:
> plot([seq(fn(5*i,x),i=1..4),exp(x)],x=-1..4,
        linestyle=[seq(SOLID,i=1..4),DASH]);
```



2.2.3 Convergence uniforme

En général, l'entier $N = N_{x,\varepsilon}$ dans l'assertion (1) dépend à la fois de x et de ε (voir l'ordre des quantificateurs). Il arrive que dans certains cas, l'entier N ne dépende pas du réel x (il est le même quel que soit le réel x choisi dans l'ensemble X) mais uniquement du réel ε. En imposant à N d'être identique pour tous les réels x choisis dans l'ensemble X, on définit une condition de convergence plus restrictive que la convergence simple. On parle de « convergence uniforme » sur l'ensemble X. Nous verrons ultérieurement que la convergence uniforme d'une suite d'applications continues est une condition suffisante permettant d'affirmer que l'application limite est elle aussi continue.

DÉFINITION 2.2 (Convergence uniforme)

✗ *On dit que la suite d'applications $(f_n)_n$ de X dans $\mathbb{R}$* converge uniformément *sur X vers l'application f de X dans $\mathbb{R}$, si*

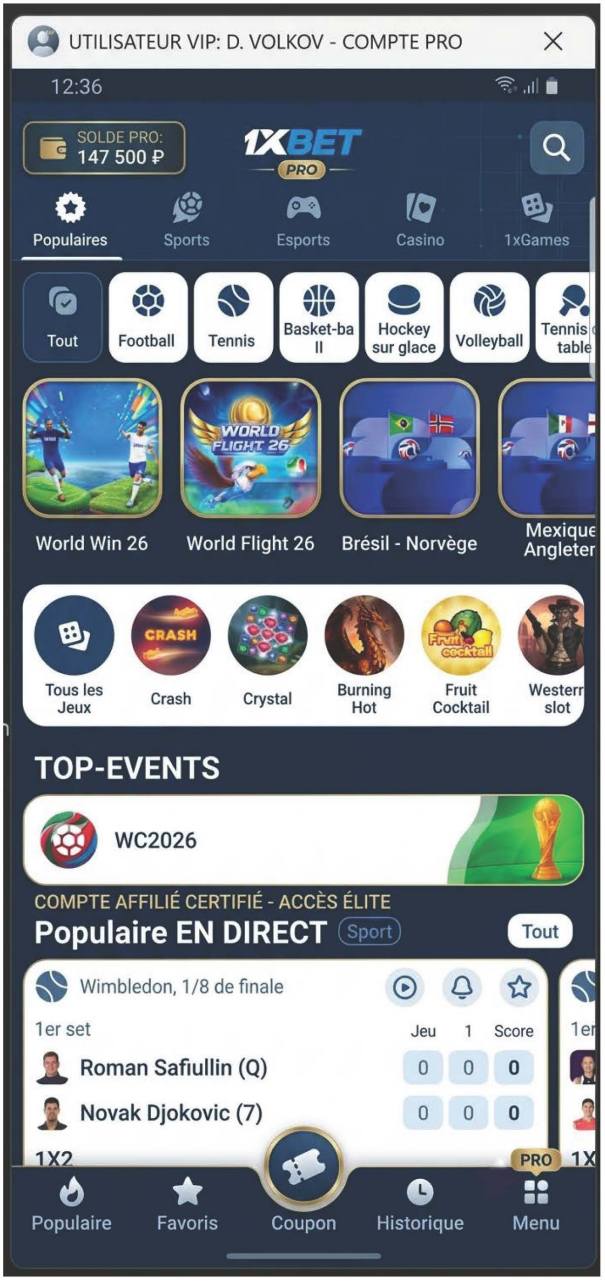
$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N_\varepsilon \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad \boxed{\forall x \in X} \quad \big(n \geqslant N_\varepsilon \implies |f_n(x) - f(x)| \leqslant \varepsilon\big). \qquad (2)$$

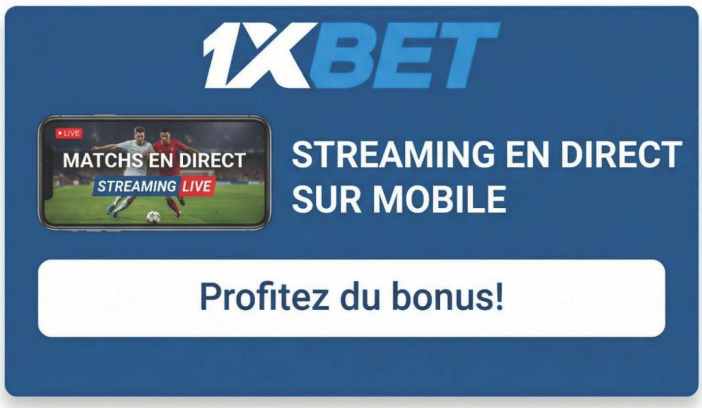
✗ *On dit que la suite d'applications $(f_n)_n$ de X dans $\mathbb{R}$* converge uniformément sur X *s'il existe une application f de X dans $\mathbb{R}$ telle que la suite d'applications $(f_n)_n$ converge uniformément vers f sur X.*

Remarques

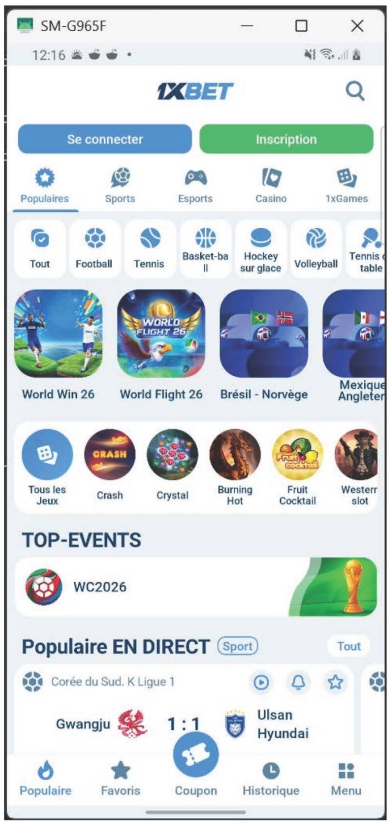
1. On peut montrer que l'on obtient une définition équivalente si on remplace dans l'assertion quantifiée (2) chacune des inégalités larges par une inégalité stricte.

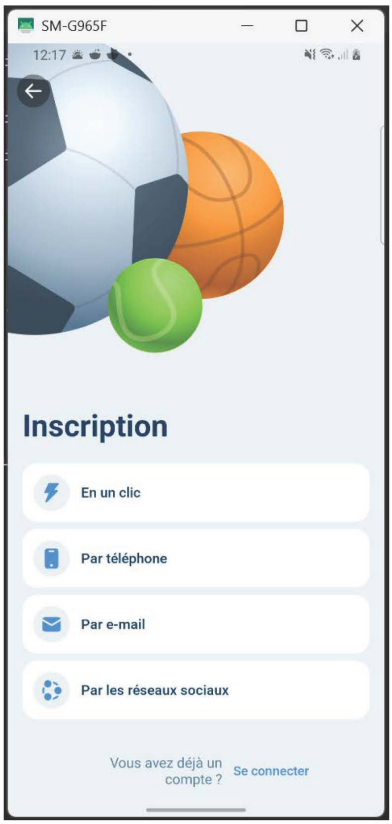

ÉTAPES D'INSCRIPTION
You Tube
pour voir comment creer un compte
CLICK TO WATCH

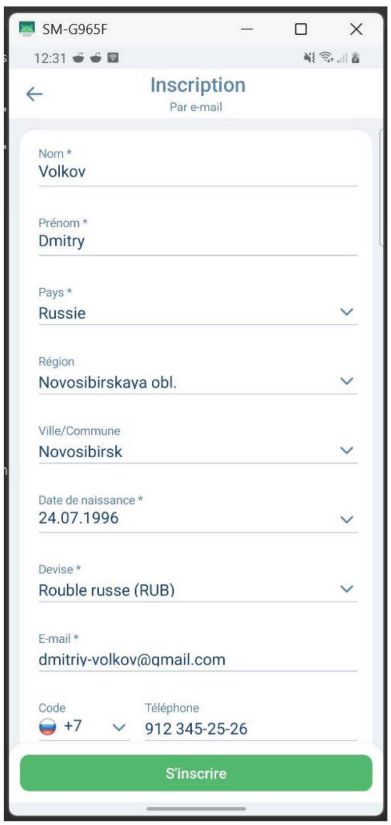

UTILISATEUR VIP: D. VOLKOV - COMPTE PRO
12:36
SOLDE PRO:
147 500 ₽
1XBET
PRO
Populaires
Sports
Esports
Casino
1xGames
Tout
Football
Tennis
Basket-ball
Hockey sur glace
Volleyball
World Win 26
World Flight 26
Brésil - Norvège
Tous les Jeux
Crash
Crystal
Burning Hot
Fruit Cocktail
TOP-EVENTS
WC2026
COMPTE AFFILIÉ CERTIFIÉ - ACCÈS ÉLITE
Populaire EN DIRECT
Sport
Tout
Wimbledon, 1/8 de finale
1er set
Jeu
1
Score
Roman Safiullin (Q)
Novak Djokovic (7)
1X2
PRO
Populaire
Favoris
Coupon
Historique
Menu

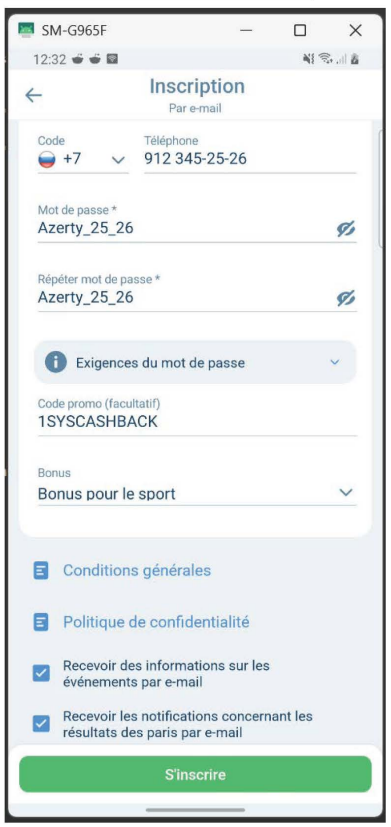

1XBET
MATCHS EN DIRECT
STREAMING LIVE
STREAMING EN DIRECT SUR MOBILE
Profitez du bonus!


CODE PROMO
1SYSCASHBACK


SM-G965F
12:16
1XBET
Se connecter
Inscription
Populaires
Sports
Esports
Casino
1xGames
Tout
Football
Tennis
Basket-ball
Hockey sur glace
Volleyball
World Win 26
World Flight 26
Brésil - Norvège
Tous les Jeux
Crash
Crystal
Burning Hot
Fruit Cocktail
TOP-EVENTS
WC2026
Populaire EN DIRECT
Sport
Tout
Corée du Sud. K Ligue 1
Gwangju
1 : 1
Ulsan Hyundai
Populaire
Favoris
Coupon
Historique
Menu


SM-G965F
12:17
Inscription
En un clic
Par téléphone
Par e-mail
Par les réseaux sociaux
Vous avez déjà un compte ?
Se connecter


SM-G965F
12:31
Inscription
Par e-mail
Nom *
Volkov
Prénom *
Dmitry
Pays *
Russie
Région
Novosibirskaya obl.
Ville/Commune
Novosibirsk
Date de naissance *
24.07.1996
Devise *
Rouble russe (RUB)
E-mail *
dmitriy-volkov@gmail.com
Code
+7
Téléphone
912 345-25-26
S'inscrire


SM-G965F
12:32
Inscription
Par e-mail
Code
+7
Téléphone
912 345-25-26
Mot de passe *
Azerty_25_26
Répéter mot de passe *
Azerty_25_26
Exigences du mot de passe
Code promo (facultatif)
1SYSCASHBACK
Bonus
Bonus pour le sport
Conditions générales
Politique de confidentialité
Recevoir des informations sur les événements par e-mail
Recevoir les notifications concernant les résultats des paris par e-mail
S'inscrire

2. Les deux assertions quantifiées (1) et (2) ne diffèrent que par la place du bloc quantificateur « $\forall x \in X$ ». On devra donc être vigilant lors de l'utilisation de ces assertions quantifiées. ◇

On peut exprimer l'assertion (2) à l'aide de la borne supérieure. Notons tout d'abord que l'assertion (2) se réécrit :

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N_\varepsilon \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad \big(n \geqslant N_\varepsilon \implies (\forall x \in X \quad |f_n(x) - f(x)| \leqslant \varepsilon)\big).$$

L'assertion quantifiée « $\forall x \in X \quad |f_n(x) - f(x)| \leqslant \varepsilon$ » indique que l'ensemble $\{|f_n(x) - f(x)| \ ; \ x \in X\}$ est majoré par ε. Par définition, la borne supérieure de cet ensemble qui est notée $\sup_{x\in X} |f_n(x) - f(x)|$, est le plus petit majorant de cet ensemble. Elle vérifie donc $\sup_{x\in X} |f_n(x) - f(x)| \leqslant \varepsilon$. Inversement si $\sup_{x\in X} |f_n(x) - f(x)| \leqslant \varepsilon$ alors l'ensemble $\{|f_n(x) - f(x)| \ ; \ x \in X\}$ est majorée par ε. L'assertion (2) est donc équivalente à

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N_\varepsilon \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad \big(n \geqslant N_\varepsilon \implies \sup_{x\in X} |f_n(x) - f(x)| \leqslant \varepsilon\big)$$

qu'on peut également exprimer sous la forme

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N_\varepsilon \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad \big(n \geqslant N_\varepsilon \implies \|f_n - f\|_{\infty,X} \leqslant \varepsilon\big). \tag{3}$$

Remarquons enfin que si l'on introduit la suite numérique $(u_n)_n$ de terme général $u_n = \sup_{x\in X} |f_n(x) - f(x)|$ alors l'assertion (3) exprime le fait[4] que la suite $(u_n)_n$ converge vers 0. On a donc la proposition suivante.

PROPOSITION 2.2 *Soit $(f_n)_n$ une suite d'applications de X dans $\mathbb{R}$ et f une application de X dans $\mathbb{R}$. Une condition nécessaire et suffisante pour que la suite $(f_n)_n$ converge uniformément vers f sur X est que la suite numérique de terme général*

$$u_n = \sup_{x\in X} |f_n(x) - f(x)|$$

converge vers 0.

Exemples

1. La suite d'applications $(f_n)_n$ de terme général $f_n : x \in [0, +\infty[\mapsto \mathrm{e}^{-nx}$ ne converge pas uniformément vers la fonction

$$f : x \in [0, +\infty[\longmapsto \begin{cases} 0 & \text{si } x > 0 \\ 1 & \text{si } x = 0 \end{cases}$$

bien que la suite d'applications $(f_n)_n$ converge simplement vers f (voir les exemples p. 71). En remarquant que $f_n(0) = f(0) = 1$, on a

$$\|f_n - f\|_{\infty,\mathbb{R}^+} = \sup_{x\in\mathbb{R}^+} |f_n(x) - f(x)| = \sup_{x\in\mathbb{R}_+^*} |f_n(x) - f(x)| = \sup_{x\in\mathbb{R}_+^*} |\mathrm{e}^{-nx}|.$$

[4] Voir la définition 5.1 p. 168 du *Cours de première année*.

Pour calculer la valeur de $\sup_{x\in\mathbb{R}_+^*} |f_n(x) - f(x)|$, il nous faut étudier pour $n \in \mathbb{N}^*$ fixé, la fonction $\phi : x \mapsto \mathrm{e}^{-nx}$. Cette fonction est positive et strictement décroissante sur $\mathbb{R}$. On a donc $u_n = \sup_{x\in\mathbb{R}_+^*} |\mathrm{e}^{-nx}| = \phi(0) = 1$. La suite $(u_n)_n$ qui est la suite constante égale à 1 ne converge pas vers 0.

2. La suite d'applications $(f_n)_n$ de terme général $f_n : x \in [1, +\infty[\mapsto \mathrm{e}^{-nx}$ converge uniformément vers la fonction nulle. En effet, on a

$$\|f_n - f\|_{\infty,[1,+\infty[} = \sup_{x\in[1,+\infty[} |\mathrm{e}^{-nx}| = \mathrm{e}^{-n}$$

et la suite $(u_n)_n$ de terme général $u_n = \mathrm{e}^{-n}$ converge vers 0.

3. La suite d'applications $(f_n)_n$ de terme général $f_n : x \in]0, +\infty[\mapsto \mathrm{e}^{-nx}$ ne converge pas uniformément vers la fonction nulle (mais elle converge simplement vers la fonction nulle). Ceci est lié au fait que

$$\sup_{x\in\mathbb{R}^+} |\mathrm{e}^{-nx}| = \sup_{x\in\mathbb{R}_+^*} |\mathrm{e}^{-nx}| = \phi(0) = 1.$$

On voit avec ces exemples que la notion de convergence uniforme est très dépendante de l'ensemble considéré. En particulier, cela n'a aucun sens de dire qu'une suite de fonctions converge uniformément. Il est indispensable de préciser sur quel ensemble a lieu la convergence uniforme.

D'un point de vue pratique, il n'est pas toujours nécessaire de calculer l'expression de $\sup_{x\in X} |f_n(x) - f(x)|$. En effet, la suite $(u_n)_n$ de terme général $u_n = \sup_{x\in X} |f_n(x) - f(x)|$ étant une suite à termes positifs, une condition nécessaire et suffisante pour que la suite réelle converge vers 0 est que la suite $(u_n)_n$ soit majorée par une suite qui converge vers 0. On a par conséquent le résultat suivant.

Corollaire 2.1 *Soit $(f_n)_n$ une suite d'applications de X dans $\mathbb{R}$ et f une application de X dans $\mathbb{R}$. Une condition nécessaire et suffisante pour que la suite d'applications $(f_n)_n$ converge uniformément vers f sur X est qu'il existe une suite numérique $(v_n)_n$ convergeant vers 0 telle que*

$$\forall n \in \mathbb{N} \qquad \sup_{x\in X} |f_n(x) - f(x)| \leqslant v_n.$$

Exemples

1. La suite d'applications $(f_n)_n$ de terme général $f_n : x \in \mathbb{R} \mapsto \frac{1}{n}\sin(nx)$ converge uniformément vers la fonction nulle puisque

$$u_n = \sup_{x\in\mathbb{R}} |f_n(x) - 0| = \sup_{x\in\mathbb{R}} \left|\frac{\sin(nx)}{n}\right| \leqslant \frac{1}{n} \quad \text{et} \quad \lim_{n\to+\infty} \frac{1}{n} = 0.$$

2. Montrons à l'aide du corollaire 2.1 que la suite d'applications $(f_n)_n$ de terme général $f_n : x \in \mathbb{R} \mapsto \sqrt{x^2 + \frac{1}{n^2}}$ converge uniformément sur $\mathbb{R}$ vers la fonction valeur absolue. Comme $x^2 + \frac{1}{n^2} \geqslant x^2$ pour tout $x \in \mathbb{R}$ et tout $n \in \mathbb{N}^*$, on a $\sqrt{x^2 + \frac{1}{n^2}} \geqslant \sqrt{x^2} = |x|$. Par ailleurs de l'inégalité $\sqrt{a+b} \leqslant \sqrt{a} + \sqrt{b}$ valable pour tous réels positifs a et b, on déduit que :

$$\sqrt{x^2 + \frac{1}{n^2}} \leqslant \sqrt{x^2} + \sqrt{\frac{1}{n^2}} = |x| + \frac{1}{n}.$$

Ainsi, pour tout réel x on a $0 \leqslant \sqrt{x^2 + \frac{1}{n^2}} - |x| \leqslant \frac{1}{n}$, ce qui implique que $\sup_{x\in\mathbb{R}} \left|\sqrt{x^2 + \frac{1}{n^2}} - |x|\right| \leqslant \frac{1}{n}$. Comme la suite numérique $(v_n)_n$ de terme général $v_n = \frac{1}{n}$ converge vers 0, on en déduit que la suite d'applications $(f_n)_n$ converge uniformément sur $\mathbb{R}$.

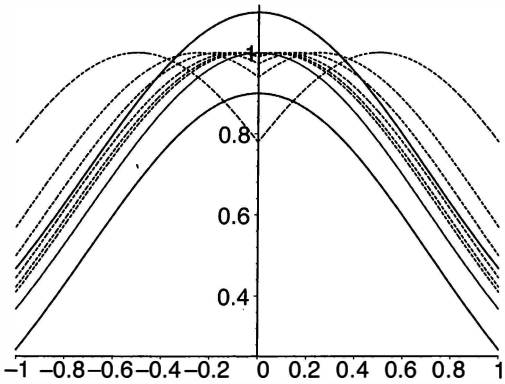
Remarque Soient $(f_n)_n$ une suite d'applications de X dans $\mathbb{R}$ convergeant uniformément vers une application f de X dans $\mathbb{R}$ et Y un sous-ensemble de X. La suite $(f_n)_n$ converge alors uniformément sur Y vers la restriction de f à l'ensemble Y. Ceci résulte du fait que si $Y \subset X$ alors $\sup_{x\in Y} |f_n(x) - f(x)| \leqslant \sup_{x\in X} |f_n(x) - f(x)|$. ◇

Interprétation graphique

Soit $(f_n)_n$ une suite d'applications définies sur un intervalle I à valeurs dans $\mathbb{R}$ qui converge vers une application f de I dans $\mathbb{R}$. Dans un plan affine rapporté à un repère orthonormé direct, considérons les représentations graphiques $\mathcal{C}_n$, $\mathcal{C}$, $\mathcal{C}_{+\varepsilon}$ et $\mathcal{C}_{-\varepsilon}$ des applications f_n, f, $f+\varepsilon$ et $f-\varepsilon$. Si la suite $(f_n)_n$ converge uniformément vers f alors pour tout $\varepsilon \in \mathbb{R}_+^*$, il existe un rang à partir duquel les représentations graphiques $\mathcal{C}_n$ sont toutes incluses dans la partie du plan comprise entre $\mathcal{C}_{+\varepsilon}$ et $\mathcal{C}_{-\varepsilon}$ (et éventuellement deux parallèles à l'axe des ordonnées si l'intervalle I est borné).

À titre d'exemple, considérons la suite $(f_n)_n$ dont le terme général est $f_n : x \in [-1,1] \mapsto \exp(-(|x| - \frac{1}{n})^2)$ qui converge uniformément sur $[-1,1]$ vers l'application $f : x \in [-1,1] \mapsto \exp(-x^2)$. Si l'on prend $\varepsilon = 10^{-1}$, la figure ci-dessous (obtenue avec MAPLE) nous indique qu'à partir du rang 10 les représentations graphiques $\mathcal{C}_n$ (tracées en pointillés) sont toutes incluses dans la partie du plan comprise entre $\mathcal{C}_{+\varepsilon}$ et $\mathcal{C}_{-\varepsilon}$ (deux des courbes en traits pleins, la troisième étant la représentation graphique de f).

```
> fn:=(n,x)->exp(-(abs(x)-1/n)^2): f:=x-> exp(-x^2):
> epsilon:=0.1:
> fig1:=plot([f(x),f(x)+epsilon,f(x)-epsilon],x=-1..1,
              color=[red],thickness=[1,2,2]):
> fig2:=plot([seq(fn(n,x),n=[2,4,6,10,14,18])],x=-1..1,
              linestyle=DOT):
> with(plots):display(fig1,fig2);
```



Afin d'illustrer graphiquement la différence entre la convergence simple et la convergence uniforme, considérons les deux suites d'applications $(g_n)_n$ et $(h_n)_n$ de terme général

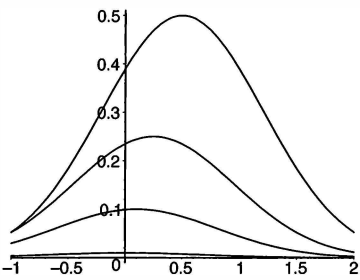
$$g_n : x \in [0,1] \mapsto \frac{1}{n}\exp(-(x-\tfrac{1}{n})^2) \quad \text{et} \quad h_n : x \in [0,1] \mapsto \exp(-n(x-\tfrac{1}{n})^2).$$

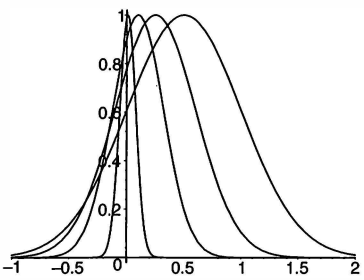
La suite $(g_n)_n$ converge simplement sur $[0,1]$ vers l'application g qui est l'application nulle et la suite $(h_n)_n$ converge simplement sur $[0,1]$ vers l'application

$$h : x \in [0,1] \longmapsto \begin{cases} 0 & \text{si } x > 0 \\ 1 & \text{si } x = 0 \end{cases}.$$

On vérifie sans peine que la suite $(g_n)_n$ converge uniformément vers g sur $[0,1]$ alors que la suite $(h_n)_n$ ne converge pas uniformément vers h sur $[0,1]$. On peut obtenir avec MAPLE les représentations graphiques des deux fonctions $x \mapsto |g(x) - g_n(x)|$ et $x \mapsto |h(x) - h_n(x)|$ pour quelques valeurs de n.

```
> gn:=(n,x)->(1/n)*exp(-(x-1/n)^2): g:= x->0:
> hn:=(n,x)->exp(-n*(x-1/n)^2):
> h:=x-> if (x <> 0) then 0 else 1 end if:
> plot([seq(abs(gn(n,x)-g(x)),n=[2,4,10,100])],x=-1..2);
> plot([seq(abs(hn(n,x)-h(x)),n=[2,4,10,100])],x=-1..2);
```

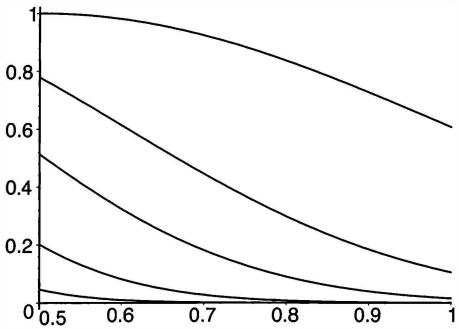




On observera sur la figure de gauche que les représentations graphiques des fonctions $x \mapsto |g(x)-g_n(x)|$ se présentent sous la forme d'une « bosse glissante » se déplaçant vers l'axe des ordonnées en « s'écrasant » sur l'axe des abscisses (cas de convergence uniforme). Les représentations graphiques des fonctions $x \mapsto |h(x) - h_n(x)|$ (à droite) se présentent elles sous la forme d'une « bosse glissante » se déplaçant vers l'axe des ordonnées sans « s'aplatir », en restant de hauteur constante (cas de non-convergence uniforme).

On vérifie qu'en restriction à l'intervalle $[\frac{1}{2}, 1]$ la suite de fonctions $(h_n)_n$ converge uniformément vers l'application nulle.

```
> plot([seq(abs(hn(n,x)-h(x)),n=[2,4,6,10,16])],x=1/2..1);
```



On peut observer qu'en restriction à l'intervalle $[\frac{1}{2}, 1]$, les représentations graphiques des fonctions $x \mapsto |h(x) - h_n(x)|$ « s'aplatissent » vers l'axe des abscisses.

Exercice 2.2 Montrer que la suite d'applications $(f_n)_n$ de terme général $f_n : x \in \mathbb{R} \mapsto x/(1+x^2)^n$ converge uniformément vers l'application nulle sur $\mathbb{R}$.

2.2.4 Propriétés

THÉORÈME 2.1 *Si la suite d'applications $(f_n)_n$ de X dans $\mathbb{R}$ converge uniformément sur X vers l'application f de X dans $\mathbb{R}$ alors elle converge aussi simplement vers f.*

Démonstration Supposons que la suite d'applications $(f_n)_n$ de X dans $\mathbb{R}$ converge uniformément sur X vers l'application f. Soient x_0 un élément de X et ε un réel strictement positif. D'après l'assertion (2) p. 73, il existe un entier N_ε tel que pour tout $x \in X$,

$$\big(n \geqslant N_\varepsilon \implies |f_n(x) - f(x)| \leqslant \varepsilon\big).$$

En particulier, comme $x_0 \in X$ on a : $\big(n \geqslant N_\varepsilon \implies |f_n(x_0) - f(x_0)| \leqslant \varepsilon\big)$.
On en déduit que

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N_\varepsilon \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad \big(n \geqslant N_\varepsilon \implies |f_n(x_0) - f(x_0)| \leqslant \varepsilon\big).$$

Cette assertion étant vraie pour tout $x_0 \in \mathbb{R}$, l'assertion (1) p. 71 est vérifiée et on en conclut que la suite $(f_n)_n$ converge simplement sur X vers f. □

Lors de l'étude de la convergence uniforme d'une suite d'applications de X dans $\mathbb{R}$, on sera attentif aux considérations suivantes.

- S'il existe un réel $x_0 \in X$ tel que la suite numérique de terme général $f_n(x_0)$ diverge alors la suite $(f_n)_n$ ne converge pas simplement sur X. Par conséquent elle ne converge pas non plus uniformément sur X.
- Si la suite $(f_n)_n$ converge simplement sur X vers une application f de X dans $\mathbb{R}$ alors, dans l'hypothèse où elle converge uniformément sur X, elle converge uniformément vers l'application f.
- Il existe de nombreux exemples de suites de fonctions qui convergent simplement sur un ensemble X donné mais pas uniformément sur cet ensemble (on pourra consulter les exemples des pp. 73 à 75). Par ailleurs, une suite de fonctions peut converger uniformément sur un ensemble donné dans $\mathbb{R}$ mais pas sur un autre.

Exercice 2.3
1 - Montrer que pour qu'une suite $(f_n)_n$ d'applications de X dans $\mathbb{R}$ ne converge pas uniformément sur X vers la fonction nulle, il suffit qu'il existe une suite $(x_n)_n$ d'éléments de X telle que la suite de terme général $f_n(x_n)$ ne converge pas vers 0.
2 - a) Montrer que la suite d'applications $(f_n)_n$ de terme général $f_n : x \in [0, \pi] \mapsto \sqrt{n}\cos(x)(\sin(x))^n$ converge simplement vers l'application nulle.
b) En considérant la suite $(x_n)_n$ de terme général $x_n = \arcsin(1 - \frac{1}{n})$, montrer que la suite $(f_n)_n$ ne converge pas uniformément sur $[0, \pi]$.

PROPOSITION 2.3 *Soient $(f_n)_n$ et $(g_n)_n$ deux suites d'applications de X dans $\mathbb{R}$ convergeant simplement (resp. uniformément) sur X vers f et g.*

✗ *La suite d'applications de terme général $f_n + g_n$ converge simplement (resp. uniformément) sur X vers $f + g$.*

✗ *Pour tout $\lambda \in \mathbb{R}$, la suite d'applications de terme général $\lambda \cdot f_n$ converge simplement (resp. uniformément) sur X vers $\lambda \cdot f$.*

Démonstration ⊵ Le cas de la convergence simple résulte des propriétés des suites numériques[5] car une suite $(f_n)_n$ d'applications de X dans $\mathbb{R}$ converge

[5] Voir la proposition 5.8 p. 175 du *Cours de première année*.

simplement sur X si et seulement si pour tout $x \in X$ la suite numérique de terme général $f_n(x)$ converge vers $f(x)$.

$\trianglerighteq$ Supposons que les deux suites d'applications $(f_n)_n$ et $(g_n)_n$ de X dans $\mathbb{R}$ convergent uniformément sur X vers les applications f et g. D'après la proposition 2.2, les suites numériques de terme général $u_n = \|f - f_n\|_{\infty,X}$ et $v_n = \|g - g_n\|_{\infty,X}$ convergent vers 0. En utilisant la troisième relation de la proposition 2.1, on obtient pour tout $n \in \mathbb{N}$,

$$\begin{aligned} \|(f+g) - (f_n + g_n)\|_{\infty,X} &= \|(f - f_n) + (g - g_n)\|_{\infty,X} \\ &\leqslant \|f - f_n\|_{\infty,X} + \|g - g_n\|_{\infty,X}. \end{aligned}$$

On déduit de cette inégalité que la suite numérique dont le terme général est $w_n = \|(f+g) - (f_n + g_n)\|_{\infty,X}$ converge vers 0. La proposition 2.2 permet de conclure que la suite d'applications de terme général $f_n + g_n$ converge uniformément sur X vers $f + g$.

$\trianglerighteq$ Pour démontrer que pour tout $\lambda \in \mathbb{R}$ la suite d'applications de terme général $\lambda \cdot f_n$ converge uniformément sur X vers $\lambda \cdot f$, on procède de manière analogue en ayant recours cette fois à la seconde relation de la proposition 2.1. □

Comme la suite nulle converge simplement (resp. uniformément), la proposition 2.3 exprime le fait que l'ensemble des suites d'applications de X dans $\mathbb{R}$ qui convergent simplement (resp. uniformément) est un sous-espace vectoriel de $\mathcal{S}(X, \mathbb{R})$.

Convergence uniforme et applications bornées

Si la suite $(f_n)_n$ converge uniformément vers f sur X, alors à partir d'un certain rang toutes les applications $f_n - f$ sont bornées car d'après l'assertion (3), on a

$$\exists N_1 \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad (n \geqslant N_1 \implies \|f_n - f\|_{\infty,X} \leqslant 1). \tag{4}$$

Par contre, si une suite d'applications $(f_n)_n$ de X dans $\mathbb{R}$ converge uniformément sur X vers une application f de X dans $\mathbb{R}$, il se peut qu'aucune des applications f_n ne soit bornée (seules les applications $f_n - f$ sont bornées). Dans ce cas, l'application f n'est pas bornée.

Soit $(f_n)_n$ une suite d'éléments de l'ensemble $\mathcal{B}(X, \mathbb{R})$ des applications de X dans $\mathbb{R}$ bornées[6] ; on a

$$\forall n \in \mathbb{N} \quad \exists M_n \in \mathbb{R}^+ \quad \forall x \in X \quad |f_n(x)| \leqslant M_n.$$

Si la suite $(f_n)_n$ converge uniformément sur X vers une application f de X dans $\mathbb{R}$, alors l'application f est nécessairement bornée car en considérant l'entier N_1 défini par (4), on a

$$\|f\|_{\infty,X} \leqslant \|f - f_{N_1}\|_{\infty,X} + \|f_{N_1}\|_{\infty,X} \leqslant 1 + M_{N_1}.$$

[6] Voir le *Cours de première année* p. 566.

Il n'existe pas de suite d'applications non bornées qui converge uniformément vers une application bornée (mais ceci peut être vrai dans le cas de la convergence simple).

L'exercice 2.4 illustre ces différents comportements.

Exercice 2.4
1 - Montrer que les applications $f_n : x \in \mathbb{R} \mapsto \min(|x|, n)$ sont bornées et que la suite $(f_n)_n$ converge simplement vers une application qui n'est pas bornée. La convergence de la suite $(f_n)_n$ est-elle uniforme ?
2 - Montrer que les applications $f_n : x \in \mathbb{R} \mapsto x + \frac{1}{n}$ ne sont pas bornées mais que la suite $(f_n)_n$ converge uniformément sur $\mathbb{R}$. La fonction limite est-elle une application bornée ?

Convergence du produit de deux suites d'applications

PROPOSITION 2.4

✗ *Si $(f_n)_n$ et $(g_n)_n$ sont deux suites d'applications de X dans $\mathbb{R}$ qui convergent simplement sur X vers f et g respectivement, alors la suite d'applications de terme général $f_n \times g_n$ converge simplement sur X vers $f \times g$.*

✗ *Si $(f_n)_n$ et $(g_n)_n$ sont deux suites d'applications de $\mathcal{B}(X, \mathbb{R})$ qui convergent uniformément sur X vers f et g respectivement alors la suite d'applications de terme général $f_n \times g_n$ converge uniformément sur X vers $f \times g$.*

Démonstration $\trianglerighteq$ Les deux suites d'applications $(f_n)_n$ et $(g_n)_n$ convergent simplement sur X vers les applications f et g si pour tout $x \in X$ fixé, les suites de réels $(f_n(x))_{n\in\mathbb{N}}$ et $(g_n(x))_{n\in\mathbb{N}}$ ont pour limites les réels $f(x)$ et $g(x)$. Il résulte des propriétés des suites numériques[7] que la suite réelle de terme général $f_n(x)\, g_n(x)$ converge vers $f(x)\, g(x)$. Par définition de la convergence simple, ceci signifie que la suite d'applications de terme général $f_n \times g_n$ converge simplement sur X vers $f \times g$.

$\trianglerighteq$ Supposons que les deux suites d'applications bornées $(f_n)_n$ et $(g_n)_n$ qui convergent uniformément sur X vers f et g. D'après la proposition 2.2, les suites numériques de terme général $u_n = \|f - f_n\|_{\infty,X}$ et $v_n = \|g - g_n\|_{\infty,X}$ convergent vers 0. En utilisant la proposition 2.1, on obtient pour tout $n \in \mathbb{N}$,

$$\begin{aligned}\|(f \times g) - (f_n \times g_n)\|_{\infty,X} &= \|f \times (g - g_n) + g \times (f - f_n) \\ &\qquad - (f - f_n) \times (g - g_n)\|_{\infty,X} \\ &\leqslant \|f - f_n\|_{\infty,X} \times \|g - g_n\|_{\infty,X} + \|f\|_{\infty,X} \times \|g - g_n\|_{\infty,X} \\ &\qquad + \|g\|_{\infty,X} \times \|f - f_n\|_{\infty,X}.\end{aligned}$$

[7] Voir la proposition 5.8 p. 175 du *Cours de première année*.

Puisque $(f_n)_n$ (resp. $(g_n)_n$) est une suite d'applications bornées de X dans $\mathbb{R}$ qui converge uniformément sur X vers f (resp. g), l'application f (resp. g) est bornée (voir la discussion précédente). Le théorème d'encadrement permet d'en déduire que la suite numérique $(w_n)_n$ dont le terme général est $w_n = \|(f \times g) - (f_n \times g_n)\|_{\infty,X}$ converge vers 0. La proposition 2.2 permet alors de conclure que la suite d'applications de terme général $f_n \times g_n$ converge uniformément sur X vers $f \times g$. □

2.2.5 Critère de Cauchy pour la convergence uniforme

DÉFINITION 2.3 (Critère de Cauchy uniforme)
Une suite $(f_n)_n$ d'applications de X dans $\mathbb{R}$ est qualifiée de suite d'applications uniformément de Cauchy *sur X (ou* suite d'applications vérifiant uniformément le critère de Cauchy *sur X) si*

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N_\varepsilon \in \mathbb{N} \quad \forall (m,n) \in \mathbb{N}^2 \quad \boxed{\forall x \in X} \quad (n \geqslant m \geqslant N_\varepsilon \implies |f_n(x) - f_m(x)| \leqslant \varepsilon). \qquad (5)$$

Remarque L'assertion quantifiée (5) est équivalente à :

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N_\varepsilon \in \mathbb{N} \quad \forall (m,n) \in \mathbb{N}^2 \quad (n \geqslant m \geqslant N_\varepsilon \implies \|f_n - f_m\|_{\infty,X} \leqslant \varepsilon).$$

On peut aussi exprimer qu'une suite d'applications $(f_n)_n$ est uniformément de Cauchy sur X de la manière suivante :

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N_\varepsilon \in \mathbb{N} \quad \forall (n,p) \in \mathbb{N}^2 \quad \forall x \in X \quad (n \geqslant N_\varepsilon \implies |f_{n+p}(x) - f_n(x)| \leqslant \varepsilon).$$

◇

THÉORÈME 2.2 *Une condition nécessaire et suffisante pour qu'une suite d'applications de X dans $\mathbb{R}$ converge uniformément sur X est que cette suite soit uniformément de Cauchy sur X.*

Démonstration ⊵ Considérons une suite $(f_n)_n$ d'applications de X dans $\mathbb{R}$ qui converge uniformément sur X vers une application f de X dans $\mathbb{R}$. Soit ε un réel strictement positif ; d'après l'assertion (3) p. 74, il existe $N \in \mathbb{N}$ tel que[8] :

$$\forall p \in \mathbb{N} \quad (p \geqslant N \implies \|f_p - f\|_{\infty,X} \leqslant \tfrac{1}{2}\varepsilon). \qquad (6)$$

Si m et n sont deux entiers naturels vérifiant $n \geqslant m \geqslant N$, on a d'après la relation (6) (considérée successivement avec $p = n$ et $p = m$) :

$$\|f_n - f\|_{\infty,X} \leqslant \tfrac{1}{2}\varepsilon \quad \text{et} \quad \|f_m - f\|_{\infty,X} \leqslant \tfrac{1}{2}\varepsilon.$$

[8] On considère l'assertion (3) p. 74 avec $\varepsilon = \frac{1}{2}\varepsilon$ (sic).

Puisque[9] $\|f_n - f_m\|_{\infty,X} \leqslant \|f_n - f\|_{\infty,X} + \|f_m - f\|_{\infty,X}$, on en déduit que $\|f_n - f_m\|_{\infty,X} \leqslant \varepsilon$. On a ainsi montré que :

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N_\varepsilon \in \mathbb{N} \quad \forall (m,n) \in \mathbb{N}^2 \quad (n \geqslant m \geqslant N_\varepsilon \implies \|f_n - f_m\|_{\infty,X} \leqslant \varepsilon),$$

ce qui prouve que la suite $(f_n)_n$ est donc uniformément de Cauchy sur X.

$\trianglerighteq$ Réciproquement, considérons une suite $(f_n)_n$ qui soit uniformément de Cauchy sur X, c'est-à-dire telle que :

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N_\varepsilon \in \mathbb{N} \quad \forall (m,n) \in \mathbb{N}^2 \quad \forall x \in X \quad (n \geqslant m \geqslant N_\varepsilon \implies |f_n(x) - f_m(x)| \leqslant \varepsilon). \quad (7)$$

Commençons par établir que la suite $(f_n)_n$ converge simplement sur X. De l'assertion (7), on déduit que[10] :

$$\forall x \in X \quad \forall \varepsilon \in \mathbb{R}_+^* \quad \exists N_\varepsilon \in \mathbb{N} \quad \forall (m,n) \in \mathbb{N}^2 \quad (n \geqslant m \geqslant N_\varepsilon \implies |f_n(x) - f_m(x)| \leqslant \varepsilon),$$

c'est à dire que pour tout $x \in X$ la suite numérique $(f_n(x))_n$ est une suite numérique de Cauchy[11]. Ceci implique que la suite numérique $(f_n(x))_n$ converge[12]. On peut définir une application f de X dans $\mathbb{R}$ qui au réel $x \in X$ associe la limite de la suite $(f_n(x))_n$.

Montrons que la suite $(f_n)_n$ converge uniformément sur X vers l'application f. Soient ε un réel strictement positif et $(m,n) \in \mathbb{N}^2$ tels que $n \geqslant m \geqslant N_\varepsilon$ où l'entier N_ε est défini par l'assertion (7). Pour tout $x \in X$, on a

$$|f_n(x) - f_m(x)| \leqslant \varepsilon \quad \textit{i.e.} \quad f_m(x) - \varepsilon \leqslant f_n(x) \leqslant f_m(x) + \varepsilon.$$

Cette inégalité est vraie pour tout $n \geqslant N_\varepsilon$; puisque la suite numérique $(f_n(x))_n$ converge, l'inégalité est respectée par sa limite[13]. Cette limite étant $f(x)$, pour tout $x \in X$ on a

$$f(x) - \varepsilon \leqslant f_n(x) \leqslant f(x) + \varepsilon \quad \textit{i.e.} \quad |f_m(x) - f(x)| \leqslant \varepsilon.$$

On a donc montré que

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N_\varepsilon \in \mathbb{N} \quad \forall m \in \mathbb{N} \quad \forall x \in X \quad (m \geqslant N_\varepsilon \implies |f_m(x) - f(x)| \leqslant \varepsilon),$$

ce qui permet de conclure que la suite $(f_n)_n$ converge uniformément sur X vers f. □

---

[9] Voir la proposition 2.1 p. 68.

[10] Noter la place du bloc quantifié $\forall x \in X$.

[11] Voir la proposition 5.8 p.191 du *Cours de première année*.

[12] Voir le théorème 5.4 p.191 du *Cours de première année* ; on notera qu'il y a équivalence entre la suite numérique $(f_n(x))_n$ est une suite numérique de Cauchy et la suite numérique $(f_n(x))_n$ converge.

[13] Voir la proposition 5.12 p. 178 du *Cours de première année*.

On sera vigilant à la place du quantificateur « $\forall x \in X$ » dans l'assertion (5). En particulier, l'assertion

$$\forall x \in X \quad \forall \varepsilon \in \mathbb{R}_+^* \quad \exists N_\varepsilon \in \mathbb{N} \quad \forall (m,n) \in \mathbb{N}^2 \\ (n \geqslant m \geqslant N_\varepsilon \implies |f_n(x) - f_m(x)| \leqslant \varepsilon), \quad (8)$$

qui a bien un sens d'un point de vue mathématique, n'est pas équivalente à l'assertion (5). Elle exprime le fait que pour tout $x \in X$ la suite numérique de terme général $f_n(x)$ est une suite numérique de Cauchy[14]. On sait[15] que ceci équivaut au fait que pour tout $x \in X$ la suite numérique converge. Pour la suite de fonctions $(f_n)_n$, cette assertion est donc équivalente à la convergence simple sur X (et non à la convergence uniforme).

Exercice 2.5 On considère la suite de fonctions $(f_n)_n$ définies par la relation

$$f_n(x) = (-1)^n \ln\left(1 + \frac{x}{n(1+x)}\right) \qquad \forall n \in \mathbb{N}^*.$$

1 - Déterminer pour $n \in \mathbb{N}^*$, l'ensemble de définition $\mathcal{D}_n$ de f_n. En déduire l'ensemble $\mathcal{D}$ sur lequel est définie la suite de fonctions $(f_n)_n$.
2 - Montrer que la suite de fonctions $(f_n)_{n\in\mathbb{N}^*}$ converge simplement sur $\mathcal{D}$ et déterminer la limite f de la suite.
3 - Étudier la convergence uniforme de la suite de fonctions $(f_n)_{n\in\mathbb{N}^*}$ sur $\mathcal{D}$ puis sur $]-\frac{1}{2}, +\infty[$.

2.3 Propriétés de la limite d'une suite d'applications réelles

2.3.1 Limite simple d'une suite d'applications

Certaines propriétés que possèdent les termes d'une suite d'applications sont conservées par la limite de cette suite.

Proposition 2.5 *Soient I un intervalle de $\mathbb{R}$ et $(f_n)_n$ une suite d'applications de I dans $\mathbb{R}$* convergeant simplement sur I *vers une application f de I dans $\mathbb{R}$.*

✗ *Si pour tout $n \in \mathbb{N}$, l'application f_n est croissante (resp. décroissante) sur I alors f est croissante (resp. décroissante) sur I.*

✗ *Si pour tout $n \in \mathbb{N}$, l'application f_n est convexe[16] (resp. concave) sur I alors f est convexe (resp. concave) sur I.*

[14] Voir la définition 5.8 p. 191 du *Cours de première année*.
[15] Voir le théorème 5.4 p. 191 du *Cours de première année*.

Démonstration $\trianglerighteq$ Considérons une suite d'applications $(f_n)_n$ de I dans $\mathbb{R}$ convergeant simplement vers une application f de I dans $\mathbb{R}$ et supposons que pour tout $n \in \mathbb{N}$ l'application f_n est croissante[17] sur I, c'est-à-dire qu'elle vérifie :

$$\forall (x,y) \in I^2 \quad \big(x \leqslant y \implies f_n(x) \leqslant f_n(y)\big).$$

On veut montrer que f est croissante sur I, c'est-à-dire que :

$$\forall (x,y) \in I^2 \quad \big(x \leqslant y \implies f(x) \leqslant f(y)\big).$$

Considérons deux réels x et y appartenant à I tels que $x \leqslant y$. D'après les hypothèses, $f_n(x) \leqslant f_n(y)$ pour tout $n \in \mathbb{N}$ et les suites numériques $(f_n(x))_n$ et $(f_n(y))_n$ convergent vers $f(x)$ et $f(y)$. Il résulte du théorème d'encadrement[18] que les limites de ces deux suites vérifient $f(x) \leqslant f(y)$. L'application f est donc bien croissante.

Le cas où les applications f_n sont décroissantes se traite d'une manière analogue. On peut aussi déduire ce résultat de celui que l'on vient de montrer en utilisant le fait que si une application est croissante, son opposée est décroissante.

$\trianglerighteq$ Considérons une suite d'applications de I dans $\mathbb{R}$ convergeant simplement vers une application f de I dans $\mathbb{R}$ et supposons que pour tout $n \in \mathbb{N}$, l'application f_n est convexe sur I, c'est-à-dire qu'elle vérifie :

$$\forall (x,y) \in I^2 \quad \forall \lambda \in [0,1] \quad f_n(\lambda x + (1-\lambda)y) \leqslant \lambda f_n(x) + (1-\lambda) f_n(y).$$

Montons que f est convexe, c'est-à-dire que :

$$\forall (x,y) \in I^2 \quad \forall \lambda \in [0,1] \quad f(\lambda x + (1-\lambda)y) \leqslant \lambda f(x) + (1-\lambda) f(y).$$

Considérons deux réels x et y dans I, un réel λ dans $[0,1]$ et notons $z = \lambda x + (1-\lambda)y$. On a $z \in I$ et d'après les hypothèses, la suite numérique $(f_n(z))_n$ converge vers $f(z)$. Par ailleurs, puisque les suites numériques $(f_n(x))_n$ et $(f_n(y))_n$ convergent vers $f(x)$ et $f(y)$, la suite numérique de terme général $u_n = \lambda f_n(x) + (1-\lambda) f_n(y)$ converge vers $\lambda f(x) + (1-\lambda) f(y)$. Il résulte du théorème d'encadrement[17] que les limites de ces deux suites vérifient : $f(z) \leqslant \lambda f(x) + (1-\lambda) f(y)$. On en conclut que f est convexe sur I.

Le cas où les applications f_n sont concaves se déduit en utilisant le fait que si une application est concave, son opposée est convexe. □

---

[16] Voir la définition 16.7 p. 737 du *Cours de première année*.

[17] Voir les définitions p. 563 du *Cours de première année*.

[18] Voir le théorème 5.1 p. 179 du *Cours de première année*.

On sera vigilant au fait que ce sont les applications f_n qui sont supposées être monotones et non pas la suite $(f_n)_n$. On sera également attentif au fait qu'une suite de fonctions qui sont toutes strictement croissantes peut converger vers une fonction qui n'est pas strictement croissante. C'est le cas par exemple de la suite $(f_n)_n$ de terme général $f_n : x \in [0,1] \mapsto x/n$; pour tout entier $n \in \mathbb{N}^*$, l'application f_n est strictement croissante, mais la suite $(f_n)_n$ converge vers l'application nulle qui n'est pas strictement croissante.

Si certaines des propriétés que possèdent tous les termes d'une suite de fonctions sont conservées par la fonction qui en est limite, il n'en va pas ainsi pour toutes les propriétés. Nous avons vu p. 71 que la limite de la suite d'applications de terme général $f_n : x \in [0,+\infty[\mapsto \mathrm{e}^{-nx}$ n'était pas une application continue alors que toutes les applications f_n sont continues. Autrement dit, une suite d'applications continues peut converger simplement vers une application qui n'est pas continue. La propriété de continuité n'est pas toujours conservée lors d'une convergence simple. Il en va de même des propriétés de dérivabilité et d'intégrabilité.

Nous allons donner dans les paragraphes suivants des conditions suffisantes pour que soient conservées par passage à la limite les propriétés de continuité, de dérivabilité et d'intégrabilité.

2.3.2 Limite uniforme d'une suite d'applications continues

THÉORÈME 2.3 *Soient I un intervalle ouvert de $\mathbb{R}$ et $(f_n)_n$ une suite d'applications de I dans $\mathbb{R}$* convergeant uniformément sur I *vers l'application f de I dans $\mathbb{R}$.*

✗ *Si pour tout $n \in \mathbb{N}$ l'application f_n est continue en $x_0 \in I$ alors f est continue en x_0.*

✗ *Si pour tout $n \in \mathbb{N}$ l'application f_n est continue sur I alors f est continue sur I.*

Démonstration Une application est continue sur l'intervalle I si pour tout $x_0 \in I$ elle est continue en x_0. La deuxième assertion du théorème est donc une conséquence immédiate de la première assertion.

La suite $(f_n)_n$ étant une suite d'applications de I dans $\mathbb{R}$ convergeant uniformément sur I vers l'application f, on a

$$\forall \varepsilon_2 \in \mathbb{R}_+^* \ \exists N_2 \in \mathbb{N} \ \forall x \in I \ \forall n \in \mathbb{N} \quad (n \geqslant N_2 \implies |f_n(x) - f(x)| \leqslant \varepsilon_2). \quad (9)$$

De plus, pour tout $n \in \mathbb{N}$, l'application f_n étant continue en $x_0 \in I$, on a[19] :

$$\forall n \in \mathbb{N} \ \forall \varepsilon_1 \in \mathbb{R}_+^* \ \exists \eta_1 \in \mathbb{R}_+^* \ \forall x \in I \quad (|x - x_0| \leqslant \eta_1 \implies |f_n(x) - f_n(x_0)| \leqslant \varepsilon_1). \quad (10)$$

[19] Voir la définition 13.4 p. 588 du *Cours de première année*.

Pour montrer que f est continue en x_0, il faut montrer que[19] :

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists \eta \in \mathbb{R}_+^* \quad \forall x \in I \quad (|x - x_0| \leqslant \eta \implies |f(x) - f(x_0)| \leqslant \varepsilon). \quad (11)$$

Pour ε réel strictement positif fixé, déterminons un réel η strictement positif tel que pour tout réel $x \in I$ avec $|x - x_0| \leqslant \eta$ on ait $|f(x) - f(x_0)| \leqslant \varepsilon$.

Prenons $\varepsilon_2 = \frac{1}{3}\varepsilon$ dans l'assertion (9) : il existe $N_2 \in \mathbb{N}$ tel que pour tout réel $x \in I$ et pour tout entier $n \geqslant N_2$ on ait $|f_n(x) - f(x)| \leqslant \frac{1}{3}\varepsilon$. Pour $n \geqslant N_2$ considérons à présent l'assertion (10) avec $\varepsilon_1 = \frac{1}{3}\varepsilon$: il existe un réel η_1 strictement positif tel que pour tout réel x vérifiant $|x - x_0| \leqslant \eta_1$ on ait $|f_n(x) - f_n(x_0)| \leqslant \frac{1}{3}\varepsilon$.

On en déduit que pour $n \geqslant N_2$ et $x \in I$ tel que $|x - x_0| \leqslant \eta_1$ on a

$$\begin{aligned} |f(x) - f(x_0)| &= |f(x) - f_n(x) + f_n(x) - f_n(x_0) + f_n(x_0) - f(x_0)| \\ &\leqslant |f(x) - f_n(x)| + |f_n(x) - f_n(x_0)| + |f_n(x_0) - f(x_0)|. \end{aligned}$$

Or, on a les majorations suivantes des différents termes du membre de droite :

- $|f(x) - f_n(x)| \leqslant \frac{1}{3}\varepsilon$ d'après (9) ;
- $|f_n(x) - f_n(x_0)| \leqslant \frac{1}{3}\varepsilon$ d'après (10) ;
- $|f_n(x_0) - f(x_0)| \leqslant \frac{1}{3}\varepsilon$ d'après (9) considérée avec le réel x_0.

On obtient ainsi,

$$|f(x) - f(x_0)| \leqslant \tfrac{1}{3}\varepsilon + \tfrac{1}{3}\varepsilon + \tfrac{1}{3}\varepsilon = \varepsilon.$$

L'assertion (11) est démontrée : le réel $\eta = \eta_1$ convient. □

Remarques

1. Le théorème 2.3 exprime le fait que si $(f_n)_n$ est une suite d'applications continues de I dans $\mathbb{R}$ convergeant uniformément sur I alors la limite de cette suite est une application de I dans $\mathbb{R}$ continue sur I. On peut noter que pour que f soit continue en x_0, il suffit que pour tout $n \in \mathbb{N}$ l'application f_n soit continue en $x_0 \in I$ et que la suite $(f_n)_n$ converge uniformément vers f sur un voisinage de x_0 (et pas nécessairement sur I en entier).

2. On peut remplacer dans l'énoncé du théorème 2.3 le terme « continue » par « continue à droite » ou par « continue à gauche »[20].

3. Le résultat du théorème 2.3 reste vrai si les applications f_n ne sont continues qu'à partir d'un certain rang.

4. La convergence uniforme d'une suite d'applications continues est une condition suffisante pour que l'application limite soit continue. Elle n'est toutefois pas nécessaire. Nous avons montré p. 74, que la suite d'applications continues $(f_n)_n$ de terme général $f_n : x \in]0, +\infty[\mapsto \mathrm{e}^{-nx}$ convergeait simplement mais pas uniformément vers la fonction nulle sur $]0, +\infty[$. Pourtant l'application nulle est bien continue sur $]0, +\infty[$. ◇

[20] Voir la définition 13.5 p. 590 du *Cours de première année*.

Par contraposition, on déduit du théorème 2.3 un critère permettant d'affirmer que la convergence d'une suite d'applications n'est pas une convergence uniforme.

COROLLAIRE 2.2 *Soient I un intervalle ouvert de $\mathbb{R}$ et $(f_n)_n$ une suite d'applications continues de I dans $\mathbb{R}$ qui converge simplement sur I vers une application f de I dans $\mathbb{R}$. Si f n'est pas une application continue sur I alors la suite $(f_n)_n$ ne converge pas uniformément sur I.*

Exemple Soit $(f_n)_n$ la suite de terme général $f_n : x \in [0, +\infty[\mapsto \mathrm{e}^{-nx}$. Nous avons établi p. 71 que cette suite converge simplement sur $[0, +\infty[$ vers :

$$f : x \in [0, +\infty[\longmapsto \begin{cases} 0 & \text{si } x > 0 \\ 1 & \text{si } x = 0 \end{cases} .$$

L'application f n'étant pas continue en 0, la suite $(f_n)_n$ ne converge pas uniformément vers f sur $[0, +\infty[$ (ce que nous avions établi par un autre procédé p. 74).

Exercice 2.6 Soit $(f_n)_n$ une suite d'applications continues sur un intervalle I et à valeurs dans $\mathbb{R}$ qui converge uniformément sur I vers une application f de I dans $\mathbb{R}$. Montrer que si $(x_n)_n$ est une suite de réels de l'intervalle I convergeant vers $x \in I$ alors $\lim_{n\to+\infty} f_n(x_n) = f(x)$. (On raisonnera à l'aide d'assertions quantifiées.)

On déduit du théorème 2.3 le résultat suivant qui donne une condition suffisante pour pouvoir permuter deux limites.

PROPOSITION 2.6 (Permutation de limites)
Soient x_0 un réel, X un voisinage de x_0 et $(f_n)_n$ une suite d'applications de X dans $\mathbb{R}$ qui converge uniformément sur X. Si à partir d'un certain rang les applications f_n admettent une limite en x_0 alors :

$$\lim_{x\to x_0} \lim_{n\to+\infty} f_n(x) = \lim_{n\to+\infty} \lim_{x\to x_0} f_n(x).$$

On a également le résultat suivant qui est l'analogue de la proposition 2.6 dans le cas d'une limite en $+\infty$ (on a un résultat similaire en $-\infty$). Il se démontre sur de la même manière que le théorème 2.3 en prenant la définition idoine de la limite.

PROPOSITION 2.7 (Permutation de limites)
Soient a un réel et $(f_n)_n$ une suite d'applications de $]a,+\infty[$ dans $\mathbb{R}$ qui converge uniformément sur $]a,+\infty[$. Si à partir d'un certain rang les applications f_n admettent une limite en $+\infty$ alors,

$$\lim_{x\to+\infty}\lim_{n\to+\infty} f_n(x) = \lim_{n\to+\infty}\lim_{x\to+\infty} f_n(x)$$

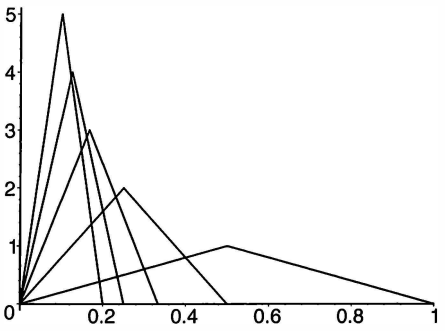
2.3.3 Limite uniforme d'une suite d'applications intégrables

Considérons une suite $(f_n)_n$ d'applications de $[a,b]$ dans $\mathbb{R}$, Riemann-intégrables sur $[a,b]$ et convergeant simplement sur $[a,b]$ vers une application f. On se pose la question suivante : l'intégrale entre a et b de f, c'est-à-dire l'intégrale entre a et b de la limite de la suite $(f_n)_n$, est-elle égale à la limite lorsque n tend vers l'infini de l'intégrale entre a et b de f_n ? Autrement dit, a-t-on toujours :

$$\int_a^b \lim_{n\to+\infty} f_n(x)\,\mathrm{d}x = \lim_{n\to+\infty}\int_a^b f_n(x)\,\mathrm{d}x \quad ? \tag{12}$$

On peut se convaincre que cette égalité n'a pas toujours lieu en considérant la suite $(f_n)_n$ où f_n est l'application définie sur $[0,1]$ par $f_n(x) = 2n^2x$ pour $x \in [0, \frac{1}{2n}[$, $f_n(x) = -2n(nx-1)$ pour $x \in [\frac{1}{2n}, \frac{1}{n}[$ et $f_n(x) = 0$ pour $x \in [\frac{1}{n}, 1]$.
Les premiers termes de la suite $(f_n)_n$ peuvent aisément être représentés à l'aide de MAPLE.

```
> fn:=(n,x)->piecewise(x>=0 and x  < 1/(2*n), 2*n**2*x,
                       x>=1/(2*n) and x < 1/n, -2*n*(n*x-1),0):
> plot([seq(fn(i,x),i=1..5)],x=0..1);
```



On vérifie que cette suite converge simplement sur $[0,1]$ vers l'application nulle ; on a donc $\int_0^1 \lim_{n\to+\infty} f_n(x)\,\mathrm{d}x = 0$. Par ailleurs,

$$\int_0^1 f_n(x)\,\mathrm{d}x = \int_0^{\frac{1}{2n}} 2n^2x\,\mathrm{d}x - 2n\int_{\frac{1}{2n}}^{\frac{1}{n}} (nx-1)\,\mathrm{d}x = \frac{1}{2}.$$

On en déduit que $\lim_{n\to+\infty} \int_0^1 f_n(x)\,\mathrm{d}x = \frac{1}{2} \neq \int_0^1 \lim_{n\to+\infty} f_n(x)\,\mathrm{d}x$.

Le théorème suivant indique qu'une condition suffisante pour avoir l'égalité (12) est que la convergence de la suite $(f_n)_n$ vers f soit uniforme sur $[a,b]$.

THÉORÈME 2.4 *Soit $(f_n)_n$ une suite d'applications de $[a,b]$ dans $\mathbb{R}$, Riemann-intégrables sur $[a,b]$. Si la suite $(f_n)_n$ converge uniformément sur $[a,b]$ vers l'application f alors f est Riemann-intégrable sur $[a,b]$ et*

$$\int_a^b f(x)\,\mathrm{d}x = \lim_{n\to+\infty} \int_a^b f_n(x)\,\mathrm{d}x.$$

Démonstration ⊵ Pour montrer que l'application f est Riemann-intégrable sur $[a,b]$, il faut montrer[21] que pour tout $\varepsilon \in \mathbb{R}_+^*$, il existe un couple $(\phi_\varepsilon, \psi_\varepsilon)$ de fonctions en escalier sur $[a,b]$ tel que : $\psi_\varepsilon \leqslant f \leqslant \phi_\varepsilon$ sur $[a,b]$ et $\int_a^b (\phi_\varepsilon - \psi_\varepsilon) < \varepsilon$.

Soit ε un réel strictement positif fixé ; par hypothèse, la suite d'applications $(f_n)_n$ convergeant uniformément sur $[a,b]$ vers l'application f, on a[22]

$$\exists N \in \mathbb{N} \quad \forall x \in [a,b] \quad \forall n \in \mathbb{N} \quad \left(n \geqslant N \implies |f_n(x) - f(x)| \leqslant \frac{\varepsilon}{3(b-a)}\right). \tag{13}$$

L'application f_N étant Riemann-intégrable[22] sur $[a,b]$, il existe un couple $(\phi_{N,\varepsilon}, \psi_{N,\varepsilon})$ de fonctions en escalier sur $[a,b]$ tel que :

$$\psi_{N,\varepsilon} \leqslant f_N \leqslant \phi_{N,\varepsilon} \text{ sur } [a,b] \quad \text{et} \quad \int_a^b (\phi_{N,\varepsilon} - \psi_{N,\varepsilon}) < \tfrac{1}{3}\varepsilon. \tag{14}$$

Pour tout $x \in [a,b]$, on a d'après (13),

$$-\frac{1}{3(b-a)}\,\varepsilon \leqslant f_N(x) - f(x) \leqslant \frac{1}{3(b-a)}\,\varepsilon. \tag{15}$$

On déduit alors de (14) et (15) que

$$f(x) = f(x) - f_N(x) + f_N(x) \leqslant \frac{1}{3(b-a)}\,\varepsilon + f_N(x) \leqslant \frac{1}{3(b-a)}\,\varepsilon + \phi_{N,\varepsilon}(x)$$

et que

$$f(x) = f(x) - f_N(x) + f_N(x) \geqslant -\frac{1}{3(b-a)}\,\varepsilon + f_N(x) \geqslant -\frac{1}{3(b-a)}\,\varepsilon + \psi_{N,\varepsilon}(x).$$

[21] Voir la définition 18.4 p. 836 du *Cours de première année*.

[22] On prend $\varepsilon = \frac{\varepsilon}{3(b-a)}$ (sic) dans la définition 2.2 p. 73.

Les fonctions $\phi_\varepsilon = \frac{1}{3(b-a)}\varepsilon + \phi_{N,\varepsilon}$ et $\psi_\varepsilon = -\frac{1}{3(b-a)}\varepsilon + \psi_{N,\varepsilon}$ sont deux fonctions en escalier sur $[a,b]$ vérifiant $\psi_\varepsilon \leqslant f \leqslant \phi_\varepsilon$ sur $[a,b]$. Par ailleurs,

$$\begin{aligned}\int_a^b (\phi_\varepsilon - \psi_\varepsilon) &= \int_a^b \left(\frac{2\varepsilon}{3(b-a)} + \phi_{N,\varepsilon} - \psi_{N,\varepsilon}\right) = \tfrac{2}{3}\varepsilon + \int_a^b (\phi_{N,\varepsilon} - \psi_{N,\varepsilon}) \\ &< \tfrac{2}{3}\varepsilon + \tfrac{1}{3}\varepsilon = \varepsilon.\end{aligned}$$

On en conclut que la fonction f est Riemann-intégrable sur $[a,b]$.

$\trianglerighteq$ Pour tout $n \in \mathbb{N}$, on a

$$\begin{aligned}0 &\leqslant \left|\int_a^b f(x)\,\mathrm{d}x - \int_a^b f_n(x)\,\mathrm{d}x\right| = \left|\int_a^b (f(x) - f_n(x))\,\mathrm{d}x\right| \\ &\leqslant \int_a^b |f(x) - f_n(x)|\,\mathrm{d}x \leqslant \int_a^b \sup_{t\in[a,b]} |f(t) - f_n(t)|\,\mathrm{d}x \\ &= (b-a) \sup_{t\in[a,b]} |f(t) - f_n(t)|.\end{aligned}$$

Comme la suite d'applications $(f_n)_n$ converge uniformément sur $[a,b]$ vers l'application f, on a $\lim_{n\to+\infty} \sup_{t\in[a,b]} |f(t) - f_n(t)| = 0$. On en déduit par le théorème d'encadrement que $\lim_{n\to+\infty} \left|\int_a^b f(x)\,\mathrm{d}x - \int_a^b f_n(x)\,\mathrm{d}x\right| = 0$, autrement dit que

$$\lim_{n\to+\infty} \int_a^b f_n(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x.$$

Le théorème est démontré. □

Remarques

1. Le théorème 2.4 donne une condition suffisante pour « intervertir les symboles d'intégration et de limite » : pour pouvoir affirmer que

$$\int_a^b \lim_{n\to+\infty} f_n(x)\,\mathrm{d}x = \lim_{n\to+\infty} \int_a^b f_n(x)\,\mathrm{d}x,$$

il suffit que la suite $(f_n)_n$ d'applications Riemann-intégrables sur $[a,b]$ converge uniformément sur $[a,b]$.

2. La convergence uniforme de la suite $(f_n)_n$ est une condition suffisante « pour intervertir les symboles d'intégration et de limite » mais cette condition n'est pas nécessaire. On peut en effet montrer (c'est le théorème de la convergence dominée de Lebesgue[23]) que si la suite $(f_n)_n$ converge simplement vers f sur $[a,b]$ et si

$$\exists M \in \mathbb{R}_+^* \quad \forall n \in \mathbb{N} \quad \forall x \in [a,b] \quad |f_n(x)| \leqslant M$$

[23] Lebesgue, Henri (1875, Beauvais - 1941, Paris). Auteur de travaux novateurs en analyse, Henri Lebesgue a introduit dans sa thèse soutenue en 1902 une théorie de l'intégrale qui généralise de façon très satisfaisante celle de Riemann et qui est également largement utilisée de nos jours. Il consacra ensuite ses travaux à étudier dans le cadre de sa nouvelle théorie les fonctions de plusieurs variables et les séries de fonctions. Il fut nommé en 1922 à l'Académie des Sciences.

(on dit que la suite $(f_n)_n$ est dominée uniformément par M sur $[a,b]$) alors

$$\int_a^b \lim_{n\to+\infty} f_n(x)\,\mathrm{d}x = \lim_{n\to+\infty} \int_a^b f_n(x)\,\mathrm{d}x.$$

◇

Exercice 2.7 On considère la suite d'applications $(f_n)_{n\in\mathbb{N}^*}$ définie par

$$f_n : x \in \mathbb{R} \longmapsto n\left(\exp\left(\frac{x^2}{n}\right) - \frac{n^2}{n^2+1}\right).$$

1 - Montrer que $(f_n)_n$ converge simplement sur $\mathbb{R}$ vers $f : x \in \mathbb{R} \mapsto x^2$.
2 - a) Montrer que $(f_n)_n$ converge uniformément sur tout intervalle de la forme $[-\alpha,\alpha]$ où $\alpha \in \mathbb{R}_+^*$.
b) $(f_n)_n$ converge-t-elle uniformément sur $\mathbb{R}$?
3 - Pour tout $n \in \mathbb{N}^*$, on note F_n la primitive de f_n qui vérifie $F_n(0) = 0$. Montrer que la suite d'applications $(F_n)_{n\in\mathbb{N}^*}$ converge simplement sur $\mathbb{R}$ vers l'application $F : x \in \mathbb{R} \mapsto x^3$.

2.3.4 Limite uniforme d'une suite d'applications dérivables

Soient I un intervalle ouvert de $\mathbb{R}$ et $(f_n)_n$ une suite d'applications de I dans $\mathbb{R}$ dérivables sur I. On se pose la question de savoir à quelle(s) condition(s) la limite f de la suite $(f_n)_n$ est elle-même une application dérivable. Contrairement à ce qui a été vu avec la continuité et l'intégrabilité, la convergence uniforme de la suite $(f_n)_n$ sur I n'est pas une condition suffisante pour assurer la dérivabilité de l'application f sur I. Pour l'établir, il suffit de considérer la suite d'applications $(f_n)_n$ de terme général $f_n : x \in \mathbb{R} \mapsto \sqrt{x^2 + 1/n^2}$. Nous avons montré p. 75 que cette suite convergeait uniformément sur $\mathbb{R}$ vers la fonction valeur absolue. Toutefois, bien que chacune des applications f_n soit dérivable sur $\mathbb{R}$, l'application valeur absolue vers laquelle cette suite converge n'est pas dérivable en 0.

Le théorème suivant donne des conditions suffisantes qui permettent d'affirmer que la limite d'une suite d'applications dérivables est elle-même dérivable.

THÉORÈME 2.5 *Soient I un intervalle ouvert et $(f_n)_n$ une suite d'applications de I dans $\mathbb{R}$. Si*

1. *pour tout $n \in \mathbb{N}$ l'application f_n est de classe C^1 sur I ;*
2. *la suite d'applications $(f_n')_n$ converge uniformément sur I vers une application g de I dans $\mathbb{R}$;*
3. *il existe un réel $x_0 \in I$ tel que la suite numérique de terme général $f_n(x_0)$ converge ;*

alors la suite $(f_n)_n$ converge simplement sur I vers une application f qui est de classe C^1 sur I et qui admet pour dérivée $f' = g$.

Démonstration Nous allons montrer que la suite $(f_n)_n$ converge simplement sur I vers l'application f définie par :

$$f(x) = \ell + \int_{x_0}^{x} g(t)\ \mathrm{d}t \tag{16}$$

où $\ell = \lim_{n\to+\infty} f_n(x_0)$. D'après la théorème 2.4, l'application g qui est limite d'une suite d'applications Riemann-intégrables sur I convergeant uniformément sur I est Riemann-intégrable sur I. Par conséquent, pour tout $x \in I$ l'application g est Riemann-intégrable sur $[x_0, x]$ et la relation (16) a bien un sens.

Pour tout $x \in I$ et pour tout $n \in \mathbb{N}$, on a[24] : $f_n(x) = f_n(x_0) + \int_{x_0}^{x} f_n'(t)\ \mathrm{d}t$. On en déduit, compte tenu de la relation (16), que

$$|f(x) - f_n(x)| \leqslant |\ell - f_n(x_0)| + \left| \int_{x_0}^{x} g(t)\ \mathrm{d}t - \int_{x_0}^{x} f_n'(t) \right|. \tag{17}$$

Soit $x \in I$; puisque pour tout $n \in \mathbb{N}$ l'application f_n est de classe $\mathcal{C}^1$ sur I, f_n' est une application continue sur l'intervalle[25] $[x_0, x]$. Elle est par conséquent Riemann-intégrable sur $[x_0, x]$. Par ailleurs, la suite d'applications $(f_n')_n$ converge uniformément sur I vers l'application g. On en déduit que sur $[x_0, x]$ la suite $(f_n')_n$ converge uniformément vers la restriction de g à l'intervalle $[x_0, x]$. D'après le théorème 2.4, sur l'intervalle $[x_0, x]$ l'application g est intégrable et

$$\lim_{n\to+\infty} \int_{x_0}^{x} f_n'(t)\ \mathrm{d}t = \int_{x_0}^{x} \lim_{n\to+\infty} f_n'(t)\ \mathrm{d}t = \int_{x_0}^{x} g(t)\ \mathrm{d}t. \tag{18}$$

Puisque $\ell := \lim_{n\to+\infty} f_n(x_0)$, la relation (18) combinée à (17) permet de conclure que $\lim_{n\to+\infty} |f(x) - f_n(x)| = 0$ et par conséquent que la suite $(f_n)_n$ converge simplement sur I vers l'application f définie par la relation (16). La relation (16) indique que f est une primitive de g sur I donc que f est dérivable sur I avec $f' = g$. Comme g est continue sur I (c'est la limite uniforme d'une suite d'applications continues, voir le théorème 2.3), l'application f est de classe $\mathcal{C}^1$ sur I. □

Il apparaît dans la démonstration que l'on peut affaiblir la seconde hypothèse du théorème. Il suffit que la suite $(f_n')_n$ converge simplement sur I et converge uniformément sur tout intervalle fermé borné (autrement dit compact) inclus dans I. On peut aussi affaiblir l'hypothèse de régularité $\mathcal{C}^1$ des applications f_n. Si les applications f_n sont uniquement supposées dérivables sur I, alors on peut montrer que la suite $(f_n)_n$ converge simplement sur I vers une application f qui est dérivable sur I et qui admet pour dérivée $f' = g$. On a le théorème suivant que nous admettrons.

---

[24] Voir la section 18.3.1 p. 847 du *Cours de première année*.

[25] On commet ici un abus de notation : il s'agit de l'intervalle $[x_0, x]$ pour $x > x_0$ et de l'intervalle $[x, x_0]$ pour $x < x_0$.

THÉORÈME 2.6 *Soit $(f_n)_n$ une suite d'applications de I dans $\mathbb{R}$. Si*

1. *pour tout $n \in \mathbb{N}$ l'application f_n est dérivable sur I ;*
2. *la suite d'applications $(f'_n)_n$ converge simplement sur I vers une application g de I dans $\mathbb{R}$ et converge uniformément vers g sur tout intervalle fermé borné inclus dans I ;*
3. *il existe un réel $x_0 \in I$ tel que la suite numérique de terme général $f_n(x_0)$ converge simplement ;*

alors la suite $(f_n)_n$ converge simplement sur I vers une application f qui est dérivable sur I et qui admet pour dérivée $f' = g$.

Exercice 2.8

1 - Montrer que la suite d'applications $(f_n)_n$ de terme général $f_n : x \in [-1,1] \mapsto \dfrac{x}{1+n^2x^2}$ converge uniformément sur $[-1,1]$.

2 - Montrer que la suite des applications dérivées $(f'_n)_n$ converge simplement sur $[-1,1]$. La suite $(f'_n)_n$ converge-t-elle uniformément sur $[-1,1]$?

2.4 Plan d'étude de la convergence d'une suite d'applications

La procédure pour l'étude d'une suite d'applications $(f_n)_n$ d'un sous-ensemble X de $\mathbb{R}$ dans $\mathbb{R}$ peut être résumée de la manière suivante.

Convergence simple

- S'il existe $x_0 \in X$ tel que la suite numérique réelle $(f_n(x_0))_n$ diverge alors la suite d'applications $(f_n)_n$ ne converge pas simplement sur X.
- Pour montrer que la suite d'applications $(f_n)_n$ converge simplement sur X, il faut montrer que pour tout $x \in X$ la suite numérique réelle $(f_n(x))_n$ converge (vers un réel ℓ_x). La suite d'applications $(f_n)_n$ converge alors simplement sur X vers l'application $f : x \in X \mapsto \ell_x$.

Convergence uniforme

- Si la suite d'applications $(f_n)_n$ ne converge pas simplement sur X, elle ne peut pas converger uniformément sur X.
- Supposons que la suite d'applications $(f_n)_n$ converge simplement sur X vers f ; si la suite d'applications $(f_n)_n$ converge uniformément sur X, ce ne peut être que vers f.
- Le théorème 2.3 (et parfois le théorème 2.4) peut servir à établir la non-convergence uniforme d'une suite d'applications : si les applications f_n sont toutes continues sur X et que la limite f n'est pas continue sur X alors la convergence de la suite d'applications $(f_n)_n$ vers f ne peut pas être uniforme sur X.

- S'il existe une suite $(x_n)_n$ d'éléments de X convergeant vers $x \in X$ telle que la suite de terme général $f_n(x_n)$ ne converge pas vers $f(x)$ alors la convergence de la suite $(f_n)_n$ vers f n'est pas uniforme sur X, voir l'exercice 2.6.
- La façon générale de procéder pour étudier la convergence uniforme de la suite d'applications $(f_n)_n$ vers f sur l'ensemble X est de considérer[26] la suite numérique de terme général $u_n = \sup_{x\in X} |f_n(x) - f(x)|$. Si cette suite converge vers 0 alors il y a convergence uniforme ; dans le cas contraire, il n'y a pas convergence uniforme. Pour déterminer la nature de la suite $(u_n)_n$, on peut :
  - soit étudier la fonction $g_n : x \mapsto f_n(x) - f(x)$ sur X pour déterminer l'expression de u_n puis en calculer la limite éventuelle quand n tend vers $+\infty$;
  - soit majorer u_n par une quantité qui converge vers 0 et utiliser le théorème d'encadrement pour établir que $(u_n)_n$ converge vers 0 ;
  - soit minorer u_n par une quantité qui ne tend pas vers 0 pour établir que $(u_n)_n$ ne converge pas vers 0.

2.5 Exercices de synthèse

Exercice 2.9 ***Calcul de la valeur de*** $\int_0^{+\infty} \mathrm{e}^{-t^2}\,\mathrm{d}t$

Soit $(f_n)_n$ la suite d'applications de terme général

$$f_n : x \in [0,+\infty[\longmapsto \begin{cases} \left(1 - \dfrac{x^2}{n}\right)^n & \text{si } x \in [0,\sqrt{n}] \\ 0 & \text{si } x \in]\sqrt{n},+\infty[\end{cases}.$$

Le but de cet exercice est de montrer que[27] $\int_0^{+\infty} \mathrm{e}^{-t^2}\,\mathrm{d}t = \frac{1}{2}\sqrt{\pi}$.

1 - a) Pour $n \in \mathbb{N}^*$ fixé, étudier l'application f_n.

b) Montrer que la suite $(f_n)_n$ converge simplement sur $[0,+\infty[$ vers $f : x \in [0,+\infty[\mapsto \mathrm{e}^{-x^2}$.

c) Soient a un réel strictement positif fixé et n un entier strictement supérieur à $\mathrm{E}(a^2)$. Calculer $\sup_{x\in[0,a]} \left(1 - \exp\left(x^2 + n\ln(1 - \frac{x^2}{n})\right)\right)$.

d) En déduire que la suite $(f_n)_n$ converge uniformément sur $[0,a]$ pour tout $a \in \mathbb{R}_+^*$.

---

[26] Lorsqu'elle est définie ; si la fonction $g_n : x \mapsto f_n(x) - f(x)$ n'est pas bornée sur X à partir d'un certain rang, il n'y a pas convergence uniforme de de la suite $(f_n)_n$ vers f.

[27] Les primitives de la fonction $f : x \mapsto \mathrm{e}^{-x^2}$ (qui existent car f est continue sur $\mathbb{R}$, voir le corollaire 18.2 p. 851 du *Cours de première année*) ne s'exprimant pas à l'aide des fonctions usuelles, on doit avoir recours à des moyens détournés pour évaluer cette intégrale. Cet exercice en propose un ; nous verrons également qu'il est possible d'évaluer cette intégrale en ayant recours à des résultats sur l'intégrale des fonctions de plusieurs variables.

e) Soit ε un réel strictement positif fixé ; montrer qu'il existe $N \in \mathbb{N}$ tel que pour tout entier $n \geqslant N$ et pour tout $x \in [0, +\infty[$ on ait $|f_n(x) - f(x)| \leqslant \varepsilon$. En déduire que la suite $(f_n)_n$ converge uniformément sur $\mathbb{R}_+^*$.

2 - On considère la suite $(F_n)_n$ de terme général $F_n : x \in [0, +\infty[\mapsto \int_0^x f_n(t)\ \mathrm{d}t$.

a) Montrer que la suite $(F_n)_n$ converge uniformément sur $[0, +\infty[$ vers $F : x \in [0, +\infty[\mapsto \int_0^x \mathrm{e}^{-t^2}\ \mathrm{d}t$.

b) Pour $n \in \mathbb{N}^*$, on note $\ell_n = \int_0^{\sqrt{n}} f_n(t)\ \mathrm{d}t$. En considérant le changement de variable $t = \sqrt{n}\sin(\theta)$, montrer que $\ell_n = \sqrt{n}\int_0^{\frac{\pi}{2}} \cos^{2n+1}(\theta)\ \mathrm{d}\theta$.

En déduire que $\ell_n = \sqrt{n}\dfrac{(2^n\, n!)^2}{(2n+1)!}$ puis, en utilisant la formule de Stirling[28], montrer que la suite $(\ell_n)_n$ converge vers $\frac{1}{2}\sqrt{\pi}$.

3 - Conclure en montrant que $\int_0^{+\infty} \mathrm{e}^{-t^2}\ \mathrm{d}t = \frac{1}{2}\sqrt{\pi}$.

Exercice 2.10

On considère la suite $(f_n)_n$ de terme général $f_n : x \in \mathbb{R} \mapsto \int_1^n t^{-x}\ \mathrm{d}t$.

1 - Calculer $f_n(x)$ pour tout $x \in \mathbb{R}$ et tout $n \in \mathbb{N}^*$.

2 - a) Calculer le développement limité à l'ordre 1 au voisinage de $x = 1$ de la fonction $x \mapsto \dfrac{1}{x-1}\left(1 - \dfrac{1}{n^{x-1}}\right)$.

b) En déduire que pour tout $n \in \mathbb{N}^*$ l'application f_n est dérivable sur $\mathbb{R}$ et calculer $f_n'(x)$ pour tout $x \in \mathbb{R}$.

c) L'application f_n est-elle de classe $\mathcal{C}^1$ sur $\mathbb{R}$?

3 - Étudier la convergence simple et uniforme de la suite $(f_n')_n$. On note g la limite de la suite $(f_n')_n$.

4 - a) Montrer que la suite de fonctions $(f_n)_n$ converge simplement sur $]1, +\infty[$ et déterminer la limite f de cette suite.

b) Vérifier que $f' = g$ et analyser cette situation au regard d'un résultat du cours.

---

[28] On rappelle la formule de Stirling : $n! \underset{+\infty}{\sim} \sqrt{2\pi n}\left(\frac{n}{\mathrm{e}}\right)^n$.

2.6 Solutions des exercices

Solution de l'exercice 2.1

Considérons l'application $f : x \in]0,2[\mapsto \mathrm{sh}(x^2 - x - 1)$. Comme la fonction sinus hyperbolique est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$, f est de classe $\mathcal{C}^\infty$ sur $]0,2[$ en tant que composée de la fonction sinus hyperbolique avec une fonction polynomiale. Pour tout $x \in \mathbb{R}$, on a $f'(x) = (2x-1)\,\mathrm{ch}(x^2 - x - 1)$. La fonction cosinus hyperbolique étant minorée par 1, le signe de f' est celui de $2x-1$. On a donc le tableau de variation suivant pour f :

x	0		$\frac{1}{2}$		2
$f'(x)$	$-\,\mathrm{ch}(1)$	$-$	0	$+$	$3\,\mathrm{ch}(1)$
$f(x)$	$\mathrm{sh}(1)$	↘	$-\mathrm{sh}(5/4)$	↗	$\mathrm{sh}(1)$

Comme la fonction sinus hyperbolique est strictement croissante sur $\mathbb{R}$, on a $\mathrm{sh}(5/4) > \mathrm{sh}(1)$. On déduit du tableau de variation que

$$\sup_{x\in]0,2[} |\mathrm{sh}(x^2 - x - 1)| = \mathrm{sh}(5/4).$$

Solution de l'exercice 2.2

Si $x = 0$ alors $f_n(x) = 0$ pour tout $n \in \mathbb{N}$ et si $x \in \mathbb{R}^*$ alors $f_n(x) \underset{n\to+\infty}{\sim} 1/x^{2n-1}$ donc $\lim_{n\to+\infty} f_n(x) = 0$. On en déduit que la suite $(f_n)_n$ converge simplement sur $\mathbb{R}$ vers l'application nulle. Pour tout $n \in \mathbb{N}^*$, on a

$$\sup_{x\in\mathbb{R}} |f_n(x) - 0| = \sup_{x\in\mathbb{R}} \frac{|x|}{(1+x^2)^n} = \sup_{x\in\mathbb{R}^+} \frac{x}{(1+x^2)^n}.$$

Soit $g_n : x \in \mathbb{R}^+ \mapsto x/(1+x^2)^n$. Pour tout $x \in \mathbb{R}^+$ on a

$$g_n'(x) = \frac{1-(2n-1)x^2}{(1+x^2)^{n+1}}$$

et g_n' s'annule sur $\mathbb{R}^+$ en $x_n = \frac{1}{\sqrt{2n-1}}$. Le tableau de variation de l'application g_n est donc le suivant :

x	0		x_n		$+\infty$
$\psi'(x)$		$+$	0	$-$	
$\psi(x)$	0	↗	$g_n(x_n)$	↘	0

On en déduit que :

$$\sup_{x\in\mathbb{R}} |f_n(x) - 0| = g_n(x_n) = \frac{1}{\sqrt{2n-1}} \frac{1}{(1+\frac{1}{2n-1})^n} = \frac{1}{\sqrt{2n-1}} \left(1 - \frac{1}{2n}\right)^n.$$

Or,

$$\left(1-\frac{1}{2n}\right)^n = \sqrt{\left(1-\frac{1}{2n}\right)^{2n}} \quad \text{et} \quad \lim_{n\to+\infty}\left(1-\frac{1}{2n}\right)^{2n} = \lim_{p\to+\infty}\left(1-\frac{1}{p}\right)^{p} = \frac{1}{\mathrm{e}}.$$

On en déduit que $g_n(x_n) \underset{+\infty}{\sim} 1/\sqrt{2\,\mathrm{e}\,n}$ ce qui implique que $\lim_{n\to+\infty} g_n(x_n) = 0$. On en conclut que la suite $(f_n)_n$ converge uniformément sur $\mathbb{R}$ vers l'application nulle.

Solution de l'exercice 2.3

1 - D'après la proposition 2.2, une condition nécessaire et suffisante pour que la suite $(f_n)_n$ converge uniformément vers la fonction nulle est que la suite de terme général $u_n = \sup_{x\in X} |f_n(x)|$ converge vers 0. Or, par définition de la borne supérieure, $\sup_{x\in X} |f_n(x)| \geqslant |f_n(x_n)|$ pour tout $n \in \mathbb{N}$. Il est donc nécessaire que la suite de terme général $|f_n(x_n)|$ converge vers 0 pour que la suite $(u_n)_n$ converge vers 0. Comme la suite de terme général $|f_n(x_n)|$ converge vers 0 si et seulement si la suite de terme général $f_n(x_n)$ converge vers 0, on en déduit qu'une condition suffisante pour que la suite $(f_n)_n$ ne converge pas uniformément sur X vers la fonction nulle, est qu'il existe une suite $(x_n)_n$ d'éléments de X telle que la suite de terme général $f_n(x_n)$ ne tende pas vers 0.

2 - a) Si $x \in [0, \frac{\pi}{2}[\cup]\frac{\pi}{2}, \pi]$ alors $0 \leqslant \sin(x) < 1$. Comme pour tout $\alpha \in [0,1[$ on a $\lim_{n\to+\infty} \sqrt{n}\,\alpha^n$, on en déduit que $\lim_{n\to+\infty} f_n(x) = 0$. Si $x = \frac{\pi}{2}$ alors $f_n(x) = 0$ pour tout $n \in \mathbb{N}$ et on a aussi $\lim_{n\to+\infty} f_n(x) = 0$. On en conclut que pour tout $x \in [0, \pi]$ la suite numérique de terme général $f_n(x)$ converge vers 0, autrement dit que la suite d'applications $(f_n)_n$ converge simplement vers l'application nulle.

b) Déterminons l'expression de $f_n(x_n)$ pour tout $n \in \mathbb{N}^*$. On a $\sin(\arcsin(1-1/n)) = 1-1/n$ et comme $x_n \in [0, \frac{\pi}{2}[$, le réel $\cos(x_n)$ est positif et s'exprime sous la forme $\cos(x_n) = \sqrt{1-\sin^2(x_n)} = 1/\sqrt{n}$. On a donc

$$f_n(x_n) = \sqrt{n} \times \frac{1}{\sqrt{n}} \times \left(1-\frac{1}{n}\right) = 1-\frac{1}{n}$$

et cette quantité tend vers 1 quand n tend vers 0. On conclut, en utilisant la question 1, que la suite $(f_n)_n$ ne converge pas uniformément sur $[0, \pi]$ vers la fonction nulle.

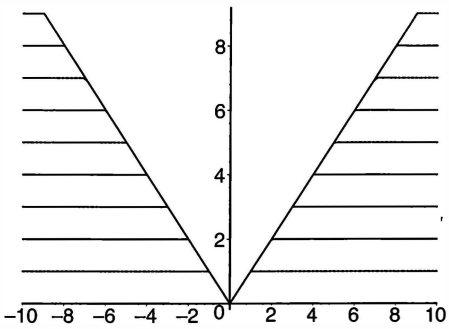
Solution de l'exercice 2.4

1 - L'application f_n est bornée puisque pour tout $x \in \mathbb{R}$ on a $|f_n(x)| \leqslant n$.

Soit x un réel fixé. Pour tout $n \in \mathbb{N}$ avec $n > \mathrm{E}(x)$ on a $f_n(x) = |x|$. La suite $(f_n)_n$ converge donc simplement vers la fonction valeur absolue qui n'est pas une fonction bornée sur $\mathbb{R}$.

On peut illustrer le comportement de la suite $(f_n)_n$ avec MAPLE.

```
> fn := (n,x)-> min(abs(x),n):
> plot([seq(fn(n,x),n=1..9)],x=-10..10,linestyle=[dash,solid]);
```



Pour tous $n \in \mathbb{N}$ et $x \in \mathbb{R}$, on a

$$|f_n(x) - f(x)| = |x| - \min(|x|, n) = \begin{cases} 0 & \text{si } x \in [-n, n] \\ |x| - n & \text{sinon} \end{cases}.$$

On en déduit que $\sup_{x\in\mathbb{R}} |f_n(x) - f(x)| = \sup_{x\in]n,+\infty[}(|x| - n) = +\infty$. Il n'y a donc pas convergence uniforme de la suite $(f_n)_n$ vers la fonction valeur absolue.

On dispose avec cet exemple d'une suite d'applications bornées qui converge simplement vers une application non bornée. La convergence n'est pas uniforme.

2 - L'application f_n n'est pas bornée puisque[29] pour tout $M \in \mathbb{R}^+$ le réel $x = M$ vérifie $|f(x)| > M$.

Pour $x \in \mathbb{R}$ fixé, la suite numérique réelle de terme général $f_n(x) = x + \frac{1}{n}$ converge vers x. On en déduit que la suite $(f_n)_n$ converge simplement sur $\mathbb{R}$ vers l'application identité qui n'est pas bornée sur $\mathbb{R}$. Cette convergence est également uniforme puisque $\sup_{x\in\mathbb{R}} |f_n(x) - f(x)| = \frac{1}{n}$ converge vers 0.

On a avec cet exemple, une suite d'applications non bornées qui converge uniformément. La limite est une application non bornée.

Solution de l'exercice 2.5

1 - La quantité $f_n(x)$ est définie pour $x \in \mathbb{R}$ tel que $1 + \frac{x}{n(1+x)} > 0$, c'est-à-dire pour x tel que

$$\frac{x}{1+x} > -n. \tag{19}$$

- Si $1 + x < 0$, *i.e.* si $x \in]-\infty, -1[$, alors (19) équivaut à $x < -n(1 + x)$, autrement dit à $x < -\frac{n}{n+1}$.

---

[29] On rappelle qu'une application f de X dans $\mathbb{R}$ est bornée sur l'ensemble X si : $\exists M \in \mathbb{R}^+ \ \forall x \in X \ |f(x)| \leqslant M$. En prenant la négation de cette assertion quantifiée, on obtient qu'une application f de X dans $\mathbb{R}$ n'est pas bornée sur l'ensemble X si : $\forall M \in \mathbb{R}^+ \ \exists x \in X \ |f(x)| > M$.

- Si $1+x>0$, *i.e.* si $x\in]-1,+\infty[$, alors (19) équivaut à $x>-n(1+x)$, autrement dit à $x>-\frac{n}{n+1}$.

Comme $-\frac{n}{n+1}<-1$, on en déduit que f_n est défini sur

$$\mathcal{D}_n=]-\infty,-1[\cup]-\frac{n}{n+1},+\infty[.$$

La suite de fonctions $(f_n)_n$ est donc définie sur

$$\mathcal{D}=\bigcap_{n\in\mathbb{N}^*}\mathcal{D}_n=]-\infty,-1[\cup]-\tfrac{1}{2},+\infty[.$$

2 - Montrer que la suite de fonctions $(f_n)_{n\in\mathbb{N}^*}$ converge simplement sur $\mathcal{D}$, revient à montrer que pour tout $x\in\mathcal{D}$ fixé, la suite numérique réelle $(f_n(x))_n$ converge. Soit x un élément de $\mathcal{D}$; comme $\lim\limits_{n\to+\infty}\frac{x}{n(1+x)}=0$ et $\ln(1+u)\underset{0}{\sim}u$, on a $f_n(x)\underset{n\to+\infty}{\sim}(-1)^n\frac{x}{n(1+x)}$. Deux suites réelles dont les termes sont équivalents en $+\infty$ convergent vers une même limite ou divergent simultanément[30], donc pour tout $x\in\mathcal{D}$ fixé, la suite numérique réelle $(f_n(x))_n$ converge vers 0. Par conséquent, la suite de fonctions $(f_n)_{n\in\mathbb{N}^*}$ converge simplement sur $\mathcal{D}$ vers la fonction nulle.

3 - La suite de fonctions $(f_n)_{n\in\mathbb{N}^*}$ converge uniformément sur $\mathcal{D}$ vers la fonction nulle f si la suite numérique réelle $(u_n)_n$ de terme général

$$u_n=\sup_{x\in\mathcal{D}}|f_n(x)-f(x)|=\sup_{x\in\mathcal{D}}\left|\ln\left(1+\frac{x}{n(1+x)}\right)\right|$$

converge vers 0.

Afin d'étudier la nature de la suite $(u_n)_n$, commençons par étudier, pour $n\in\mathbb{N}^*$ fixé, l'application $\psi_n:x\in\mathcal{D}\mapsto\ln\left(1+\frac{x}{n(1+x)}\right)$. Puisque l'application $x\in\mathcal{D}\mapsto1+\frac{x}{n(1+x)}\in]0,+\infty[$ est de classe $\mathcal{C}^1$ sur $\mathcal{D}$ et que la fonction logarithme est de classe $\mathcal{C}^1$ sur $]0,+\infty[$, l'application ψ_n est de classe $\mathcal{C}^1$ sur $\mathcal{D}$. De plus, pour $x\in\mathcal{D}$ on a

$$\psi_n'(x)=\frac{1}{(n+1)(x+1)\left(x+\frac{n}{n+1}\right)}.$$

Le signe de $\psi_n'(x)$ est celui du produit $(x+1)\times\left(x+\frac{n}{n+1}\right)$. On obtient le tableau de variation suivant pour ψ_n :

x	$-\infty$		-1	$\|\|$	$\|\|$	$-\frac{1}{2}$		$+\infty$
$\psi_n'(x)$		$+$		$\|\|$	$\|\|$		$+$	
$\psi_n(x)$	$\ln(1+\frac{1}{n})$	$\nearrow$	$+\infty$	$\|\|$	$\|\|$	$\ln(1-\frac{1}{n})$	$\nearrow$	$\ln(1+\frac{1}{n})$

[30] Voir la proposition 15.10 p. 701 du *Cours de première année*.

Comme ψ_n n'est pas bornée sur $\mathcal{D}$, on a $\sup_{x\in\mathcal{D}} |f_n(x) - f(x)| = +\infty$ et par conséquent la suite $(f_n)_n$ ne converge pas uniformément sur $\mathcal{D}$.

La suite de fonctions $(f_n)_{n\in\mathbb{N}^*}$ convergera uniformément sur $]-\frac{1}{2}, +\infty[$ vers la fonction nulle f si la suite numérique réelle $(v_n)_n$ de terme général

$$v_n = \sup_{x\in]-\frac{1}{2},+\infty[} |f_n(x) - f(x)| = \sup_{x\in]-\frac{1}{2},+\infty[} |\psi_n(x)| = \max\left(\left|\ln(1-\tfrac{1}{n})\right|, \ln(1+\tfrac{1}{n})\right)$$

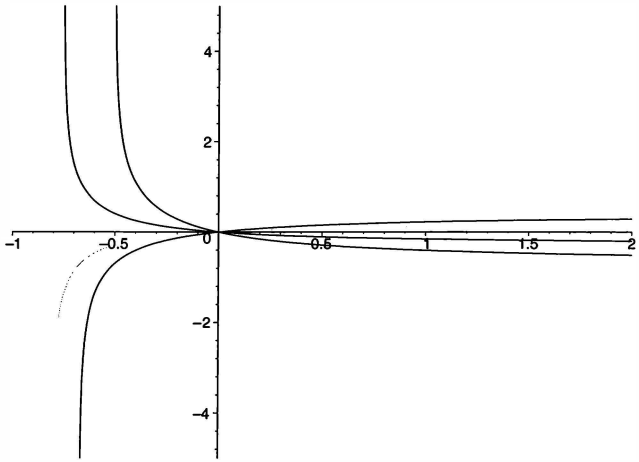
converge vers 0. Comme $\lim_{n\to+\infty} \left|\ln(1-\frac{1}{n})\right| = \lim_{n\to+\infty} \ln(1+\frac{1}{n}) = 0$, la suite $(v_n)_n$ converge vers 0 et par conséquent la suite $(f_n)_n$ converge uniformément sur $]-\frac{1}{2}, +\infty[$. Même s'il n'a pas été utile de déterminer l'expression de v_n, on peut remarquer que

$$\left|\ln(1-\tfrac{1}{n})\right| = \left|-\ln\left(\tfrac{n}{n-1}\right)\right| = \ln\left(\tfrac{n}{n-1}\right) \quad \text{et} \quad \ln(1+\tfrac{1}{n}) = \ln\left(\tfrac{n+1}{n}\right)$$

ce qui permet d'établir que $v_n = \ln\left(\frac{n}{n-1}\right)$ car $\frac{n+1}{n} < \frac{n}{n-1}$ et la fonction logarithme est croissante.

On peut aisément représenter graphiquement le comportement de la suite $(f_n)_n$ avec MAPLE :

```
> fn:= (n,x)->(-1)**n*log(1+x/(n*(1+x))):
> plot([fn(1,x),fn(2,x),fn(3,x),fn(4,x)], x=-1..5);
```



Solution de l'exercice 2.6

Pour montrer que $\lim_{n\to+\infty} f_n(x_n) = f(x)$, montrons que[31] :

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad (n \geqslant N \implies |f_n(x_n) - f(x)| \leqslant \varepsilon). \qquad (20)$$

[31] Voir la définition 5.1 p. 168 du *Cours de première année*.

Soit ε un réel strictement positif fixé ; pour tout $n \in \mathbb{N}$ on a :

$$|f_n(x_n) - f(x)| \leqslant |f_n(x_n) - f_n(x)| + |f_n(x) - f(x)|.$$

Par hypothèse, la suite $(f_n)_n$ converge uniformément vers f sur I ce qui implique que[32]

$$\exists N_1 \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad \forall y \in I \quad (n \geqslant N_1 \implies |f_n(y) - f(y)| \leqslant \tfrac{1}{2}\varepsilon). \tag{21}$$

Par ailleurs, les applications f_n sont continues sur I donc en x. On a donc également, pour tout $n \in \mathbb{N}$, l'assertion suivante[33] :

$$\exists \eta \in \mathbb{R}_+^* \quad \forall y \in I \quad (|x - y| \leqslant \eta \implies |f_n(y) - f_n(x)| \leqslant \tfrac{1}{2}\varepsilon). \tag{22}$$

Enfin, la convergence de la suite $(x_n)_n$ vers x implique que[34] :

$$\exists N_2 \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad (n \geqslant N_2 \implies |x_n - x| \leqslant \eta). \tag{23}$$

On déduit de (23) que pour $n \geqslant N_2$ on a $|x_n - x| \leqslant \eta$ ce qui entraîne d'après (22) que $|f_n(x_n) - f_n(x)| \leqslant \frac{1}{2}\varepsilon$. En combinant cette majoration avec (21), on en déduit que pour $n \geqslant N$ où $N = \max\{N_1, N_2\}$ on a $|f_n(x_n) - f(x)| \leqslant \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon = \varepsilon$ et l'assertion (20) est démontrée.

Solution de l'exercice 2.7

1 - Pour tout $x \in \mathbb{R}^*$ fixé, on a

$$\mathrm{e}^{x^2/n} = 1 + \frac{x^2}{n} + \frac{x^4}{2n^2} + \mathcal{O}_{+\infty}\left(n^{-2}\right) \quad \text{et} \quad \frac{n^2}{n^2+1} = \frac{1}{1+\frac{1}{n^2}} = 1 - \frac{1}{n^2} + \mathcal{O}_{+\infty}\left(n^{-2}\right).$$

On en déduit que :

$$f_n(x) = x^2 + \frac{1}{n}\left(\frac{x^4}{2} + 1\right) + \mathcal{O}_{+\infty}\left(n^{-1}\right). \tag{24}$$

On a donc $\lim_{n\to+\infty} f_n(x) = x^2$. Par ailleurs, pour $x = 0$ et $n \in \mathbb{N}^*$, on a

$$f_n(0) = n\left(1 - \frac{n^2}{n^2+1}\right) = \frac{n}{n^2+1}$$

d'où $\lim_{n\to+\infty} f_n(0) = 0$. On en conclut que la suite $(f_n)_n$ converge simplement vers l'application $f : x \in \mathbb{R} \mapsto x^2$.

2 - a) Pour montrer que $(f_n)_n$ converge uniformément sur tout intervalle de la forme $[-\alpha, \alpha]$ avec $\alpha \in \mathbb{R}_+^*$, il faut montrer que la suite numérique de terme général $u_n = \sup_{x\in[-\alpha,\alpha]} |f_n(x) - f(x)|$ converge vers 0.

---

[32] Voir la définition 2.2 p. 73, considérée ici avec $\varepsilon = \frac{1}{2}\varepsilon$.

[33] Voir la définition 13.4 p. 588 du *Cours de première année*.

[34] Voir la définition 5.1 p. 168 du *Cours de première année*.

Commençons par déterminer l'expression de u_n. Pour tout $x \in \mathbb{R}$, on a

$$f_n(x) - f(x) = n \exp\left(\frac{x^2}{n}\right) - \frac{n^3}{n^2+1} - x^2.$$

Pour étudier les variations de $f_n - f$, étudions pour $n \in \mathbb{N}^*$ fixé la fonction ϕ_n : $x \mapsto n \exp\left(x^2/n\right) - x^2 - \frac{n^3}{n^2+1}$. Cette fonction est définie sur $\mathbb{R}$ et est paire. Elle est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$ et pour tout $x \in \mathbb{R}$ on a $\phi_n'(x) = 2x\left(\exp\left(x^2/n\right) - 1\right)$. La fonction exponentielle est strictement croissante sur $\mathbb{R}$ et prend la valeur 1 en 0. Comme x^2/n est un réel positif, on a $\left(\exp\left(x^2/n\right) - 1\right) > 0$ pour tout $x \in \mathbb{R}^*$ et $\phi_n'(x)$ est du signe de x. On a donc le tableau de variation suivant :

x	0		α
$\phi_n'(x)$	0	$+$	
$\phi_n(x)$	n	↗	

On déduit de ce tableau de variation que ϕ_n est une fonction positive qui atteint son maximum sur $[-\alpha, \alpha]$ en α. On a donc

$$u_n = \phi_n(\alpha) = f_n(\alpha) - f(\alpha) = n\left(\exp\left(\frac{\alpha^2}{n}\right) - \frac{\alpha^2}{n} - \frac{n^2}{n^2+1}\right).$$

On déduit de la relation (24) que

$$u_n = \frac{1}{n}\left(\frac{\alpha^4}{2} + 1\right) + o_{+\infty}\left(n^{-1}\right) \underset{n\to+\infty}{\sim} \frac{1}{n}\left(\frac{\alpha^4}{2} + 1\right)$$

ce qui permet d'affirmer que la suite $(u_n)_n$ converge vers 0. Il y a donc convergence uniforme de $(f_n)_n$ vers f sur tout intervalle de la forme $[-\alpha, \alpha]$.

b) Comme $\lim_{x\to+\infty} \phi_n(x) = +\infty$, l'application ϕ_n n'est pas bornée sur $\mathbb{R}$. On en déduit qu'il n'y a pas convergence uniforme de $(f_n)_n$ vers f sur $\mathbb{R}$.

3 - La primitive F_n de f_n qui vérifie $F_n(0) = 0$ est donnée par

$$\forall x \in \mathbb{R} \qquad F_n(x) = \int_0^x f_n(t)\, \mathrm{d}t.$$

Comme f_n est une application paire, pour tout $x \in \mathbb{R}$ on a :

$$F_n(-x) = \int_0^{-x} f_n(t)\, \mathrm{d}t = -\int_0^x f_n(-t)\, \mathrm{d}t = -\int_0^x f_n(t)\, \mathrm{d}t = -F_n(x),$$

ce qui signifie que F_n est impaire. Étudions l'application F_n sur $\mathbb{R}^+$. Pour un réel $x > 0$ fixé, f_n étant continue sur $\mathbb{R}$, elle est continue sur $[0, x]$ et par conséquent Riemann-intégrable sur $[0, x]$. De plus, la suite $(f_n)_n$ convergeant uniformément sur tout intervalle de la forme $[-\alpha, \alpha]$ avec $\alpha \in \mathbb{R}_+^*$, elle converge aussi uniformément sur $[0, x]$. Le théorème 2.4 p. 90 permet d'écrire :

$$\begin{aligned} \lim_{n\to+\infty} F_n(x) &= \lim_{n\to+\infty} \int_0^x f_n(t)\, \mathrm{d}t = \int_0^x \lim_{n\to+\infty} f_n(t)\, \mathrm{d}t \\ &= \int_0^x f(t)\, \mathrm{d}t = \int_0^x t^2\, \mathrm{d}t = \frac{1}{3}x^3. \end{aligned}$$

La suite d'applications $(F_n)_{n\in\mathbb{N}^*}$ converge donc simplement sur $\mathbb{R}$ vers l'application $F : x \in \mathbb{R} \mapsto \frac{1}{3}x^3$.

Solution de l'exercice 2.8

1 - Si $x = 0$ alors pour tout $n \in \mathbb{N}$ on a $f_n(x) = 0$ et la suite $(f_n)_n$ converge en 0 vers l'application nulle. Par ailleurs, pour $x \in [-1, 0[\cup]0, 1]$ fixé on a $(1 + n^2x^2) \underset{n\to+\infty}{\sim} n^2x^2$ et par conséquent $f_n(x) \underset{n\to+\infty}{\sim} 1/(n^2x)$.

Comme $\lim_{n\to+\infty} 1/(n^2x) = 0$, on en déduit que la suite $(f_n)_n$ converge simplement sur $[-1, 1]$ vers l'application nulle.

Montrons que la suite $(f_n)_n$ converge uniformément sur $[-1, 1]$ vers l'application nulle. Il s'agit pour cela de montrer que $u_n == \sup_{x\in[-1,1]} |f_n(x)|$ tend vers 0 quand n tend vers $+\infty$. Comme pour tout $n \in \mathbb{N}$ fixé, l'application f_n est paire on peut remarquer que $u_n = \sup_{x\in[0,1]} |f_n(x)|$.

Pour déterminer l'expression de u_n, étudions l'application f_n sur $[0, 1]$. Pour tout $n \in \mathbb{N}$ fixé, l'application f_n est de classe $\mathcal{C}^\infty$ sur $[-1, 1]$ et

$$\forall x \in [-1, 1] \qquad f_n'(x) = \frac{1 - n^2x^2}{(1 + n^2x^2)^2}.$$

On a le tableau de variation suivant :

x	0		$\frac{1}{n}$		1
$f_n'(x)$		$+$	0	$-$	
$f_n(x)$	0	↗	$\frac{1}{2n}$	↘	$\frac{1}{1+n^2}$

On déduit de ce tableau de variation que $u_n = \sup_{x\in[0,1]} |f_n(x)| = \frac{1}{2n}$. Comme u_n tend vers 0 quand n tend vers $+\infty$, on en conclut que la suite $(f_n)_n$ converge uniformément sur $[-1, 1]$ vers l'application nulle.

2 - Pour tout $n \in \mathbb{N}$ on a $f_n'(0) = 1$ donc $\lim_{n\to+\infty} f_n'(0) = 1$. Par ailleurs, pour tout $x \in [-1, 0[\cup]0, 1]$ fixé on a $f_n'(x) \underset{n\to+\infty}{\sim} -\frac{1}{n^2x^2}$ ce qui implique que $\lim_{n\to+\infty} f_n'(x) = 0$. On en déduit que la suite $(f_n')_n$ converge simplement sur $[-1, 1]$ vers l'application g qui vaut 1 en 0 et 0 sur $[-1, 0[\cup]0, 1]$.

Pour tout $n \in \mathbb{N}$, l'application f_n' est continue sur $[-1, 1]$ alors que l'application g ne l'est pas. On en déduit du théorème 2.3 que la suite $(f_n')_n$ ne converge pas uniformément vers g sur $[-1, 1]$.

Solution de l'exercice 2.9

1 - a) Pour $n \in \mathbb{N}^*$ fixé, l'application f_n est dérivable sur $]0, \sqrt{n}[$ et

$$\forall x \in]0, \sqrt{n}[\qquad f'(x) = -2x\left(1 - \frac{x^2}{n}\right)^{n-1}$$

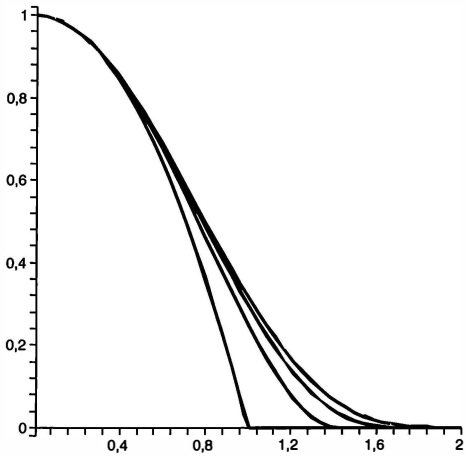
On a donc $f_n' < 0$ sur $]0, \sqrt{n}[$. Le tableau de variation de f_n est le suivant :

x	0		$\sqrt{n}$		$+\infty$
$f_n'(x)$	0	$-$	α_n	0	
$f_n(x)$	1	$\searrow$	0	$\longrightarrow$	0

Pour $n \geqslant 2$ on a $\alpha_n = 0$. Pour $n = 1$, la fonction f_1 n'est pas dérivable en $\sqrt{n} = 1$: elle admet pour dérivée à gauche -2 et pour dérivée à droite 0.

Utilisons MAPLE pour tracer la représentation graphique de ces fonctions :

```
> f_n:=(n,x)-> piecewise(x>0 and x<sqrt(n), (1-x^2/n)^n, 0):
> plot([seq(f_n(n,x),n=1..4)],x=0..2);
```



b) Montrer que la suite $(f_n)_n$ converge simplement vers l'application $f : x \in [0, +\infty[\mapsto \mathrm{e}^{-x^2}$ revient à montrer que quel que soit le réel x fixé la suite numérique de terme général $u_n = f_n(x)$ admet pour limite e^{-x^2}. Soit x un réel positif fixé ; pour n assez grand $(n \geqslant \mathrm{E}(x^2) + 1)$ on a $x \leqslant \sqrt{n}$ et par conséquent

$$f_n(x) = \left(1 - \frac{x^2}{n}\right)^n = \exp\left(n \ln(1 - \frac{x^2}{n})\right).$$

Pour $x > 0$ on a $\ln(1 - \frac{x^2}{n}) \underset{+\infty}{\sim} -\frac{x^2}{n}$ et par conséquent $\lim\limits_{n \to +\infty} n \ln(1 - \frac{x^2}{n}) = -x^2$. On a donc par continuité de la fonction exponentielle : $\lim\limits_{n \to +\infty} f_n(x) = \mathrm{e}^{-x^2}$. Si $x = 0$ alors $u_n = f_n(0) = 1$; la suite constante égale à 1 converge bien vers $\mathrm{e}^0 = 1$. On en conclut que la suite $(f_n)_n$ converge simplement vers l'application $f : x \in [0, +\infty[\mapsto \mathrm{e}^{-x^2}$.

c) Pour calculer $\sup_{x\in[0,a]} \exp\Big(x^2 + n\ln(1-\frac{x^2}{n})\Big)$, étudions les variations de la fonction $g : x \mapsto x^2 + n\ln(1-\frac{x^2}{n})$ sur $[0,a]$. Remarquons que pour n strictement supérieur à $\mathrm{E}(a^2)$ et $x \in [0,a]$ on a $0 \leqslant \frac{x^2}{n} < 1$; le réel $g(x)$ est donc bien défini. Comme la fonction logarithme est de classe $\mathcal{C}^\infty$ sur son ensemble de définition, l'application g est de classe $\mathcal{C}^\infty$ sur $[0,a]$. Pour tout $x \in]0,a[$,

$$g'(x) = -\frac{2x^3}{n-x^2}$$

et cette quantité est négative car $n - x^2 > 0$ pour $x \in]0,a[$ avec $n > \mathrm{E}(a^2)$. On en déduit que g est décroissante sur $[0,a]$. Comme la fonction exponentielle est croissante, pour tout $(x,y) \in [0,a]^2$ on a

$$x \leqslant y \implies g(y) \leqslant g(x) \implies \mathrm{e}^{g(y)} \leqslant \mathrm{e}^{g(x)} \implies 1 - \mathrm{e}^{g(x)} \leqslant 1 - \mathrm{e}^{g(y)}.$$

ce qui signifie que l'application $\phi : x \in [0,a] \mapsto 1 - \exp\Big(x^2 + n\ln(1-\frac{x^2}{n})\Big)$ est croissante sur $[0,a]$. On a donc

$$0 = \phi(0) \leqslant \phi(x) \leqslant \phi(a) \qquad \text{pour tout } x \in [0,a]$$

et par conséquent

$$\sup_{x\in[0,a]} \Big(1 - \exp\big(x^2 + n\ln(1-\tfrac{x^2}{n})\big)\Big) = \phi(a) = 1 - \mathrm{e}^{a^2}\left(1-\frac{a^2}{n}\right)^n.$$

d) Intéressons-nous aux termes de la suite $(f_n)_n$ à partir du rang $N = \mathrm{E}(a^2) + 1$. Pour tout entier $n \geqslant N$ on a $a \leqslant \sqrt{n}$ et

$$\forall x \in [0,a] \qquad f_n(x) = \left(1-\frac{x^2}{n}\right)^n.$$

La suite $(f_n)_n$ converge uniformément sur $[0,a]$ vers l'application f si et seulement si la suite numérique de terme général $u_n = \sup_{x\in[0,a]} |f_n(x) - f(x)|$ converge vers 0. Pour $n \geqslant N$ on a

$$\begin{aligned} u_n &= \sup_{x\in[0,a]} \left|\left(1-\frac{x^2}{n}\right)^n - \mathrm{e}^{-x^2}\right| = \sup_{x\in[0,a]} \left|\exp\big(n\ln(1-\tfrac{x^2}{n})\big) - \mathrm{e}^{-x^2}\right| \\ &= \sup_{x\in[0,a]} \mathrm{e}^{-x^2}\left|\exp\big(x^2 + n\ln(1-\tfrac{x^2}{n})\big) - 1\right| \leqslant 1 - \mathrm{e}^{-a^2}\left(1-\frac{a^2}{n}\right)^n \end{aligned}$$

car $\mathrm{e}^{-x^2} \leqslant 1$ pour tout $x \in [0,a]$. Comme la suite de terme général $v_n = \big(1 - a^2/n\big)^n$ converge vers e^{-a^2}, par le théorème d'encadrement on peut en conclure que la suite $(u_n)_n$ converge vers 0. La suite $(f_n)_n$ converge donc uniformément sur $[0,a]$ vers l'application f.

e) Soit $\varepsilon \in \mathbb{R}_+^*$; pour tout $a \in \mathbb{R}_+^*$, la suite $(u_n)_n$ de terme général $u_n = \sup_{x\in[0,a]} |f_n(x) - f(x)|$ convergeant vers 0, il existe[35] un entier naturel N tel que pour tout entier $n \geqslant N$ on ait $|u_n| \leqslant \varepsilon$. Pour tout $n \geqslant N$ et pour tout $x \in [0,a]$ on a donc $|f(x) - f_n(x)| \leqslant \varepsilon$.

[35] Voir la définition 5.1 p. 168 du *Cours de première année*.

Intéressons-nous maintenant à l'intervalle $[a, +\infty[$. Comme pour tous $n \in \mathbb{N}$ et $x \in \mathbb{R}^+$ on a $f_n(x) \in [0,1]$ et comme la fonction exponentielle est croissante, on obtient pour tout $x \in [a, +\infty[$

$$0 \leqslant f(x) - f_n(x) = \mathrm{e}^{-x^2} - f_n(x) \leqslant \mathrm{e}^{-x^2} \leqslant \mathrm{e}^{-a^2}.$$

Si a est tel que $\mathrm{e}^{-a^2} \leqslant \varepsilon$ alors pour tout $x \in [a, +\infty[$ on a $|f(x) - f_n(x)| \leqslant \varepsilon$.

On en conclut que pour tout $\varepsilon \in \mathbb{R}_+^*$, il existe un entier naturel N tel que pour tout entier $n \geqslant N$ et pour tout $x \in [0, +\infty[$ on a $|f(x) - f_n(x)| \leqslant \varepsilon$. D'après la définition 2.2, cela signifie que la suite $(f_n)_n$ converge uniformément sur $\mathbb{R}_+^*$.

2 - a) La suite $(F_n)_n$ converge uniformément sur $\mathbb{R}_+$ vers l'application F si la suite numérique de terme général $v_n = \sup_{x \in \mathbb{R}_+} |F_n(x) - F(x)|$ converge vers 0. Pour tout $x \in \mathbb{R}_+$ on a

$$\begin{aligned} |F_n(x) - F(x)| &= \left| \int_0^x f_n(t)\, \mathrm{d}t - \int_0^x f(t)\, \mathrm{d}t \right| = \left| \int_0^x (f_n(t) - f(t))\, \mathrm{d}t \right| \\ &\leqslant \int_0^x |f_n(t) - f(t)|\, \mathrm{d}t. \end{aligned}$$

Soit a un réel strictement positif fixé. Pour tout $n \in \mathbb{N}$ avec $n > \mathrm{E}(a^2)$, en notant $\alpha_n = 1 - e^{-a^2} \left(1 - \frac{a^2}{n}\right)^n$, on déduit de la question 1-d) que

- pour $x \in [0, a]$,

$$|F_n(x) - F(x)| \leqslant \int_0^x \sup_{t \in [0,a]} |f_n(t) - f(t)|\, \mathrm{d}t = x\, \alpha_n \leqslant a\, \alpha_n,$$

- et pour $x \in]a, +\infty[$,

$$\begin{aligned} |F_n(x) - F(x)| &\leqslant \int_0^a |f_n(t) - f(t)|\, \mathrm{d}t + \int_a^x |\underbrace{f_n(t)}_{=0} - f(t)|\, \mathrm{d}t \\ &\leqslant a\, \alpha_n + \int_a^x |f(t)|\, \mathrm{d}t \leqslant a\, \alpha_n + \int_a^{+\infty} \mathrm{e}^{-t^2}\, \mathrm{d}t. \end{aligned}$$

Cela implique que pour tout $n \in \mathbb{N}$ avec $n > \mathrm{E}(a^2)$ on a

$$v_n = \max\Big(\sup_{x \in [0,a]} |F_n(x) - F(x)|, \sup_{x \in]a, +\infty[} |F_n(x) - F(x)| \Big) \leqslant a\, \alpha_n + \int_a^{+\infty} \mathrm{e}^{-t^2}\, \mathrm{d}t.$$

Comme $\lim_{n \to +\infty} \alpha_n = 0$, la limite ℓ (si elle existe) de la suite $(v_n)_n$ vérifie $0 \leqslant \ell \leqslant \int_a^{+\infty} \mathrm{e}^{-t^2}\, \mathrm{d}t$. Cette relation est vraie pour tout réel $a > 0$. Comme l'intégrale généralisée $\int_0^{+\infty} \mathrm{e}^{-t^2}\, \mathrm{d}t$ converge, on a $\lim_{a \to +\infty} \int_a^{+\infty} \mathrm{e}^{-t^2}\, \mathrm{d}t = 0$. Le théorème d'encadrement permet d'en déduire que $\ell = 0$, autrement dit que la suite $(v_n)_n$ converge vers 0. On en conclut que la suite $(F_n)_n$ converge uniformément sur $[0, +\infty[$.

b) Considérons le changement de variable défini par l'application $\phi : \theta \in [0, \frac{\pi}{2}] \mapsto \sqrt{n}\sin(\theta) \in [0, \sqrt{n}]$. L'application ϕ est clairement de classe $\mathcal{C}^1$ et en vertu de la formule de changement de variable pour une intégrale[36] on obtient

$$\begin{aligned}\ell_n &= \int_{\phi(0)}^{\phi(\frac{\pi}{2})} f_n(t)\,\mathrm{d}t = \int_0^{\frac{\pi}{2}} f_n(\phi(\theta))\,\phi'(\theta)\,\mathrm{d}\theta \\ &= \int_0^{\frac{\pi}{2}} (1-\sin^2(\theta))\,\sqrt{n}\cos(\theta)\,\mathrm{d}\theta = \sqrt{n}\int_0^{\frac{\pi}{2}} \cos^{2n+1}(\theta)\,\mathrm{d}\theta.\end{aligned}$$

Pour montrer que $\ell_n = \sqrt{n}\frac{(2^n\,n!)^2}{(2n+1)!}$, procédons par récurrence. La relation est vraie pour $n = 1$ puisque

$$\ell_1 = \int_0^{\frac{\pi}{2}} \cos^3(\theta)\,\mathrm{d}\theta = \int_0^{\frac{\pi}{2}} (1-\sin^2(\theta))\cos(\theta)\,\mathrm{d}\theta = \left[\sin(\theta) - \tfrac{1}{3}\sin^3(\theta)\right]_0^{\frac{\pi}{2}} = \frac{2}{3}$$

et si $n = 1$ alors $\sqrt{n}\frac{(2^n\,n!)^2}{(2n+1)!} = \frac{2^2}{3!} = \frac{2}{3}$. Supposons, c'est notre hypothèse de récurrence, que $\ell_n = \sqrt{n}\frac{(2^n\,n!)^2}{(2n+1)!}$. En intégrant par parties, on obtient :

$$\begin{aligned}\ell_{n+1} &= \sqrt{n+1}\int_0^{\frac{\pi}{2}} \cos^{2n+3}(\theta)\,\mathrm{d}\theta \\ &= \sqrt{n+1}\,(2n+2)\int_0^{\frac{\pi}{2}} \cos^{2n+1}(\theta)\,\underbrace{\sin^2(\theta)}_{=1-\cos^2(\theta)}\,\mathrm{d}\theta \\ &= \sqrt{n+1}\,(2n+2)\left(\frac{1}{\sqrt{n}}\ell_n - \frac{1}{\sqrt{n+1}}\ell_{n+1}\right).\end{aligned}$$

On en déduit que

$$\begin{aligned}\ell_{n+1} &= \sqrt{\frac{n+1}{n}}\,\frac{2n+2}{2n+3}\ell_n = \sqrt{n+1}\,(2n+2)^2\,\frac{(2^n\,n!)^2}{(2n+1)!(2n+2)(2n+3)} \\ &= \sqrt{n+1}\,\frac{(2^{n+1}\,(n+1)!)^2}{(2n+3)!}.\end{aligned}$$

La relation est démontrée pour l'entier $n+1$. On déduit de ce raisonnement par récurrence que pour tout $n \in \mathbb{N}^*$ on a $\ell_n = \sqrt{n}\frac{(2^n\,n!)^2}{(2n+1)!}$.

[36] Voir le théorème 18.2 p. 863 du *Cours de première année*.

D'après la formule de Stirling, on a $n! \underset{+\infty}{\sim} \sqrt{2\pi n}\left(\frac{n}{\mathrm{e}}\right)^n$ et par conséquent

$$\begin{aligned}
\ell_n &\underset{+\infty}{\sim} \sqrt{n}\,(2^n)^2 \left(\sqrt{2\pi n}\left(\frac{n}{\mathrm{e}}\right)^n\right)^2 \frac{1}{\sqrt{2\pi(2n+1)}}\left(\frac{\mathrm{e}}{2n+1}\right)^{2n+1} \\
&= \sqrt{n}\,(2^n)^2\, 2\pi n \left(\frac{n}{\mathrm{e}}\right)^{2n} \frac{1}{\sqrt{2\pi}\sqrt{2n+1}}\left(\frac{\mathrm{e}}{2n+1}\right)^{2n+1} \\
&= (2^n)^2 \frac{\sqrt{n}}{\sqrt{2n+1}}\sqrt{2\pi}\,\mathrm{e}\left(\frac{n}{2n+1}\right)^{2n+1} \\
&= \frac{\sqrt{n}}{\sqrt{2n+1}}\frac{\sqrt{2\pi}\,\mathrm{e}}{2}\left(\frac{n}{2n+1}\right)^{2n+1}
\end{aligned}$$

On a $\lim_{n\to+\infty} \frac{\sqrt{n}}{\sqrt{2n+1}} = \frac{1}{\sqrt{2}}$ et $\left(\frac{n}{2n+1}\right)^{2n+1} = \exp(-(2n+1)\ln(1+\frac{1}{2n}))$ admet pour limite $1/\mathrm{e}$ car $(2n+1)\ln(1+\frac{1}{2n}) \underset{+\infty}{\sim} \frac{2n+1}{2n} \underset{+\infty}{\sim} 1$. On en déduit que la suite $(\ell_n)_n$ converge vers $\frac{1}{2}\sqrt{\pi}$.

3 - Appliquons la proposition 2.7 à la suite $(F_n)_n$. Cette suite converge uniformément sur $[0,+\infty[$ vers F et les applications F_n admettent toutes une limite quand x tend vers $+\infty$ qui est $\int_0^{+\infty} f_n(t)\,\mathrm{d}t = \int_0^{\sqrt{n}} f_n(t)\,\mathrm{d}t = \ell_n$. On a donc,

$$\begin{aligned}
\int_0^{+\infty} \mathrm{e}^{-t^2}\,\mathrm{d}t &= \lim_{x\to+\infty} F(x) = \lim_{x\to+\infty}\lim_{n\to+\infty} F_n(x) \\
&= \lim_{n\to+\infty}\lim_{x\to+\infty} F_n(x) = \lim_{n\to+\infty} \ell_n = \frac{1}{2}\sqrt{\pi}.
\end{aligned}$$

On en conclut que $\int_0^{+\infty} \mathrm{e}^{-t^2}\,\mathrm{d}t = \frac{1}{2}\sqrt{\pi}$.

Solution de l'exercice 2.10

1 - Pour $x \neq 1$ on a

$$f_n(x) = \left[\frac{1}{1-x}\frac{1}{t^{x-1}}\right]_1^n = \frac{1}{1-x}\left(\frac{1}{n^{x-1}-1}\right) \quad \text{et} \quad f_n(1) = \int_1^n \frac{1}{t}\,\mathrm{d}t = \ln(n).$$

2 - a) Pour $x \neq 1$ on a

$$\frac{1}{x-1}\left(1 - \frac{1}{n^{x-1}}\right) = f_n(x) = \frac{1}{x-1}\left(1 - \exp((1-x)\ln(n))\right).$$

Or $\mathrm{e}^u = 1 + u + \frac{1}{2}u^2 + o_0(u^2)$ donc

$$\exp((1-x)\ln(n)) = 1 + (1-x)\ln(n) + \tfrac{1}{2}(1-x)^2\ln^2(n) + o_1((1-x)^2)$$

et

$$f_n(x) = \ln(n) - \frac{1}{2}(1-x)\ln^2(n) + o_1(1-x).$$

b) L'application f_n est dérivable sur $\mathbb{R}\setminus\{1\}$ car c'est le produit de la fonction rationnelle $x \mapsto \frac{1}{1-x}$ dont le seul pôle est 1 par la fonction $x \mapsto ne^{-x\ln(n)} - 1$ qui est dérivable sur $\mathbb{R}$ car la fonction exponentielle est dérivable sur $\mathbb{R}$. Par ailleurs, l'application f_n admettant en 1 un développement limité d'ordre 1, elle est dérivable[37] en 1 et

$$f_n'(1) = -\frac{1}{2}\ln^2(n).$$

On en conclut que f_n est dérivable sur $\mathbb{R}$, et par conséquent continue sur $\mathbb{R}$.

Pour $x \in \mathbb{R} \setminus \{1\}$, on a

$$f_n'(x) = -\frac{1}{(x-1)^2} + \frac{1}{(x-1)\,n^{x-1}}\left(\ln(n) + \frac{1}{x-1}\right).$$

c) L'application f_n est de classe $\mathcal{C}^1$ sur $\mathbb{R}$ si f_n' est continue sur $\mathbb{R}$. Sur $\mathbb{R}\setminus\{1\}$, l'application f_n' est continue car d'une part les fonctions rationnelles $x \mapsto \frac{1}{1-x}$ et $x \mapsto \frac{1}{(1-x)^2}$, qui possèdent pour unique pôle 1, sont continues sur $\mathbb{R}\setminus\{1\}$ et d'autre part la fonction $x \mapsto ne^{-x\ln(n)}$ est continue sur $\mathbb{R}$ car la fonction exponentielle est continue sur $\mathbb{R}$. Il reste à établir que f_n' est continue en 1. On a vu que :

$$\frac{1}{n^{x-1}} = \exp((1-x)\ln(n)) = 1 + (1-x)\ln(n) + \tfrac{1}{2}(1-x)^2\ln(n)^2 + o_1((1-x)^2).$$

On en déduit que :

$$\begin{aligned} f_n'(x) &= -\frac{1}{(x-1)^2} + \frac{1}{x-1}\left(\ln(n) + \frac{1}{x-1}\right) \times \Big(1 + (1-x)\ln(n) + \\ &\qquad \tfrac{1}{2}(1-x)^2\ln(n)^2 + o_1((1-x)^2)\Big) \\ &= -\frac{1}{2}\ln(n)^2 + o_1(1). \end{aligned}$$

Cela implique que $\lim_{x\to 1} f_n'(x) = -\frac{1}{2}\ln(n)^2 = f_n'(1)$, et par conséquent que f_n' est continue en 1. On en conclut que f_n est de classe $\mathcal{C}^1$ sur $\mathbb{R}$.

3 - Pour $x > 1$ on a $\lim_{n\to+\infty} \exp((1-x)\ln(n)) = 0$ et pour $x < 1$ on a $\lim_{n\to+\infty} \exp((1-x)\ln(n)) = +\infty$. Par ailleurs, pour $x > 1$ on a

$$\lim_{n\to+\infty} \ln(n)\,\exp((1-x)\ln(n)) = \lim_{p\to+\infty} p\,\exp((1-x)p) = 0.$$

On déduit de l'expression de $f_n'(x)$ que la suite numérique de terme général $f_n'(x)$ converge si et seulement si $x > 1$. Sa limite est $-\frac{1}{(x-1)^2}$. Cela signifie que la suite de fonctions $(f_n')_n$ converge simplement sur $]1,+\infty[$ vers $g : x \in]1,+\infty[\mapsto -\frac{1}{(x-1)^2}$.

[37] Voir la proposition 17.3 p. 776 du *Cours de première année*.

La suite $(f_n')_n$ convergera uniformément vers g sur $]1,+\infty[$ si la quantité $u_n = \sup_{x\in]1,+\infty[} |f_n'(x) - g(x)|$ existe et admet pour limite 0 quand n tend vers $+\infty$. Considérons l'application $\phi : x \in]1,+\infty[\mapsto f_n'(x) - g(x)$ pour $n \in \mathbb{N}^*$ fixé. Pour tout $x \in]1,+\infty[$, on a

$$\phi(x) = \frac{1}{(x-1)\,n^{x-1}} \left(\ln(n) + \frac{1}{x-1}\right)$$

$$\text{et} \quad \phi'(x) = -\frac{1}{(x-1)^3\,n^{x-1}} \left(2 + 2(x-1)\ln(n) + (x-1)^2 \ln(n)^2\right).$$

Pour tout $x \in]1,+\infty[$, on a $\phi'(x) < 0$; on en déduit que ϕ est strictement décroissante sur $]1,+\infty[$. Comme $\lim_{x\to 1^+} \phi(x) = +\infty$, l'application ϕ n'est pas bornée et par conséquent la convergence sur $]1,+\infty[$ de la suite $(f_n')_n$ vers g n'est pas uniforme.

4 - a) Pour $x > 1$ on a $\lim_{n\to+\infty} \exp((1-x)\ln(n)) = 0$ et pour $x < 1$ on a $\lim_{n\to+\infty} \exp((1-x)\ln(n)) = +\infty$. On déduit de l'expression de $f_n(x)$ que la suite numérique de terme général $f_n(x)$ converge si et seulement si $x > 1$. Sa limite est $\frac{1}{x-1}$. Cela signifie que la suite de fonctions $(f_n)_n$ converge simplement sur $]1,+\infty[$ vers $f : x \in]1,+\infty[\mapsto \frac{1}{x-1}$.

b) Le théorème 2.5 indique que si $(f_n)_n$ est une suite d'applications de $I =]1,+\infty[$ dans $\mathbb{R}$ telle que :

1. pour tout $n \in \mathbb{N}$ l'application f_n est de classe $\mathcal{C}^1$ sur I ;
2. la suite d'applications $(f_n')_n$ converge uniformément sur I vers une application g de I dans $\mathbb{R}$;
3. il existe un réel $x_0 \in I$ tel que la suite numérique de terme général $f_n(x_0)$ converge ;

alors la suite $(f_n)_n$ converge simplement sur I vers une application f qui est de classe $\mathcal{C}^1$ sur I et qui admet pour dérivée $f' = g$.

Dans le cas traité ici, les conditions 1 et 3 sont satisfaites mais pas la condition 2. La conclusion du théorème est toutefois vraie. La seconde condition du théorème est donc une condition suffisante mais ce n'est pas une condition nécessaire.

CHAPITRE 3

Séries de fonctions

Dans le chapitre précédent nous avons mené l'étude des suites de fonctions. De même que nous avons introduit dans le chapitre 1 la notion de série numérique à partir de celle de suite numérique, nous allons dans ce chapitre utiliser les suites de fonctions pour définir la notion de série de fonctions. Nous nous limitons dans ce chapitre aux séries de fonctions réelles d'une variable réelle. Nous aborderons la notion de séries de fonctions complexes d'une variable complexe dans le chapitre suivant.

3.1 Convergence d'une série de fonctions

On désigne par X un sous-ensemble non vide de $\mathbb{R}$, par $\mathcal{A}(X,\mathbb{R})$ l'ensemble des applications de X dans $\mathbb{R}$ et par $\mathcal{S}(X,\mathbb{R})$ l'ensemble des suites d'applications de X dans $\mathbb{R}$.

3.1.1 Définitions

DÉFINITION 3.1 *Soient $(f_n)_n$ une suite d'applications de X dans $\mathbb{R}$ et $(\phi_n)_n$ la suite d'applications de terme général $\phi_n = \sum_{k=0}^{n} f_k$.*
On appelle série d'applications de X dans $\mathbb{R}$ de terme général f_n *et on note* $\sum_n f_n$, *le couple* $((f_n)_n, (\phi_n)_n)$ *d'éléments de* $\mathcal{S}(X,\mathbb{R}) \times \mathcal{S}(X,\mathbb{R})$.

On note $\Sigma(X,\mathbb{R})$ l'ensemble des séries d'applications de X dans $\mathbb{R}$.

Remarques

1. On confond très souvent du point de vue terminologique *série de fonctions* avec *série d'applications*. Nous utiliserons le terme « série d'applications » lorsque les termes de la série seront définis sur un même ensemble qui sera précisé. Le terme « série de fonctions » sera utilisé comme terme générique. Ainsi, on parlera de la série de fonctions de terme général $f_n : x \mapsto \ln(x - \frac{1}{n})$ et de la série d'applications de terme général $f_n : x \in]1, +\infty[\mapsto \ln(x - \frac{1}{n})$.

2. L'ensemble X sur lequel est considérée la série $\sum_n f_n$ doit correspondre à un ensemble sur lequel toutes les fonctions f_n sont définies. Par exemple, pour tout $n \in \mathbb{N}^*$, la fonction $f_n : x \mapsto \frac{1}{2^n} \tan\left(\frac{x}{2^n}\right)$ est $2^{n-1}\pi$ périodique et définie

sur $D_n =]-2^{n-1}\pi, 2^{n-1}\pi[$. La série de fonctions $\sum_n f_n$ ne peut être étudiée que sur $X = \bigcap_{n\in\mathbb{N}} D_n =]-\frac{\pi}{2}, \frac{\pi}{2}[$. ◇

On peut associer à une suite d'applications $(f_n)_n$, sa *suite des sommes partielles* $(\phi_n)_n$ dont le terme général est $\phi_n = \sum_{k=0}^{n} f_k$, ce qui permet de définir la série $\sum_n f_n$. Inversement, à une suite d'applications $(\phi_n)_n$ on peut associer une suite d'applications $(f_n)_n$, et par conséquent une série d'applications $\sum_n f_n$, dont $(\phi_n)_n$ est la suite des sommes partielles de la manière suivante :

$$\text{on pose } \quad f_0 = \phi_0 \quad \text{ et pour tout } n \in \mathbb{N}^* \quad f_n = \phi_n - \phi_{n-1}.$$

On peut par ce procédé appliquer aux séries d'applications la plupart des résultats du chapitre précédent consacré aux suites d'applications. On peut en particulier définir la notion de convergence d'une série d'applications et celle de somme d'une série d'applications.

DÉFINITION 3.2 (Convergence simple)

✗ *On dit que la série d'applications $\sum_n f_n$ de X dans $\mathbb{R}$* converge simplement *sur X si la suite d'applications $(\phi_n)_n$ de terme général $\phi_n = \sum_{k=0}^{n} f_k$ converge simplement sur X.*

✗ *On appelle alors* somme de la série d'applications $\sum_n f_n$ *la limite ϕ de la suite d'applications $(\phi_n)_n$; pour $x \in X$, on note*

$$\phi(x) = \sum_{n=0}^{+\infty} f_n(x).$$

Si la série d'applications $\sum_n f_n$ de X dans $\mathbb{R}$ converge simplement sur X vers l'application ϕ de X dans $\mathbb{R}$, alors pour tous $x \in X$ et $n \in \mathbb{N}$ on a

$$\phi(x) = \sum_{k=0}^{+\infty} f_k(x) = \sum_{k=0}^{n} f_k(x) + \sum_{k=n+1}^{+\infty} f_k(x) = \phi_n(x) + \sum_{k=n+1}^{+\infty} f_k(x).$$

On déduit donc de la définition 3.2 la proposition suivante.

PROPOSITION 3.1 *Une condition nécessaire et suffisante pour que la série d'applications $\sum_n f_n$ de X dans $\mathbb{R}$ converge simplement sur X est que la suite d'applications $(\rho_n)_n$ de terme général $\rho_n = \sum_{k=n+1}^{+\infty} f_k$ converge sur X vers l'application nulle.*

D'après la définition 2.1 p. 71, la suite d'applications $(\phi_n)_n$ de terme général $\phi_n = \sum_{k=0}^{n} f_k$ converge simplement vers ϕ sur X si pour tout $x \in X$ la suite numérique réelle $(\phi_n(x))_{n\in\mathbb{N}}$ converge vers $\phi(x)$. Or, la suite numérique réelle $(\phi_n(x))_{n\in\mathbb{N}}$ peut être vue comme la suite des sommes partielles associée à la série numérique réelle $\sum_n f_n(x)$. On peut ainsi également relier les propriétés

de convergence d'une série d'applications à celles des séries numériques étudiées au chap. 1. En particulier, on a le résultat suivant.

PROPOSITION 3.2 *La série d'applications $\sum_n f_n$ de X dans $\mathbb{R}$ converge simplement sur X vers l'application ϕ de X dans $\mathbb{R}$ si et seulement si pour tout $x \in X$ la série numérique $\sum_n f_n(x)$ converge vers le réel $\phi(x)$.*

Exemples

1. La série d'applications $\sum_n f_n$ de terme général $f_n : x \in \mathbb{R} \mapsto \frac{1}{n!} x^n$ converge simplement sur $\mathbb{R}$ et a pour somme l'application exponentielle, voir les exemples de la p. 5.

2. La série d'applications $\sum_n f_n$ de terme général $f_n : x \in \mathbb{R} \mapsto \frac{(-1)^n}{n} x^n$ converge simplement sur $[0,1]$. En effet, pour tout $x \in [0,1[$ fixé, on a

$$\lim_{n\to+\infty} \left| \frac{f_{n+1}(x)}{f_n(x)} \right| = |x| < 1$$

et le critère de D'Alembert permet de conclure que la série numérique $\sum_n f_n(x)$ converge. Par ailleurs, $\sum_n f_n(1)$ est la série harmonique alternée qui est convergente. La série d'applications $\sum_n f_n(1)$ converge donc simplement sur $[0,1]$. Montrons que la somme de cette série est $\phi : x \in [0,1] \mapsto -\ln(1+x)$. En 0, on a la série nulle qui converge bien vers $\phi(0) = 0$. Intéressons-nous à la convergence sur $]0,1]$. L'application ϕ est de classe $\mathcal{C}^\infty$ sur $]0,1]$ et pour tout $k \in \mathbb{N}^*$, on vérifie en utilisant un raisonnement par récurrence, que la dérivée k–ième de ϕ est donnée par

$$\phi^{(k)}(x) = (-1)^k \frac{(k-1)!}{(1+x)^k}.$$

La formule de Taylor-Lagrange[1] à l'ordre n appliquée à ϕ entre 0 et x, où x est fixé dans $]0,1]$, nous assure l'existence d'un réel $c \in]0,x[$ tel que

$$\phi(x) = \sum_{k=0}^{n} \phi^{(k)}(0) \frac{x^k}{k!} + \frac{x^{n+1}}{(n+1)!} \phi^{(n+1)}(c) = \sum_{k=1}^{n} (-1)^k \frac{x^k}{k} + R_n(x) \qquad (1)$$

où $R_n(x) = \dfrac{(-1)^{n+1}}{n+1} \dfrac{x^{n+1}}{(1+c)^{n+1}}$. On a donc $\phi(x) = \sum_{k=1}^{n} f_k(x) + R_n(x)$. Pour $c \in]0,x|$ et $x \in]0,1]$ on a les majorations

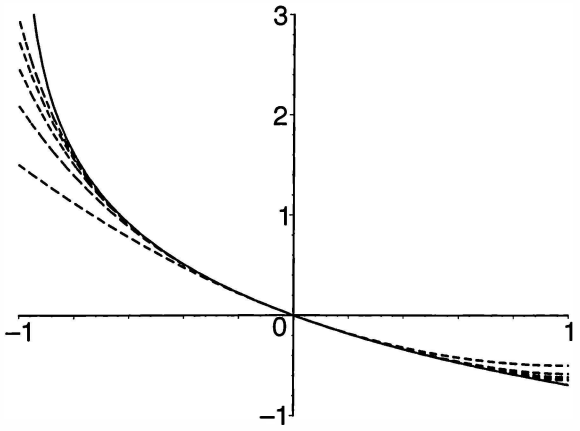
$$0 \leqslant |R_n(x)| = \frac{1}{n+1} \frac{x^{n+1}}{(1+c)^{n+1}} \leqslant \frac{x^{n+1}}{n+1} \leqslant \frac{1}{n+1}$$

qui impliquent que $\lim_{n\to+\infty} R_n(x) = 0$ pour tout $x \in]0,1]$. On en conclut que la somme de la série numérique $\sum_n f_n(x)$ est $\phi(x)$ pour tout $x \in]0,1]$.

On peut visualiser graphiquement la convergence de cette série d'applications à l'aide de MAPLE.

[1] Voir le théorème 16.4 p. 750 du *Cours de première année*.

```
> f_n:= (n,x)-> (-1)^n*x^n/n: phi:= x-> -ln(1+x):
> phi_n:= (n,x)-> sum(f_n(k,x),k=1..n):
> E:=seq(2*n,n=1..5):
> plot([seq(phi_n(n,x),n=E),phi(x)],x=-1..1,y=-3..3,
                    linestyle=[seq(DASH,n=E),SOLID]);
```



On pourra observer sur la figure que la série de fonctions $\sum_n f_n$ semble également converger vers ϕ sur $]-1,0[$; ce résultat est toutefois plus délicat à établir que dans la cas de l'intervalle $[0,1]$.

3. Considérons la série d'applications $\sum_n f_n$ où $f_n : x \in]-1,1[\mapsto x^n$. Le terme général de la suite $(\phi_n)_n$ des sommes partielles associée à cette série est donné par[2] :

$$\forall x \in]-1,1[\qquad \phi_n(x) = \sum_{k=0}^{n} x^k = \frac{1-x^{n+1}}{1-x}.$$

La suite d'applications $(\phi_n)_n$ converge simplement sur $]-1,1[$ vers l'application $\phi : x \in]-1,1[\mapsto 1/(1-x)$. La série $\sum_n f_n$ converge donc simplement sur $]-1,1[$ et a pour somme l'application ϕ.

4. Intéressons-nous à la convergence simple de la série d'applications $\sum_n f_n$ de terme général $f_n : x \in]-\pi/2,\pi/2[\mapsto \frac{1}{2^n}\tan(x/2^n)$. Pour tout $y \in]-\frac{\pi}{2},0[\cup]0,\frac{\pi}{2}[$ on a la relation[3] : $\tan\left(\frac{y}{2}\right) = \operatorname{cotan}\left(\frac{y}{2}\right) - 2\operatorname{cotan}(y)$. On en déduit les égalités

[2] Voir la proposition 5.21 p. 195 du *Cours de première année*.

[3] On rappelle que la fonction cotangente, notée cotan, est l'inverse de la fonction tangente : $\operatorname{cotan} = \frac{1}{\tan}$. Ne pas la confondre avec la fonction arc-tangente (notée arctan) qui est la bijection réciproque de la fonction tangente, voir le *Cours de première année* p. 653. Pour établir cette relation, on pourra étudier la fonction $x \mapsto \tan\left(\frac{y}{2}\right) - \operatorname{cotan}\left(\frac{y}{2}\right) + 2\operatorname{cotan}(y)$ et suivre la méthode utilisée dans la démonstration de la proposition 14.19 p. 655 du *Cours de première année*.

suivantes pour tout $n \in \mathbb{N}^*$ et pour tout $x \in]-\frac{\pi}{2}, 0[\cup]0, \frac{\pi}{2}[$:

$$\begin{aligned}
\phi_n(x) &= \sum_{k=0}^{n} f_k(x) = \sum_{k=0}^{n} \frac{1}{2^k} \tan\left(\frac{x}{2^k}\right) \\
&= \tan(x) + \sum_{k=1}^{n} \left(\frac{1}{2^k} \operatorname{cotan}\left(\frac{x}{2^k}\right) - \frac{1}{2^{k-1}} \operatorname{cotan}\left(\frac{x}{2^{k-1}}\right) \right) \\
&= \tan(x) + \sum_{k=1}^{n} \frac{1}{2^k} \operatorname{cotan}\left(\frac{x}{2^k}\right) - \sum_{k=0}^{n-1} \frac{1}{2^k} \operatorname{cotan}\left(\frac{x}{2^k}\right) \\
&= \tan(x) + \frac{1}{2^n} \operatorname{cotan}\left(\frac{x}{2^n}\right) - \operatorname{cotan}(x). \qquad (2)
\end{aligned}$$

L'équivalent $\tan(u) \underset{0}{\sim} u$ permet d'obtenir que pour tout $x \in]-\frac{\pi}{2}, 0[\cup]0, \frac{\pi}{2}[$,

$$\frac{1}{2^n} \operatorname{cotan}\left(\frac{x}{2^n}\right) = \frac{1}{2^n \tan\left(\frac{x}{2^n}\right)} \underset{n\to+\infty}{\sim} \frac{1}{x}.$$

On en déduit que pour tout $x \in]-\frac{\pi}{2}, 0[\cup]0, \frac{\pi}{2}[$,

$$\lim_{n\to+\infty} \phi_n(x) = \tan(x) + \frac{1}{x} - \operatorname{cotan}(x).$$

Comme de plus, pour tout $n \in \mathbb{N}$, on a $f_n(0) = 0$, la suite numérique $(\phi_n(0))_n$ converge vers 0. On en conclut que la suite d'applications $(\phi_n)_n$, et par conséquent la série $\sum_n f_n$, convergent simplement sur $]-\frac{\pi}{2}, \frac{\pi}{2}[$. La limite de la suite $(\phi_n)_n$, qui coïncide avec la somme de la série $\sum_n f_n$, est l'application

$$\phi : x \in]-\frac{\pi}{2}, \frac{\pi}{2}[\longmapsto \begin{cases} \tan(x) + 1/x - \operatorname{cotan}(x) & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases}.$$

5. La série d'applications $\sum_n f_n$ de terme général $f_n : x \mapsto n\mathrm{e}^{-nx}$ converge simplement sur $\mathbb{R}_+^*$. En effet, comme pour tout $x \in \mathbb{R}_+^*$ la quantité

$$\frac{f_{n+1}(x)}{f_n(x)} = \frac{n+1}{n} \mathrm{e}^{-x}$$

admet pour limite e^{-x} quand n tend vers $+\infty$ et comme $\mathrm{e}^{-x} < 1$ pour $x \in \mathbb{R}_+^*$, le critère de D'Alembert assure la convergence de la série numérique $\sum_n n\mathrm{e}^{-nx}$ pour tout $x \in \mathbb{R}_+^*$. Remarquons qu'il y a divergence grossière sur $\mathbb{R}^-$ car pour tout $x \in \mathbb{R}^-$ on a $\lim_{n\to+\infty} n\mathrm{e}^{-nx} = +\infty$. Nous montrerons ultérieurement la somme de cette série d'applications est $\phi : x \in \mathbb{R}_+^* \mapsto \mathrm{e}^x/(\mathrm{e}^x - 1)^2$.

Exercice 3.1 Étudier la convergence simple et déterminer la somme des séries de fonctions $\sum_n f_n$ dont le terme général est :
1 - $f_n : x \mapsto \cos(x)^n$ 2 - $f_n : x \mapsto \exp(-nx)$ 3 - $f_n : x \mapsto x(1-x)^n$

On veillera à ne pas confondre la somme de la série d'applications $\sum_n f_n$ et la limite de la suite d'applications $(f_n)_n$. La somme de la série d'applications $\sum_n f_n$ est la limite de la suite d'applications $(\phi_n)_n$ de terme général $\phi_n = \sum_{k=0}^{n} f_k$. Par ailleurs, le résultat suivant indique que si la série d'applications $\sum_n f_n$ converge simplement alors la limite de la suite d'applications $(f_n)_n$ est nécessairement l'application nulle.

PROPOSITION 3.3 *Si la série d'applications $\sum_n f_n$ de X dans $\mathbb{R}$ converge simplement sur X alors la suite d'applications $(f_n)_n$ converge simplement sur X vers l'application nulle.*

Démonstration Cette propriété résulte de la proposition 1.2 p. 11 qu'il suffit d'appliquer pour tout $x \in X$ à la série numérique $\sum_n f_n(x)$. □

Comme pour les séries numériques, il existe pour les séries de fonctions plusieurs notions de convergence, dont celle de convergence absolue.

DÉFINITION 3.3 (Convergence absolue)
On dit que la série d'applications $\sum_n f_n$ de X dans $\mathbb{R}$ converge absolument *sur X si la série d'applications $\sum_n |f_n|$ converge simplement sur X.*

D'après la proposition 3.2, la série d'applications $\sum_n |f_n|$ converge simplement sur X si et seulement si pour tout $x \in X$ la série numérique $\sum_n |f_n(x)|$ converge, autrement dit si pour tout $x \in X$ la série numérique $\sum_n f_n(x)$ converge absolument. On a donc la proposition suivante qui permet d'établir la convergence absolue d'une série d'applications en ayant recours aux méthodes vues dans le chap. 1 (on pensera en particulier aux critères de Cauchy et de D'Alembert).

PROPOSITION 3.4 *La série d'applications $\sum_n f_n$ de X dans $\mathbb{R}$ converge absolument sur X si et seulement si pour tout $x \in X$ la série numérique $\sum_n f_n(x)$ converge absolument.*

Reprenons les différents exemples de la p. 115 et étudions la convergence absolue de ces séries d'applications.

Exemples

1. La série d'applications $\sum_n f_n$ de terme général $f_n : x \in \mathbb{R} \mapsto \frac{1}{n!} x^n$ converge absolument sur $\mathbb{R}$. En effet, pour tous $n \in \mathbb{N}$ et $x \in \mathbb{R}$, la quantité

$$\left| \frac{f_{n+1}(x)}{f_n(x)} \right| = \frac{|x|}{n+1}$$

tend vers 0 lorsque n tend vers $+\infty$. Le critère de D'Alembert indique que pour tout $x \in \mathbb{R}$ la série numérique $\sum_n f_n(x)$ converge absolument ce qui signifie que la série d'applications $\sum_n f_n$ converge absolument sur $\mathbb{R}$.

2. Nous avons vu que la série d'applications $\sum_n f_n$ dont le terme général est $f_n : x \in \mathbb{R} \mapsto \frac{(-1)^n}{n} x^n$ convergeait simplement sur $[0,1]$. Le critère de D'Alembert permet d'établir que la série d'applications $\sum_n f_n$ converge absolument sur $[0,1[$ car pour tous $n \in \mathbb{N}$ et $x \in [0,1[$, la quantité

$$\left| \frac{f_{n+1}(x)}{f_n(x)} \right| = \frac{n}{n+1} x$$

tend vers x quand n tend vers $+\infty$ et $x < 1$. Pour $x = 1$, le critère de D'Alembert ne permet pas de conclure ; toutefois $f_n(1) = \frac{(-1)^n}{n}$ et la série harmonique alternée $\sum_n \frac{(-1)^n}{n}$ est semi-convergente. Il n'y a pas donc convergence absolue en 1 de la série d'applications $\sum_n f_n$.

3. Le critère de D'Alembert permet également d'établir que la série d'applications $\sum_n f_n$ de terme général $f_n : x \in]-1,1[\mapsto x^n$ converge absolument sur $]-1,1[$.

4. La série d'applications $\sum_n f_n$ de terme général $f_n : x \in]-\frac{\pi}{2}, \frac{\pi}{2}[\mapsto \frac{1}{2^n} \tan\left(\frac{x}{2^n}\right)$ converge absolument sur $]-\frac{\pi}{2}, \frac{\pi}{2}[$. En effet, pour $x \in]-\frac{\pi}{2}, 0[\cup]0, \frac{\pi}{2}[$ fixé on a $\lim_{n \to +\infty} x/2^n = 0$ ce qui, compte tenu de la relation $\tan(u) \underset{0}{\sim} u$, implique que[4]

$$|f_n(x)| \underset{n \to +\infty}{\sim} \frac{|x|}{4^n}.$$

Comme la série numérique $\sum_n \frac{|x|}{4^n}$ converge pour tout $x \in \mathbb{R}$ (on peut utiliser le critère de D'Alembert pour l'établir), la série numérique $\sum_n f_n(x)$ converge absolument pour tout $x \in]-\frac{\pi}{2}, 0[\cup]0, \frac{\pi}{2}[$. Pour $x = 0$, la série $\sum_n f_n(x)$ est la série nulle qui converge également absolument.

La suite des sommes partielles associée à une série d'applications permet de définir la notion de convergence uniforme pour les séries d'applications à partir de la notion de convergence uniforme pour les suites d'applications, voir la définition 2.2 p. 73.

---

[4] On sera vigilant au fait que le cas $x = 0$ soit être traité à part car $\frac{|x|}{4^n}$ est alors nul, voir le chap. 15 du *Cours de première année*, en particulier la définition 15.3 p. 688.

DÉFINITION 3.4 (Convergence uniforme)
On dit que la série d'applications $\sum_n f_n$ *de* X *dans* $\mathbb{R}$ converge uniformément *sur* X *si la suite des sommes partielles* $(\phi_n)_n$ *de terme général* $\phi_n : x \in X \mapsto \sum_{k=0}^{n} f_k(x)$ *est une suite d'applications qui converge uniformément sur* X.

On déduit de la proposition 2.2 p. 74 le résultat suivant.

PROPOSITION 3.5 *Soit* $\sum_n f_n$ *une série d'applications de* X *dans* $\mathbb{R}$ *convergeant simplement sur* X *et admettant pour somme l'application* ϕ *de* X *dans* $\mathbb{R}$.
Une condition nécessaire et suffisante pour que la série $\sum_n f_n$ *converge uniformément sur* X *est que la suite numérique de terme général*

$$u_n = \sup_{x \in X} \left| \phi(x) - \sum_{k=0}^{n} f_k(x) \right|$$

soit définie et converge vers 0.

En introduisant la suite d'applications $(\rho_n)_n$ de X dans $\mathbb{R}$, où pour tous $n \in \mathbb{N}$ et $x \in X$

$$\rho_n(x) = \phi(x) - \sum_{k=0}^{n} f_k(x) = \sum_{k=n+1}^{+\infty} f_k(x),$$

on peut reformuler la proposition 3.5 de la manière suivante.

COROLLAIRE 3.1 *Une condition nécessaire et suffisante pour que la série d'applications* $\sum_n f_n$ *de* X *dans* $\mathbb{R}$ *converge uniformément sur* X *est que la suite d'applications* $(\rho_n)_n$ *de terme général* $\rho_n : x \in X \longmapsto \sum_{k=n+1}^{+\infty} f_k(x)$ *converge uniformément sur* X *vers l'application nulle.*

On pourra comparer l'énoncé du corollaire 3.1 à celui de la proposition 3.1.

Reprenons les différents exemples de la p. 115 et étudions la convergence uniforme de ces séries d'applications.

Exemples

1. Montrons que la série d'applications $\sum_n f_n$ dont le terme général est $f_n : x \in \mathbb{R} \mapsto \frac{(-1)^n}{n} x^n$ converge uniformément sur $[0,1]$. Nous avons établi que la somme de cette série était l'application $\phi : x \in [0,1] \mapsto -\ln(1+x)$. D'après la relation (1) p. 115, on a pour tous $n \in \mathbb{N}$ et $x \in [0,1]$

$$\rho_n(x) = \phi(x) - \sum_{k=0}^{n} f_k(x) = \frac{(-1)^{n+1}}{n+1} \frac{x^{n+1}}{(1+c)^{n+1}},$$

ce qui implique que

$$\sup_{x\in[0,1]} |\rho_n(x)| \leqslant \sup_{x\in[0,1]} \left|\frac{x^{n+1}}{n+1}\right| = \frac{1}{n+1}.$$

D'après le corollaire 2.1 p. 75, on en déduit que la suite d'applications $(\rho_n)_n$ converge uniformément sur $[0,1]$ vers l'application nulle, ce qui permet de conclure que la série d'applications $\sum_n f_n$ converge uniformément sur $[0,1]$.

Remarquons que pour cette série d'applications on a convergence uniforme sur $[0,1]$ mais convergence absolue sur $[0,1[$ uniquement.

2. La série d'applications $\sum_n f_n$ de terme général $f_n : x \in \mathbb{R} \mapsto x^n/n!$ converge simplement sur $\mathbb{R}$ et a pour somme la fonction exponentielle mais il n'y a pas convergence uniforme sur $\mathbb{R}$. Pour tous $n \in \mathbb{N}$ et $x \in \mathbb{R}$, on a

$$\rho_n(x) = \sum_{k=n+1}^{+\infty} f_k(x) = \sum_{k=n+1}^{+\infty} \frac{x^k}{k!} \geqslant \frac{x^{n+1}}{(n+1)!}$$

et comme l'application $x \in \mathbb{R} \mapsto x^{n+1}/(n+1)!$ n'est pas bornée sur $\mathbb{R}$, il en va de même de l'application ρ_n. La suite d'applications $(\rho_n)_n$ ne converge donc pas uniformément sur $\mathbb{R}$ vers l'application nulle, ce qui d'après le corollaire 3.1 implique que la série d'applications $\sum_n f_n$ ne converge pas uniformément sur $\mathbb{R}$.

3. Nous avons vu que la série $\sum_n x^n$ était simplement convergente sur $]-1,1[$ de somme l'application $\phi. : x \in]-1,1[\mapsto \frac{1}{1-x}$. Montrons que pour tout réel $r \in]0,1[$ fixé, elle converge uniformément sur $[-r,r]$ mais qu'elle ne converge pas uniformément sur $]-1,1[$. Pour tout $n \in \mathbb{N}$, on a

$$\sup_{x\in[-r,r]} \left|\rho_n(x)\right| = \sup_{x\in[-r,r]} \left|\sum_{k=n+1}^{=+\infty} x^k\right| = \sup_{x\in[-r,r]} \left|\frac{x^{n+1}}{1-x}\right| = \frac{r^{n+1}}{1-r}.$$

Pour $r \in]0,1[$, on a $\lim_{n\to+\infty} \frac{r^{n+1}}{1-r} = 0$, ce qui, d'après la proposition 2.2 p. 74, implique que la suite d'applications $(\rho_n)_n$ converge uniformément sur $[-r,r]$ vers l'application nulle. On en conclut que la série d'applications $\sum_n f_n$ converge uniformément sur $[-r,r]$.

Remarquons que comme la fonction $x \mapsto x^{n+1}/(1-x)$ n'étant pas bornée sur $]-1,1[$, la suite d'applications $(\rho_n)_n$ ne converge pas uniformément sur $]-1,1[$ vers l'application nulle et par conséquent la série d'applications $\sum_n f_n$ ne converge pas uniformément sur $]-1,1[$.

4. Nous avons vu que la série d'applications $\sum_n f_n$ de terme général $f_n : x \in]-\frac{\pi}{2},\frac{\pi}{2}[\mapsto \frac{1}{2^n}\tan\left(\frac{x}{2^n}\right)$ converge simplement sur $]-\frac{\pi}{2},\frac{\pi}{2}[$ et a pour somme l'application

$$\phi : x \in]-\tfrac{\pi}{2},\tfrac{\pi}{2}[\longmapsto \begin{cases} \tan(x) + \frac{1}{x} - \operatorname{cotan}(x) & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases}.$$

Étudions la convergence uniforme de cette série sur $]-\frac{\pi}{2},\frac{\pi}{2}[$. Pour $n \in \mathbb{N}$ et $x \in]-\frac{\pi}{2},0[\cup]0,\frac{\pi}{2}[$, on a d'après la relation (2) p. 117,

$$\rho_n(x) = \sum_{k=n+1}^{+\infty} f_n(x) = \phi(x) - \phi_n(x) = \frac{1}{x} - \frac{1}{2^n}\operatorname{cotan}\left(\frac{x}{2^n}\right).$$

L'application ρ_n est impaire et elle est continue et dérivable sur $]0,\frac{\pi}{2}[$, avec pour tout $x \in]0,\frac{\pi}{2}[$,

$$\rho_n'(x) = -\frac{1}{x^2} + \frac{1}{4^n \sin^2(x/2^n)} = \frac{(x/2^n)^2 - \sin^2(x/2^n)}{x^2 \sin^2(x/2^n)}.$$

Comme pour tout $y \in \mathbb{R}_+^*$, on a $\sin(y) < y$, l'application ρ_n' est strictement positive sur $]0,\frac{\pi}{2}[$ et par conséquent ρ_n est strictement croissante sur cet intervalle. Par ailleurs, $\rho_n(0) = 0$. On déduit de ce qui précède que :

$$u_n = \sup_{x\in]-\frac{\pi}{2},\frac{\pi}{2}[} |\rho_n(x)| = \sup_{x\in]0,\frac{\pi}{2}[} |\rho_n(x)| = \rho_n(\frac{\pi}{2}) = \frac{2}{\pi} - \frac{1}{2^n}\operatorname{cotan}\left(\frac{\pi}{2^{n+1}}\right).$$

Comme $\tan(u) \underset{0}{\sim} u$, on a $\frac{1}{2^n}\operatorname{cotan}\left(\pi/2^{n+1}\right) \underset{+\infty}{\sim} 2/\pi$ ce qui implique que la suite numérique $(u_n)_n$ converge vers 0. On en conclut que la suite d'applications $(\rho_n)_n$ converge uniformément sur $]-\frac{\pi}{2},\frac{\pi}{2}[$ vers l'application nulle et donc que la série $\sum_n f_n$ converge uniformément sur $]-\frac{\pi}{2},\frac{\pi}{2}[$.

Exercice 3.2 On considère la série de fonctions $\sum_n f_n$ de terme général

$$f_n : x \longmapsto \frac{x}{(1+x^2)^n}.$$

1 - Déterminer l'ensemble sur lequel cette série de fonctions converge simplement.
2 - En utilisant la proposition 3.5 (ou le corollaire 3.1), montrer que[5] la série de fonctions $\sum_n f_n$ converge uniformément sur $[a,+\infty[$ pour tout réel $a > 0$ mais ne converge pas uniformément sur $[0,+\infty[$.

Comme on a pu le constater sur les différents exemples traités, il n'est pas toujours aisé d'établir qu'une série d'applications converge uniformément. Nous allons donner dans la suite de ce chapitre des critères permettant d'établir la convergence uniforme ou la non-convergence uniforme d'une série d'applications de manière plus pratique.

3.1.2 Propriétés

Le théorème suivant relie entre elles les différentes notions de convergence pour une série d'applications.

[5] On pourra consulter l'exercice 3.6 p. 135 pour une autre méthode permettant d'établir ce résultat.

THÉORÈME 3.1

✗ *Si une série d'applications $\sum_n f_n$ de X dans $\mathbb{R}$ converge absolument sur X alors elle converge simplement sur X.*

✗ *Si une série d'applications $\sum_n f_n$ de X dans $\mathbb{R}$ converge uniformément sur X alors elle converge simplement sur X.*

Démonstration La première assertion résulte des propriétés des séries numériques. La série d'applications $\sum_n f_n$ de X dans $\mathbb{R}$ converge absolument sur X si pour tout $x \in X$, la série numérique $\sum_n f_n(x)$ converge absolument. Or, d'après la proposition 1.11 p. 29, si la série numérique $\sum_n f_n(x)$ converge absolument alors la série numérique $\sum_n f_n(x)$ converge. D'après la définition 3.2, cela signifie que la série d'applications $\sum_n f_n$ converge simplement sur X.

La seconde assertion est une conséquence du théorème 2.1 p. 78 qu'il suffit d'appliquer à la suite $(\phi_n)_n$ de terme général $\phi_n = \sum_{k=0}^{n} f_k$. □

Remarque Quel que soit le type de convergence (simple, absolue, uniforme) d'une série d'applications, la somme de la série est la même. Par ailleurs, si $\sum_n f_n$ est une série d'applications de X dans $\mathbb{R}$ et si Y est un sous-ensemble de X, on dit que la série $\sum_n f_n$ converge simplement (resp. absolument ou uniformément) sur Y si la série d'applications de Y dans $\mathbb{R}$ dont le terme général $f_n|_Y$ est la restriction à l'ensemble Y de l'application f_n converge simplement (resp. absolument ou uniformément). ◇

En appliquant la proposition 2.3 p. 79 à la suite $(\phi_n)_n$ de terme général $\phi_n = \sum_{k=0}^{n} f_k$, on obtient le résultat suivant.

PROPOSITION 3.6 *Soient $\sum_n f_n$ et $\sum_n g_n$ deux séries d'applications de X dans $\mathbb{R}$ convergeant simplement (resp. uniformément) sur X, de somme les applications ϕ et ψ de X dans $\mathbb{R}$.*

✗ *La série d'applications de terme général $f_n + g_n$ converge simplement (resp. uniformément) sur X et a pour somme $\phi + \psi$.*

✗ *Pour tout $\lambda \in \mathbb{R}$, la série d'applications de terme général $\lambda \cdot f_n$ converge simplement (resp. uniformément) sur X et a pour somme $\lambda \cdot \phi$.*

Par définition, la série d'applications $\sum_n f_n$ de X dans $\mathbb{R}$ converge uniformément sur X si et seulement si la suite d'applications de terme général $\phi_n = \sum_{k=0}^{n} f_k$ converge uniformément sur X. Le théorème 2.2 p. 82 appliqué à la suite $(\phi_n)_n$ fournit le résultat suivant appelé *critère de Cauchy uniforme pour les séries d'applications.*

THÉORÈME 3.2 (Critère de Cauchy uniforme pour les séries)
Une condition nécessaire et suffisante pour que la série d'applications $\sum_n f_n$ *de* X *dans* $\mathbb{R}$ *converge uniformément sur* X *est :*

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N \in \mathbb{N} \quad \forall (m,n) \in \mathbb{N}^2 \quad \boxed{\forall x \in X} \quad \left(n \geqslant m \geqslant N \implies \left|\sum_{k=m}^{n} f_k(x)\right| \leqslant \varepsilon\right). \tag{3}$$

Remarque Le critère de Cauchy uniforme pour les séries d'applications s'énonce également de la manière suivante :

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N \in \mathbb{N} \quad \forall (n,p) \in \mathbb{N}^2 \quad \forall x \in X \quad \left(n \geqslant N \implies \left|\sum_{k=n}^{n+p} f_k(x)\right| \leqslant \varepsilon\right),$$

ou encore de la manière suivante :

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N \in \mathbb{N} \quad \forall (m,n) \in \mathbb{N}^2 \quad \left(n \geqslant m \geqslant N \implies \sup_{x \in X} \left|\sum_{k=n}^{m} f_k(x)\right| \leqslant \varepsilon\right).$$

Selon les situations, il peut être avantageux d'utiliser l'un ou l'autre des trois énoncés. ◇

La proposition 1.2 p. 11 indique qu'une condition nécessaire pour qu'une série numérique $\sum_n u_n$ converge est que la suite $(u_n)_n$ converge vers 0. Ce résultat implique qu'une condition nécessaire pour qu'une série d'applications $\sum_n f_n$ de X dans $\mathbb{R}$ converge simplement sur X est que la suite d'application $(f_n)_n$ converge simplement sur X vers l'application nulle. Le critère de Cauchy uniforme pour les séries d'applications va nous permettre d'établir un résultat analogue pour la convergence uniforme.

PROPOSITION 3.7 *Si la série d'applications* $\sum_n f_n$ *de* X *dans* $\mathbb{R}$ *converge uniformément sur* X *alors* la suite d'applications $(f_n)_n$ *converge uniformément vers l'application nulle sur* X.

Démonstration On applique le critère de Cauchy uniforme en prenant $n = m$; on a $\sum_{k=m}^{n} f_k(x) = f_n(x)$ et puisque la série $\sum_n f_n$ converge uniformément sur X, il résulte du théorème 3.2 que

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N \in \mathbb{N} \quad \forall x \in X \quad \forall n \in \mathbb{N} \quad (n \geqslant N \implies |f_n(x)| \leqslant \varepsilon).$$

D'après la définition 2.2 p. 73, cette assertion indique que la suite d'applications $(f_n)_n$ converge uniformément vers l'application nulle sur X. □

Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU

Exemple Nous avons vu p. 117 que la série d'applications $\sum_n f_n$ de terme général $f_n : x \in \mathbb{R}_+^* \mapsto ne^{-nx}$ converge simplement sur $\mathbb{R}_+^*$. Pour $n \in \mathbb{N}$, on a

$$\sup_{x \in \mathbb{R}_+^*} |f_n(x) - 0| = \sup_{x \in \mathbb{R}_+^*} \left(ne^{-nx}\right) = n$$

et cette quantité ne tend pas vers 0 quand n tend vers $+\infty$. Cela signifie que la suite d'applications $(f_n)_n$ ne converge pas uniformément vers l'application nulle sur $\mathbb{R}_+^*$; d'après la proposition 3.7, on en conclut que la série d'applications $\sum_n f_n$ ne converge pas uniformément sur $\mathbb{R}_+^*$.

Dans la proposition 3.7, la condition de convergence uniforme de la suite d'applications $(f_n)_n$ de X dans $\mathbb{R}$ vers l'application nulle est une condition nécessaire pour que la série $\sum_n f_n$ converge uniformément sur X. Cette condition n'est toutefois pas suffisante. Par exemple, l'exercice 3.2 a permis de montrer que la série d'applications $\sum_n f_n$ de terme général $f_n : x \in \mathbb{R}^+ \mapsto x/(1+x^2)^n$ ne converge pas uniformément sur $\mathbb{R}^+$ alors que l'on a établi au cours de l'exercice 2.2 p. 78, que la suite d'applications $(f_n)_n$ converge uniformément sur $\mathbb{R}$, et donc sur $\mathbb{R}^+$, vers l'application nulle.

3.2 Critères de convergence uniforme

Lorsqu'on étudie une série d'applications, il est assez rare que l'on puisse déterminer l'application qui est la somme de la série. Établir la convergence uniforme d'une série d'applications en utilisant la définition 3.4, ou la proposition 3.5, est alors impossible. Nous allons donc à présent établir des critères qui permettent de montrer qu'une série d'applications converge uniformément sans qu'il soit nécessaire de déterminer au préalable la somme de cette série.

3.2.1 Convergence normale

DÉFINITION 3.5 *On dit que la série d'applications $\sum_n f_n$ de X dans $\mathbb{R}$ converge normalement sur X si*

- *à partir d'un certain rang N toutes les applications f_n sont bornées*[6] *sur X*
- *et la série numérique $\sum_{n \geqslant N} u_n$ de terme général $u_n = \sup_{x \in X} |f_n(x)|$ converge.*

Remarque On notera que la série $\sum_n u_n$ est une série numérique à termes positifs. La nature de cette série peut être établie en utilisant les critères de convergence pour les séries à termes positifs détaillés au chap. 1. ◇

[6] Sous forme d'assertion quantifiée cela se traduit de la manière suivante :
$\exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad (n \geqslant N \implies \exists M_n \in \mathbb{R}^+ \quad \forall x \in X \quad |f_n(x)| \leqslant M_n)$.

Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU

Exemples

1. Considérons la série d'applications $\sum_{n\geqslant 1} f_n$ dont le terme général est $f_n : x \in \mathbb{R} \mapsto \cos(nx)/n^2$. Comme la fonction cosinus est bornée par 1, pour tout $n \in \mathbb{N}^*$, on a

$$u_n = \sup_{x\in\mathbb{R}} \left|\frac{1}{n^2}\cos(nx)\right| = \frac{1}{n^2}\sup_{x\in\mathbb{R}} \left|\cos(nx)\right| \leqslant \frac{1}{n^2}.$$

Puisque la série $\sum_{n\geqslant 1} 1/n^2$ converge, il résulte des critères de comparaison pour les séries numériques à termes positifs[7] que la série $\sum_n u_n$ converge. On en conclut que la série d'applications $\sum_n f_n$ converge normalement sur $\mathbb{R}$.

Plus généralement, on peut remarquer que les séries d'applications $\sum_n f_n$ de terme général $f_n : x \in \mathbb{R} \mapsto \alpha_n \cos(nx)$ et $f_n : x \in \mathbb{R} \mapsto \alpha_n \sin(nx)$, où $(\alpha_n)_n$ est une suite numérique réelle, convergent normalement sur $\mathbb{R}$ si et seulement si la série $\sum_n \alpha_n$ est absolument convergente. Ces séries sont appelée séries trigonométriques.

2. Pour $r \in \mathbb{R}_+^*$, intéressons-nous à la série d'applications $\sum_n f_n$ dont le terme général est $f_n : x \in [-r, r] \mapsto x^n$. Pour tout $n \in \mathbb{N}$, on a

$$u_n = \sup_{x\in[-r,r]} |x^n| = r^n.$$

La série géométrique $\sum_n r^n$ de raison r converge si et seulement si $r < 1$. On en déduit que la série d'applications $\sum_n x^n$ converge normalement sur $[-r, r]$ si et seulement si $r \in]0, 1[$. Cette série ne converge donc pas normalement sur $[-2, 2]$, sur $[-1, 1]$ ou sur $\mathbb{R}$ par exemple.

3. Pour $r \in \mathbb{R}_+^*$ fixé, la série d'applications $\sum_n f_n$ dont le terme général est $f_n : x \in [-r, r] \mapsto x^n/n!$ converge normalement sur $[-r, r]$ car

$$u_n = \sup_{x\in[-r,r]} \left|\frac{x^n}{n!}\right| = \frac{r^n}{n!}$$

et la série numérique $\sum_n r^n/n!$ converge, voir p. 5. Par contre, comme $\lim_{n\to+\infty} f_n(x) = +\infty$, l'application f_n n'est pas bornée sur $\mathbb{R}$ et la série d'applications $\sum_n f_n$ ne converge donc pas normalement sur $\mathbb{R}$. Rappelons que l'on a établi que la série $\sum_n f_n$ converge simplement et absolument sur $\mathbb{R}$.

4. Nous avons vu que la série d'applications $\sum_n f_n$ de terme général $f_n : x \in [0, 1] \mapsto \frac{(-1)^n}{n} x^n$ converge simplement et uniformément sur $[0, 1]$. Toutefois, pour $n \in \mathbb{N}^*$ on a $u_n = \sup_{x\in[0,1]} |f_n(x)| = 1/n$ et comme la série harmonique $\sum_n 1/n$ diverge, il n'y a pas convergence normale de la série $\sum_n f_n$ sur $[0, 1]$ (pas plus que sur $]0, 1[$).

---

[7] Voir la proposition 1.2 p. 14. Signalons que l'on peut déterminer l'expression de u_n en remarquant que $u_n = \sup_{x\in\mathbb{R}} |f_n(x)| \geqslant f_n(2\pi/n) = 1/n^2$. Comme par ailleurs $u_n \leqslant 1/n^2$, on en déduit que que $u_n = \frac{1}{n^2}$.

Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU

Exercice 3.3 Montrer que la série d'applications $\sum_n f_n$ de terme général $f_n : x \in \mathbb{R}_+^* \mapsto ne^{-nx}$ converge normalement sur $[y, +\infty[$ pour tout réel $a > 0$ mais ne converge pas normalement sur $]0, +\infty[$.

PROPOSITION 3.8 *Soit $(f_n)_n$ une suite d'applications de X dans $\mathbb{R}$. Une condition nécessaire et suffisante pour que la série de fonctions $\sum_n f_n$ converge normalement sur X est qu'il existe une suite numérique $(v_n)_n$ telle que*

1. *pour tout $n \in \mathbb{N}$ on ait $\sup_{x \in X} |f_n(x)| \leqslant v_n$;*
2. *la série numérique $\sum_n v_n$ converge.*

Démonstration La série numérique $\sum_n u_n$ dont le terme général est $u_n = \sup_{x \in X} |f_n(x)|$ est une série à termes positifs qui vérifie $u_n \leqslant v_n$ pour tout $n \in \mathbb{N}$. D'après le théorème 1.2 p. 14, la convergence de la série numérique $\sum_n v_n$ entraîne la convergence de la série numérique $\sum_n u_n$. □

Exemple Considérons la série d'applications dont le terme général est $f_n : x \in]1, +\infty[\mapsto (n^2x + (-1)^n)^{-1}$. Pour tout $n \in \mathbb{N}$ et pour tout $x \in]1, +\infty[$ on a : $n^2x + (-1)^n \geqslant n^2x - 1 \geqslant n^2 - 1$. On en déduit que pour tout entier $n \geqslant 2$

$$\sup_{x \in]1, +\infty[} |f_n(x)| = \sup_{x \in]1, +\infty[} \left| \frac{1}{n^2x + (-1)^n} \right| \leqslant \frac{1}{n^2 - 1}.$$

Comme $\frac{1}{n^2-1} \underset{+\infty}{\sim} \frac{1}{n^2}$ et que la série de Riemann $\sum_n \frac{1}{n^2}$ converge, la série $\sum_n \frac{1}{n^2-1}$ converge. On en conclut que la série $\sum_n f_n$ converge normalement sur $]1, +\infty[$.

THÉORÈME 3.3 *Si la série d'applications $\sum_n f_n$ de X dans $\mathbb{R}$ converge normalement sur X alors elle converge uniformément sur X.*

Démonstration Soit $(f_n)_n$ une suite d'applications de X dans $\mathbb{R}$ telle que la série $\sum_n f_n$ converge normalement sur X. Pour montrer que la série $\sum_n f_n$ converge uniformément sur X, nous allons montrer que $\sum_n f_n$ vérifie le critère de Cauchy uniforme donné par le théorème 3.2.

Soient m et n deux entiers naturels tels que $n \geqslant m$. On a

$$0 \leqslant \sup_{x \in X} \left| \sum_{k=m}^{n} f_k(x) \right| \leqslant \sup_{x \in X} \left(\sum_{k=m}^{n} |f_k(x)| \right) \leqslant \sum_{k=m}^{n} \sup_{x \in X} |f_k(x)|. \qquad (4)$$

Par hypothèse, la série d'applications $\sum_n f_n$ converge normalement sur X, ce qui signifie que la série numérique $\sum_n u_n$ de terme général $u_n = \sup_{x \in X} |f_n(x)|$

Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU

converge. La série $\sum_n u_n$ vérifie donc le critère de Cauchy pour les séries numériques, voir le théorème 1.1 p. 9, *i.e.*

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N \in \mathbb{N} \quad \forall (m,n) \in \mathbb{N}^2 \quad \left(n \geqslant m \geqslant N \implies \left| \sum_{k=m}^{n} u_k \right| \leqslant \varepsilon \right).$$

On déduit de la relation (4) que

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N \in \mathbb{N} \quad \forall (m,n) \in \mathbb{N}^2 \quad \left(n \geqslant m \geqslant N \implies \sup_{x \in X} \left| \sum_{k=m}^{n} f_k(x) \right| \leqslant \varepsilon \right).$$

La série d'applications $\sum_n f_n$ vérifie donc le critère de Cauchy uniforme et par conséquent elle converge uniformément sur X. □

Remarque Il est souvent beaucoup plus simple d'étudier la convergence normale d'une série d'applications que d'étudier sa convergence uniforme. D'un point de vue pratique, lorsque l'on cherche à montrer qu'une série converge uniformément on regarde dans un premier temps si cette série converge normalement. Si c'est le cas on peut conclure à la convergence uniforme. On sera vigilant au fait que si la série ne converge pas normalement, elle peut toutefois converger uniformément (c'est le cas comme nous l'avons vu de la série de fonctions $\sum_n (-1)^n x^n/n$ qui converge uniformément sur $[0,1]$ mais qui ne converge pas normalement sur cet intervalle). Pour établir la convergence uniforme dans le cas où il n'y a pas convergence normale, il faut avoir recours à d'autres critères, par exemple revenir à la définition ou utiliser les résultats qui seront énoncés dans la suite de ce chapitre. ◇

On a le résultat suivant qui relie la convergence normale à la convergence absolue.

THÉORÈME 3.4 *Si la série d'applications $\sum_n f_n$ de X dans $\mathbb{R}$ converge normalement sur X alors elle converge absolument sur X.*

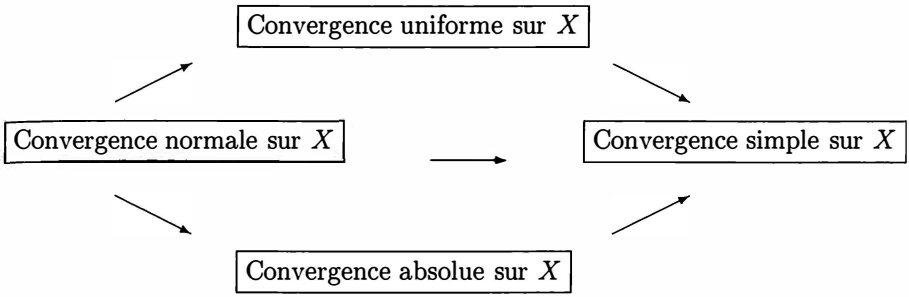
Démonstration Soit $(f_n)_n$ une suite d'applications de X dans $\mathbb{R}$ telle que la série $\sum_n f_n$ converge normalement sur X, *i.e.* telle que la série numérique de terme général $\sup_{y \in X} |f_n(y)|$ converge. Pour tous $n \in \mathbb{N}$ et $x \in X$ on a

$$0 \leqslant |f_n(x)| \leqslant \sup_{y \in X} |f_n(y)|.$$

Pour tout $x \in X$, la série $\sum_n |f_n(x)|$ est donc une série numérique à termes positifs dont le terme général est majoré par celui d'une série convergente. Le théorème 1.2 p. 14 permet d'en déduire que la série $\sum_n |f_n(x)|$ converge et par conséquent que la série $\sum_n f_n(x)$ est absolument convergente. □

Le schéma suivant indique les liens existants entre les différents types de convergence pour une série de fonctions.

Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU



On notera qu'il n'y a pas de lien entre la convergence uniforme et la convergence absolue. Nous avons par exemple vu que la série de fonctions $\sum_n f_n$ de terme général $f_n : x \mapsto \frac{(-1)^n}{n} x^n$ convergeait uniformément sur $[0,1]$ mais ne convergeait pas absolument sur $[0,1]$ (mais seulement sur $[0,1[$) alors que la série de fonctions $\sum_n f_n$ de terme général $f_n : x \mapsto \frac{1}{n!} x^n$ convergeait absolument sur $\mathbb{R}$ mais ne convergeait pas uniformément sur $\mathbb{R}$ (mais seulement sur $[-r,r]$ pour tout $r \in \mathbb{R}_+^*$).

Exercice 3.4 Soit $(f_n)_n$ une suite d'applications dérivables de $[a,b]$ dans $\mathbb{R}$. On suppose que la série numérique $\sum_n f_n(a)$ converge absolument et que la série de fonctions $\sum_n f_n'$ converge normalement sur $[a,b]$. En utilisant le théorème des accroissements finis[8], montrer que la série de fonctions $\sum_n f_n$ converge normalement sur $[a,b]$.

La fonction ζ de Riemann

Intéressons-nous à la fonction ζ (zeta) de Riemann qui est la somme de la série de fonctions de terme général $f_n : x \mapsto 1/n^x$.

Puisque la série de Riemann $\sum_n 1/n^\alpha$ converge si et seulement si $\alpha > 1$, il y a convergence simple (et absolue car f_n est à valeurs positives) de la série de fonctions $\sum_n 1/n^x$ sur $]1,+\infty[$ et divergence sur $]-\infty,1]$. L'ensemble de définition de la fonction ζ de Riemann est $]1,+\infty[$.

La série d'applications $\sum_n f_n$ ne converge pas uniformément sur $]1,+\infty[$. Pour l'établir, montrons que la série de fonctions $\sum_n f_n$ ne satisfait pas au critère de Cauchy uniforme donné par le théorème 3.2. Montrons que :

$$\forall N \in \mathbb{N} \quad \exists (m,n) \in \mathbb{N}^2 \quad \Big(n \geqslant m \geqslant N \text{ et } \sup_{x \in]1,+\infty[} \Big| \sum_{k=m}^{n} \frac{1}{k^x} \Big| > \frac{1}{2} \Big). \qquad (5)$$

Pour $(m,n) \in \mathbb{N}^2$ avec $n \geqslant m$, l'application $\varphi : x \in]1,+\infty[\mapsto \sum_{k=m}^{n} 1/k^x$ est positive et décroissante (c'est une somme finie d'applications positives et

[8] Voir le théorème 16.2 p. 731 du *Cours de première année*.

Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU

décroissantes) donc

$$\sup_{x\in]1,+\infty[}\left|\sum_{k=m}^{n}\frac{1}{k^x}\right| = \sup_{x\in]1,+\infty[}\varphi(x) = \varphi(1) = \sum_{k=m}^{n}\frac{1}{k}.$$

Considérons pour $N \in \mathbb{N}$ fixé, les entiers $m = 2^N$ et $n = 2^{N+1} - 1$. On a,

$$\sup_{x\in]1,+\infty[}\left|\sum_{k=m}^{n}\frac{1}{k^x}\right| = \sum_{k=2^N}^{2^{N+1}-1}\frac{1}{k} > 2^N\frac{1}{2^{N+1}} = \frac{1}{2},$$

car la somme contient 2^N termes tous strictement plus grands que $1/2^{N+1}$. La relation (5) est démontrée.

Toutefois, la série de fonctions $\sum_n f_n$ converge normalement sur $[a, +\infty[$ pour tout $a > 1$ puisque

$$\sup_{x\in[a,+\infty[}\left|\frac{1}{n^x}\right| = \frac{1}{n^a}$$

et la série de Riemann $\sum_{n\geqslant 1} 1/n^a$ converge pour tout $a > 1$. La représentation graphique de la fonction ζ de Riemann est donnée p. 141.

On peut remarquer avec cet exemple que la convergence normale d'une série de fonction $\sum_n f_n$ sur tous les intervalles de la forme $[a, +\infty[$ avec $a > 1$ n'implique pas la convergence normale de la série sur $]1, +\infty[$.

3.2.2 Convergence uniforme des séries alternées

Le théorème suivant donne des conditions suffisantes pour pouvoir affirmer qu'une série d'applications qui est alternée, c'est-à-dire une série d'applications $\sum_n f_n$ de terme général f_n de la forme $(-1)^n g_n$ avec g_n une application positive, converge uniformément. Ce résultat est connu sous le nom de *critère spécial de convergence uniforme des séries alternées.*

THÉORÈME 3.5 (Convergence uniforme des séries alternées)
Soit $(f_n)_n$ une suite d'applications d'un sous-ensemble X de $\mathbb{R}$ dans $\mathbb{R}$ dont le terme général est de la forme

$$f_n(x) = (-1)^n\, g_n(x)$$

où $(g_n)_n$ est une suite d'applications de X dans $\mathbb{R}$. Si,

1. *pour tout $x \in X$ fixé, la suite numérique $(g_n(x))_n$ est une suite à termes positifs, décroissante ;*
2. *la suite d'applications $(g_n)_n$ converge uniformément sur X vers l'application nulle ;*

alors la série d'applications $\sum_n f_n$ converge uniformément sur X.

Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU

Démonstration Pour $x \in X$ fixé, considérons la suite $(U_n)_n$ des sommes partielles associée à la série alternée $\sum_n u_n$ de terme général $u_n = (-1)^n g_n(x)$. Nous avons montré au cours de la démonstration du théorème 1.3 p. 42 que la suite extraite de $(U_n)_n$ correspondant aux termes d'indices pairs et la suite extraite de $(U_n)_n$ correspondant aux termes d'indices impairs étaient deux suites adjacentes convergeant vers la somme de la série. Il résulte des propriétés des suites adjacentes que pour tout $p \in \mathbb{N}$ on a

$$U_{2p+1} \leqslant U \leqslant U_{2p},$$

où $U = \sum_{n=0}^{+\infty} u_n = \sum_{n=0}^{+\infty} (-1)^n g_n(x)$. Ainsi, pour tout $n \in \mathbb{N}$, on a ou bien $U_{n+1} \leqslant U \leqslant U_n$ ou bien $U_{n+1} \geqslant U \geqslant U_n$, selon que n est pair ou impair. Dans les deux cas, on a

$$\Big| \sum_{k=n+1}^{+\infty} (-1)^k g_k(x) \Big| = |U - U_n| \leqslant |U_{n+1} - U_n| = |g_{n+1}(x)| = g_{n+1}(x),$$

ce qui implique que

$$\sup_{x \in X} \Big| \sum_{k=n+1}^{+\infty} (-1)^k g_k(x) \Big| \leqslant \sup_{x \in X} |g_{n+1}(x)|.$$

Comme la suite d'applications $(g_n)_n$ converge uniformément sur X vers l'application nulle, la suite d'applications $(\rho_n)_n$ de terme général

$$\rho_n : x \in X \longmapsto \sum_{k=n+1}^{+\infty} (-1)^k g_k(x)$$

converge uniformément sur X vers l'application nulle. D'après le corollaire 3.1, on en conclut que la série d'applications $\sum_n (-1)^n g_n$ converge uniformément sur X. □

Exemples

1. Montrons, en utilisant le théorème 3.5, que la série d'applications $\sum_{n \geqslant 1} f_n$ de terme général $f_n : x \in [0,1] \mapsto (-1)^n x^n / n$ converge uniformément sur $[0,1]$. Pour $n \in \mathbb{N}^*$ posons $g_n : x \in [0,1] \mapsto x^n / n$. Cette application est positive et, pour $x \in]0,1]$ fixé, on a

$$\frac{g_{n+1}(x)}{g_n(x)} = x \, \frac{n}{n+1} < x \leqslant 1.$$

On en déduit que pour tout $x \in]0,1]$, la suite $(g_n(x))_n$ est strictement décroissante. Comme par ailleurs la suite $(g_n(0))_n$ est la suite nulle, la suite numérique réelle $(g_n(x))_n$ est une suite décroissante pour tout $x \in [0,1]$. De plus, $\sup_{x \in [0,1]} |g_n(x)| = \frac{1}{n}$ admet pour limite 0 quand n tend vers $+\infty$, ce qui indique que la suite d'applications $(g_n)_n$ converge uniformément sur $[0,1]$ vers l'application nulle. Le critère spécial de convergence uniforme des séries

Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU

alternées permet de conclure à la converge uniforme sur $[0,1]$ de la série d'applications $\sum_{n\geqslant 1} f_n$.

2. Montrons, en utilisant le théorème 3.5, que la série d'applications $\sum_n f_n$ de terme général $f_n : x \in [0,+\infty[\mapsto \frac{(-1)^n}{\sqrt{n}}\, \mathrm{e}^{-nx}$, converge uniformément sur $[0,+\infty[$. On a $f_n = (-1)^n g_n$ où $g_n : x \in [0,+\infty[\mapsto \frac{1}{\sqrt{n}}\mathrm{e}^{-nx}$. Vérifions que la suite d'applications $(g_n)_n$ satisfait aux deux hypothèses du théorème.
– Pour tout $x \in \mathbb{R}$, le réel $g_n(x)$ est strictement positif et

$$\frac{g_{n+1}(x)}{g_n(x)} = \frac{\sqrt{n}}{\sqrt{n+1}}\, \mathrm{e}^{-x} < \mathrm{e}^{-x}.$$

Si x est positif alors $\mathrm{e}^{-x} \leqslant 1$ et par conséquent la suite numérique $(g_n(x))_n$ est décroissante. On en déduit que l'intervalle $[0,+\infty[$ est le plus grand ensemble sur lequel on peut espérer montrer la convergence uniforme de la série en ayant recours au critère spécial.
– Pour tout entier $n \geqslant 1$, on a

$$\sup_{x\in[0,+\infty[} |g_n(x)| = \frac{1}{\sqrt{n}} \sup_{x\in[0,+\infty[} \left|\mathrm{e}^{-nx}\right|.$$

Pour $n \in \mathbb{N}^*$ fixé, la fonction $x \mapsto \mathrm{e}^{-nx}$ étant décroissante sur $[0,+\infty[$, elle atteint son maximum qui vaut 1 en 0. On a donc $\sup_{x\in[0,+\infty[} |g_n(x)| = \frac{1}{\sqrt{n}}$. On déduit de la proposition 2.2 p. 74 que la suite d'applications $(g_n)_n$ converge uniformément sur $[0,+\infty[$ vers l'application nulle.

Le critère spécial de convergence uniforme des séries alternées permet de conclure que la série $\sum_{n\geqslant 1} \frac{(-1)^n}{\sqrt{n}}\, \mathrm{e}^{-nx}$ converge uniformément sur $[0,+\infty[$.

On remarquera que cette série de fonctions ne converge pas normalement sur $[0,+\infty[$ car

$$\sup_{x\in[0,+\infty[} |f_n(x)| = \sup_{x\in[0,+\infty[} \left|\frac{(-1)^n}{\sqrt{n}}\, \mathrm{e}^{-nx}\right| = \frac{1}{\sqrt{n}}$$

et la série de Riemann $\sum_{n\geqslant 1} \frac{1}{\sqrt{n}}$ diverge.

Exercice 3.5 On considère la série de fonctions $\sum_n f_n$ de terme général

$$f_n : x \longmapsto \frac{(-1)^n}{x^2+n}.$$

1 - Montrer que la série $\sum_n f_n$ converge simplement sur $\mathbb{R}$.
2 - Montrer que la série $\sum_n f_n$ converge uniformément sur $\mathbb{R}$ mais qu'il n'existe aucun intervalle de $\mathbb{R}$ où la convergence est normale.

Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU

3.2.3 Règle d'Abel pour la convergence uniforme

Nous avons établi au chapitre 1, théorème 1.4 p. 46, un résultat (la règle d'Abel) qui assure la semi-convergence d'une certaine classe de séries numériques non absolument convergentes. Nous allons donner une règle qui permet, de manière très similaire, d'établir la convergence uniforme de certaines séries de fonctions lorsque la convergence normale ne peut être établie. Ce théorème est admis.

THÉORÈME 3.6 (Règle d'Abel pour la convergence uniforme)
Soit $(f_n)_n$ une suite d'applications de X dans $\mathbb{R}$ dont le terme général est de la forme

$$f_n(x) = g_n(x) \times h_n(x)$$

où $(g_n)_n$ et $(h_n)_n$ sont deux suites d'applications de X dans $\mathbb{R}$. Si,

1. *pour tout $x \in X$ fixé, la suite numérique $(g_n(x))_n$ est une suite décroissante ;*
2. *la suite d'applications $(g_n)_n$ converge uniformément sur X vers l'application nulle ;*
3. *il existe un réel positif M tel que : $\forall x \in X \quad \forall n \in \mathbb{N} \quad |\sum_{k=0}^{n} h_k(x)| \leqslant M$;*

alors la série d'applications $\sum_n f_n$ converge uniformément sur X.

Remarques

1. Le théorème reste valable lorsque les conditions ne sont réalisées qu'à partir d'un certain rang n_0, où n_0 est un entier fixé indépendant de x.

2. Compte tenu des deux premières conditions, pour tout $x \in X$ fixé, la suite numérique $((g_n(x))_n$ est une suite à termes positifs, décroissante et convergeant vers 0. ◇

Application aux séries trigonométriques

Considérons les séries de fonctions $\sum_n \alpha_n \cos(nx)$ et $\sum_n \alpha_n \sin(nx)$ où $(\alpha_n)_n$ est une suite de réels positifs, qui est décroissante et qui converge vers 0.

Pour tout $x \in \mathbb{R} \setminus 2\pi\mathbb{Z}$ et tout $n \in \mathbb{N}$ on a

$$\sum_{k=0}^{n} \cos(kx) + \mathrm{i} \sum_{k=0}^{n} \sin(kx) = \sum_{k=0}^{n} \mathrm{e}^{\mathrm{i}kx} = \frac{\mathrm{e}^{\mathrm{i}(n+1)x} - 1}{\mathrm{e}^{\mathrm{i}x} - 1}$$

(8) Le théorème 3.6 reste vrai dans le cas de séries de fonctions de la variable complexe en remplaçant cette condition par la condition suivante : pour tout $x \in X$ fixé, la suite numérique réelle $(|g_n(x)|)_n$ est une suite décroissante.

Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU

donc[9]

$$\left|\sum_{k=0}^{n}\cos(kx)\right| \leqslant \left|\frac{e^{i(n+1)x}-1}{e^{ix}-1}\right| \leqslant \frac{2}{|e^{ix}-1|},$$

et

$$\left|\sum_{k=0}^{n}\sin(kx)\right| \leqslant \left|\frac{e^{i(n+1)x}-1}{e^{ix}-1}\right| \leqslant \frac{2}{|e^{ix}-1|}.$$

En utilisant le fait que $|ie^{i\frac{x}{2}}| = 1$ et les relations d'Euler[10], on obtient :

$$\frac{2}{|e^{ix}-1|} = \frac{|2ie^{i\frac{x}{2}}|}{|e^{ix}-1|} = \left|\frac{2i}{e^{i\frac{x}{2}}-e^{-i\frac{x}{2}}}\right| = \frac{1}{|\sin(\frac{x}{2})|}.$$

On en déduit que pour tout $x \in [2k\pi+\eta, 2(k+1)\pi-\eta]$

$$\left|\sum_{k=0}^{n}\cos(kx)\right| \leqslant \frac{1}{\sin(\frac{\eta}{2})} \quad \text{et} \quad \left|\sum_{k=0}^{n}\sin(kx)\right| \leqslant \frac{1}{\sin(\frac{\eta}{2})}.$$

La condition 3 du théorème 3.6 est donc satisfaite en prenant $M = \sin(\frac{\eta}{2})^{-1}$. On en conclut que pour tout $\eta \in]0,\pi[$, les séries de fonctions $\sum_n \alpha_n \cos(nx)$ et $\sum_n \alpha_n \sin(nx)$ convergent uniformément sur l'intervalle $[2k\pi+\eta, 2(k+1)\pi-\eta]$ (où $k \in \mathbb{Z}$).

3.3 Propriétés de la somme d'une série de fonctions

Certaines propriétés que possèdent les termes d'une série d'applications sont conservées par l'application qui en est la somme. Dans certains cas, ces propriétés sont transmises à la somme de la série sans condition, dans d'autres cas certaines hypothèses sur le type de convergence sont requises.

D'une manière générale, les résultats que nous énoncerons pour les séries d'applications se déduisent des résultats que nous avons établis au chap. 2 pour les suites d'applications par le biais de l'argument suivant : la série d'applications $\sum_n f_n$ converge sur X et a pour somme l'application ϕ si et seulement si la suite d'applications $(\phi_n)_n$ de terme général $\phi_n : x \in X \mapsto \sum_{k=0}^{n} f_k(x)$ converge sur X vers ϕ

Commençons par la propriété de continuité de la somme d'une série de fonctions.

[9] Pour tout $z \in \mathbb{C}$ on a $|\mathcal{R}e(z)| \leqslant |z|$ et $|\mathcal{I}m(z)| \leqslant |z|$.

[10] Voir la proposition 4.6 p. 138 du *Cours de première année*.

Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU

THÉORÈME 3.7 *Soient I un intervalle ouvert de $\mathbb{R}$ et $\sum_n f_n$ une série d'applications de I dans $\mathbb{R}$ qui converge uniformément sur I.*

✗ *Si pour tout $n \in \mathbb{N}$ l'application f_n est continue en $x_0 \in I$ alors la somme de la série $\phi : x \in I \mapsto \sum_{n=0}^{+\infty} f_n(x)$ est une application continue en x_0.*

✗ *Si pour tout $n \in \mathbb{N}$ l'application f_n est continue sur I alors la somme ϕ de la série est une application continue sur I.*

Démonstration La série d'applications $\sum_n f_n$ converge uniformément sur I si et seulement si la suite d'applications $(\phi_n)_n$ dont le terme général est $\phi_n : x \in I \mapsto \sum_{k=0}^{n} f_k(x)$ converge uniformément sur I. Si pour tout $k \in \mathbb{N}$, l'application f_k est continue en x_0 (resp. sur I) alors pour tout $n \in \mathbb{N}$ l'application ϕ_n est continue en x_0 (resp. sur I) en tant que somme finie d'applications continues. D'après le théorème 2.3 p. 86, la suite $(\phi_n)_n$ converge vers une application continue en x_0 (resp. sur I) et d'après la définition 3.4, cette application est la somme de la série $\sum_n f_n$. Le théorème est démontré. □

Remarques

1. Le théorème 3.7 s'étend sans difficulté au cas de la continuité à gauche ou à droite[(11)] permettant ainsi de considérer un réel x_0 adhérent à I.

2. La condition de convergence uniforme de la série d'applications $\sum_n f_n$ est une condition suffisante pour que l'application somme soit continue mais elle n'est pas une condition nécessaire. Par exemple la série d'applications $\sum_n f_n$ de terme général $f_n : x \in \mathbb{R} \mapsto x^n/n!$ est une série d'applications continues sur $\mathbb{R}$ qui converge simplement sur $\mathbb{R}$ vers l'application exponentielle qui est elle aussi une application continue sur $\mathbb{R}$ mais il n'y a pas convergence uniforme de la série d'applications sur $\mathbb{R}$ comme nous l'avons vu p. 120. ◇

Exercice 3.6

1 - Montrer que la série $\sum_n f_n$ de terme général $f_n : x \in \mathbb{R} \mapsto \frac{x}{(1+x^2)^n}$ converge simplement sur $\mathbb{R}$ et déterminer sa somme.

2 - En utilisant le théorème 3.7, établir que[(12)] la série $\sum_n f_n$ ne converge pas uniformément sur $\mathbb{R}$.

COROLLAIRE 3.2 *Soient I un intervalle ouvert, $\sum_n f_n$ une série d'applications de I dans $\mathbb{R}$ qui converge uniformément sur I et x_0 un point adhérent à I. Si pour tout $n \in \mathbb{N}$ l'application f_n admet pour limite en x_0 le réel u_n, alors la série numérique $\sum_n u_n$ converge et sa somme vérifie*

$$\sum_{n=0}^{+\infty} u_n = \lim_{x \to x_0} \sum_{n=0}^{+\infty} f_n(x).$$

(11) Voir la définition 13.5 p. 590 du *Cours de première année*.

(12) Cette propriété a aussi été établie dans l'exercice 3.2 ; on comparera les deux méthodes.

Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU

Remarques

1. Le corollaire 3.2 donne une condition suffisante pour pouvoir permuter une limite et une somme infinie puisque l'égalité s'écrit encore :

$$\sum_{n=0}^{+\infty} \lim_{x \to x_0} f_n(x) = \lim_{x \to x_0} \sum_{n=0}^{+\infty} f_n(x).$$

2. Le résultat reste valable dans le cas où x_0 est remplacé par $+\infty$ où $-\infty$, voir la proposition 2.7 p. 89. ◇

Exemple Considérons la série d'applications $\sum_n f_n$ de terme général $f_n : x \in [1, +\infty[\mapsto \frac{1}{n} \mathrm{e}^{-nx}$. Cette série converge normalement sur $[1, +\infty[$ car $\sup_{x \in [1,+\infty[} |f_n(x)| = \frac{1}{n\mathrm{e}^n}$ et le critère de D'Alembert indique que la série numérique $\sum_{n \geqslant 1} \frac{1}{n\mathrm{e}^n}$ est une série numérique convergente.

Pour $n \in \mathbb{N}^*$, notons $u_n = \lim_{x \to +\infty} f_n(x)$; la série $\sum_n u_n$ est la série nulle. On déduit du corollaire 3.2 et de la remarque précédente que la somme ϕ de la série $\sum_n f_n$ a pour limite 0 en $+\infty$.

THÉORÈME 3.8 *Soit $\sum_n f_n$ une série d'applications de $[a, b]$ dans $\mathbb{R}$ qui sont Riemann-intégrables sur $[a, b]$. Si la série $\sum_n f_n$ converge uniformément sur $[a, b]$ alors la somme $\phi : x \in [a, b] \mapsto \sum_{n=0}^{+\infty} f_n(x)$ de la série est une application Riemann-intégrable sur $[a, b]$ et on a*

$$\int_a^b \phi(x)\, \mathrm{d}x = \sum_{n=0}^{+\infty} \int_a^b f_n(x)\, \mathrm{d}x.$$

Démonstration La série d'applications $\sum_n f_n$ converge uniformément sur $[a, b]$ si et seulement si la suite d'applications $(\phi_n)_n$ de terme général $\phi_n : x \in [a, b] \mapsto \sum_{k=0}^{n} f_k(x)$ converge uniformément sur $[a, b]$. Si pour tout $k \in \mathbb{N}$, l'application f_k est Riemann-intégrable sur $[a, b]$ alors pour tout $n \in \mathbb{N}$ l'application ϕ_n est Riemann-intégrable sur $[a, b]$ en tant que somme finie d'applications Riemann-intégrable sur $[a, b]$. D'après le théorème 2.4 p. 90 la suite $(\phi_n)_n$ converge vers une application ϕ qui est Riemann-intégrable sur $[a, b]$ et d'après la définition 3.4, cette application ϕ est la somme de la série $\sum_n f_n$. Le théorème est démontré. □

Exemple Nous avons établi pp. 117 et 125 que la série d'applications $\sum_n f_n$ de terme général $f_n : x \in \mathbb{R}_+^* \mapsto n\mathrm{e}^{-nx}$ convergeait simplement mais pas uniformément sur $\mathbb{R}_+^*$. Toutefois, elle converge normalement, et par conséquent uniformément, sur $[a, +\infty[$ pour tout $a \in \mathbb{R}_+^*$.

Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU

Soient $a \in \mathbb{R}_+^*$ et $x \in \mathbb{R}$ avec $x > a$; pour tout $n \in \mathbb{N}$, l'application f_n est continue donc Riemann-intégrable sur $[a, x]$ et la série $\sum_n f_n$ converge uniformément sur $[a, x]$. On déduit du théorème 3.8 que la somme ϕ de la série $\sum_n f_n$ vérifie

$$\int_a^x \phi(t)\ \mathrm{d}t = \sum_{n=0}^{+\infty} \int_a^x f_n(t)\ \mathrm{d}t = \sum_{n=0}^{+\infty} \int_a^x n\mathrm{e}^{-nt}\ \mathrm{d}t = \sum_{n=0}^{+\infty} \left(\mathrm{e}^{-nx} - \mathrm{e}^{-na}\right).$$

Comme la somme de la série géométrique $\sum_n \mathrm{e}^{-nx}$ vaut $\frac{1}{1-\mathrm{e}^{-x}}$, on obtient :

$$\int_a^x \phi(t)\ \mathrm{d}t = \frac{1}{1-\mathrm{e}^{-x}} - \frac{1}{1-\mathrm{e}^{-a}}.$$

Cette expression donne l'intégrale indéfinie[13] associée à ϕ. On obtient par dérivation

$$\phi(x) = \left(\frac{1}{1-\mathrm{e}^{-x}}\right)' = \frac{\mathrm{e}^{-x}}{(1-\mathrm{e}^{-x})^2} = \frac{\mathrm{e}^{x}}{(\mathrm{e}^{x}-1)^2}.$$

Ainsi, la série d'applications $\sum_n f_n$ de terme général $f_n : x \in \mathbb{R}_+^* \mapsto n\mathrm{e}^{-nx}$ converge simplement sur $\mathbb{R}_+^*$ et a pour somme $\phi : x \in \mathbb{R}_+^* \mapsto \frac{\mathrm{e}^x}{(\mathrm{e}^x-1)^2}$.

Application : calcul approché d'une intégrale

Le théorème 3.8 fournit un procédé pour calculer une valeur approchée de certaines intégrales. Considérons par exemple la série d'applications de terme général $f_n : x \in [0,1] \mapsto \frac{(-1)^n x^{2n}}{2^n n!}$. Elle converge normalement et par conséquent uniformément sur $[0,1]$ puisque $\sup_{x\in[0,1]} |f_n(x)| = \frac{1}{2^n n!}$ et que le critère de D'Alembert indique que la série numérique $\sum_n \frac{1}{2^n n!}$ converge. Pour tout $x \in [0,1]$, on a

$$\sum_{n=0}^{+\infty} \frac{(-1)^n\ x^{2n}}{2^n\ n!} = \sum_{n=0}^{+\infty} \frac{1}{n!} \left(-\frac{x^2}{2}\right)^n = \mathrm{e}^{-x^2/2}$$

car $\mathrm{e}^u = \sum_{n=0}^{+\infty} \frac{u^n}{n!}$ pour tout réel u, voir p. 5. La somme de la série $\sum_n f_n$ est donc l'application $\phi : x \in [0,1] \mapsto \mathrm{e}^{-x^2/2}$.

Cette application est omniprésente en calcul des probabilités ; il s'agit (à la constante $\sqrt{2\pi}$ près) de la fonction de densité de la loi normale centrée réduite (ou loi gaussienne). La primitive de l'application ϕ n'étant pas exprimable à l'aide des fonctions usuelles, nous allons voir comment obtenir, à l'aide du théorème 3.8, une valeur approchée de l'intégrale $I = \int_0^1 \mathrm{e}^{-x^2/2}\ \mathrm{d}x$.

Puisque pour tout $n \in \mathbb{N}$, l'application f_n est continue sur $[0,1]$, elle est

[13] Voir le *Cours de première année* p. 847.

Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU

Riemann-intégrable sur $[0,1]$ et on déduit du théorème 3.8 que

$$\int_0^1 \mathrm{e}^{-x^2/2}\,\mathrm{d}x = \sum_{n=0}^{+\infty}\int_0^1 \frac{(-1)^n\,x^{2n}}{2^n\,n!}\,\mathrm{d}x = \sum_{n=0}^{+\infty}\frac{(-1)^n}{2^n\,n!}\int_0^1 x^{2n}\,\mathrm{d}x$$
$$= \sum_{n=0}^{+\infty}(-1)^n\,\alpha_n \quad \text{où} \quad \alpha_n = \frac{1}{2^n\,(2n+1)\,n!}.$$

Une valeur approchée de l'intégrale I est $I_N = \sum_{n=0}^{N}(-1)^n\,\alpha_n$ où N est un entier que l'on choisira d'autant plus grand que l'on souhaite une valeur approchée précise. Puisque la suite $(\alpha_n)_n$ est suite de réels positifs, décroissante et convergeant vers 0, il résulte de la proposition 1.15 p. 44 que le reste d'ordre N de la série alternée $\sum_n(-1)^n\,\alpha_n$, qui donne l'erreur d'approximation $I - I_N$, est tel que $|R_N| \leqslant \alpha_{N+1}$.

On peut tester cette méthode de calcul d'une valeur approchée de $\int_0^1 \mathrm{e}^{-x^2/2}\,\mathrm{d}x$ avec une erreur inférieure à 10^{-16} à l'aide de MAPLE.

```
> alpha:= n-> 1/(2^n*(2*n+1)*n!):
> N:=ceil(fsolve(alpha(n+1)=1E-16,n,1..infinity));
                                N := 13
> u:=n-> (-1)^n*alpha(n):
> evalf(sum(u(n),n=0..N),16);
                          .8556243918921488
```

On peut comparer cette valeur à la valeur approchée de l'intégrale obtenue par MAPLE par une méthode de quadrature numérique[(14)].

```
> evalf(Int(exp(-x^2/2),x=0..1,'method' = _d01ajc),16);

                          .8556243918921489
```

On peut remarquer la concordance entre ces deux valeurs approchées de $\int_0^1 \mathrm{e}^{-x^2/2}\,\mathrm{d}x$. On se convaincra, en utilisant par exemple la commande `time` de MAPLE, que le temps de calcul en utilisant une méthode de quadrature numérique est toutefois moins important.

Intéressons-nous maintenant aux propriétés de dérivabilité de la somme d'une série de fonctions.

(14) Voir la section 18.6 p. 883 et suivantes du *Cours de première année* pour une présentation de ce type de méthode.

Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU

THÉORÈME 3.9 *Soient I un intervalle ouvert de $\mathbb{R}$ et $\sum_n f_n$ une série d'applications de I dans $\mathbb{R}$. Si*

1. *pour tout $n \in \mathbb{N}$, l'application f_n est de classe $\mathcal{C}^1$ sur I ;*
2. *la série d'applications $\sum_n f_n'$ converge uniformément sur I et a pour somme une application ψ de I dans $\mathbb{R}$;*
3. *il existe un réel $x_0 \in I$ tel que la série numérique de terme général $f_n(x_0)$ converge ;*

alors la série $\sum_n f_n$ converge simplement sur I et a pour somme l'application $\phi : x \in I \mapsto \sum_{n=0}^{+\infty} f_n(x)$ qui est de classe $\mathcal{C}^1$ sur I et qui admet pour dérivée l'application ψ.

Démonstration Considérons la suite d'applications $(\phi_n)_n$ de terme général $\phi_n : x \in I \mapsto \sum_{k=0}^{n} f_k(x)$ et la suite d'applications $(\psi_n)_n$ de terme général $\psi_n : x \in I \mapsto \sum_{k=0}^{n} f_k'(x)$. La dérivée d'une somme finie d'applications étant la somme des dérivées de ces applications, pour tout $n \in \mathbb{N}$, on a

$$\phi_n'(x) = \left(\sum_{k=0}^{n} f_k(x)\right)' = \sum_{k=0}^{n} f_k'(x) = \psi_n(x).$$

La série d'applications $\sum_n f_n'$ converge uniformément sur I si et seulement si la suite d'applications $(\psi_n)_n$ converge uniformément sur I. De plus, pour $x_0 \in I$, la série numérique de terme général $f_n(x_0)$ converge si et seulement si la suite numérique de terme général $\phi_n(x_0)$ converge. D'après le théorème 2.5 p. 92, la suite $(\phi_n)_n$ converge vers une application de classe $\mathcal{C}^1$ sur I qui d'après la définition 3.4 est la somme de la série $\sum_n f_n$. Le théorème est démontré. □

Exemple La série $\sum_{n\geqslant 1} f_n$ de terme général $f_n : x \in \mathbb{R}_+^* \mapsto \frac{1}{n}\mathrm{e}^{-nx}$ converge simplement sur $\mathbb{R}_+^*$. Nous allons voir comment le théorème 3.9 peut être utilisé pour déterminer sa somme.

Pour tout $n \in \mathbb{N}^*$, l'application f_n est de classe $\mathcal{C}^1$ sur $\mathbb{R}_+^*$ car la fonction exponentielle est de classe $\mathcal{C}^1$ sur $\mathbb{R}$ et on a $f_n' : x \in \mathbb{R}_+^* \mapsto -\mathrm{e}^{-nx}$. Comme $\sup_{x\in\mathbb{R}_+^*} |f_n'(x)| = 1$, la suite $(f_n')_n$ ne converge pas uniformément vers la fonction nulle ; d'après la proposition 3.7 cela implique que la série $\sum_{n\geqslant 1} f_n'$ ne converge pas uniformément sur $\mathbb{R}_+^*$. Toutefois, pour tout $a \in \mathbb{R}_+^*$, la série $\sum_{n\geqslant 1} f_n'$ converge normalement sur $[a, +\infty[$, et par conséquent uniformément, car $\sup_{x\in[a,+\infty[} |f_n'(x)| = \mathrm{e}^{-na}$ et le critère de D'Alembert permet d'établir que la série $\sum_n \mathrm{e}^{-na}$ converge.

Les conditions du théorème 3.9 sont donc satisfaites sur $I = [a, +\infty[$ où $a > 0$. On en déduit que la série $\sum_{n\geqslant 1} f_n$ converge sur $[a, +\infty[$ vers une application ϕ qui est de classe $\mathcal{C}^1$ sur $[a, +\infty[$ et dont la dérivée en $x \in [a, +\infty[$ vaut

$$\phi'(x) = \sum_{n=0}^{+\infty} f_n'(x) = -\sum_{n=1}^{+\infty} \mathrm{e}^{-nx} = -\frac{\mathrm{e}^{-x}}{1-\mathrm{e}^{-x}} = \frac{1}{1-\mathrm{e}^{x}}.$$

Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU

On en conclut que

$$\phi(x) = -\int \frac{\mathrm{e}^{-x}}{1-\mathrm{e}^{-x}}\,\mathrm{d}x = -\ln(1-\mathrm{e}^{-x}) + C,$$

où C désigne une constante à déterminer. On a établi p. 136 que la somme ϕ de la série $\sum_{n\geqslant 1} f_n$ a pour limite 0 en $+\infty$. Comme $\lim_{x\to+\infty} -\ln(1-\mathrm{e}^{-x})+C = 0$, on en déduit que $C = 0$. La série $\sum_{n\geqslant 1} f_n$ a donc pour somme $x \mapsto -\ln(1-\mathrm{e}^{-x})$ sur $[a, +\infty[$.

Comme la série $\sum_{n\geqslant 1} f_n$ converge simplement sur $\mathbb{R}_+^*$, intéressons-nous à présent à la somme de la série $\sum_{n\geqslant 1} f_n$ sur $\mathbb{R}_+^*$. Considérons un réel x strictement positif ; si on pose $a = x/2$ alors $a > 0$ et $x \in [a, +\infty[$. On a donc d'après ce qui précède $\sum_n f_n(x) = -\ln(1-\mathrm{e}^{-x})$. On en conclut que sur $\mathbb{R}_+^*$ la somme de la série $\sum_{n\geqslant 1} f_n$ est l'application $\phi : x \in \mathbb{R}_+^* \mapsto -\ln(1-\mathrm{e}^{-x})$.

On peut affaiblir la seconde hypothèse du théorème 3.9. Il suffit que la série $(f_n')_n$ converge simplement sur I et converge uniformément sur tout intervalle fermé borné inclus dans I. On peut également affaiblir l'hypothèse de régularité $\mathcal{C}^1$ des applications f_n. Si les applications f_n sont uniquement dérivables sur I, alors on peut montrer que la série $\sum_n f_n$ converge simplement sur I vers l'application ϕ qui est dérivable sur I avec $\phi' = \psi$. La démonstration est toutefois plus délicate.

On peut donc énoncer le résultat suivant, qui est une version plus faible du théorème 3.9, que nous admettrons.

THÉORÈME 3.10 *Soient I un intervalle ouvert de $\mathbb{R}$ et $\sum_n f_n$ une série d'applications de I dans $\mathbb{R}$. Si*

1. *pour tout $n \in \mathbb{N}$ l'application f_n est dérivable sur I ;*
2. *la série d'applications $\sum_n f_n'$ converge simplement sur I et a pour somme une application ψ de I dans $\mathbb{R}$;*
3. *la série d'applications $\sum_n f_n'$ converge uniformément sur tout intervalle fermé borné inclus dans I ;*
4. *il existe un réel $x_0 \in I$ tel que la série numérique de terme général $f_n(x_0)$ converge ;*

alors la série $\sum_n f_n$ converge simplement sur I et a pour somme l'application $\phi : x \in I \mapsto \sum_{n=0}^{+\infty} f_n(x)$ qui est une application dérivable sur I dont la dérivée est ψ.

La fonction ζ de Riemann

Nous avons vu que la série de fonctions de terme général $f_n : x \mapsto 1/n^x$ converge simplement et absolument sur $]1, +\infty[$ mais qu'elle ne converge pas uniformément sur $]1, +\infty[$. Nous allons toutefois montrer que la fonction ζ de Riemann qui est la somme de cette série est continue sur $]1, +\infty[$. Soit x_0 un réel

Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU

de $]1,+\infty[$ fixé et $a = 1+\frac{1}{2}(1-x_0)$; on a $a > 1$ et $x_0 \in [a,+\infty[$. Comme $a > 1$, la série de fonctions de terme général $f_n : x \mapsto 1/n^x$ converge normalement sur $[a,+\infty[$ et par conséquent uniformément. Les applications $f_n : x \mapsto 1/n^x$ étant continues sur $\mathbb{R}$, elles le sont en x_0 et le théorème 3.7 permet de conclure que la fonction ζ est continue en x_0.

On a $\lim\limits_{x\to+\infty} f_1(x) = 1$ et pour $n \geqslant 2$ on a $\lim\limits_{x\to+\infty} f_n(x) = \lim\limits_{x\to+\infty} \mathrm{e}^{-x\ln(n)} = 0$. On déduit donc du corollaire 3.2 que

$$\lim_{x\to+\infty} \zeta(x) = 1.$$

Pour tout $n \in \mathbb{N}^*$, l'application f_n est dérivable sur $]1,+\infty[$ et a pour dérivée l'application $f_n' : x \in]1,+\infty[\mapsto -\ln(n)/n^x$. Pour tout $x \in]1,+\infty[$ fixé, la série numérique $\sum_{n\geqslant 1} f_n'(x)$ est une série de Bertrand[(15)] qui converge ; la série d'applications $\sum_{n\geqslant 1} f_n'$ convergente donc simplement sur $]1,+\infty[$. De plus, pour tout $(a,b) \in]1,+\infty[^2$ avec $a < b$ on a

$$\sup_{x\in[a,b]} |f_n'(x)| = \frac{\ln(n)}{n^a}$$

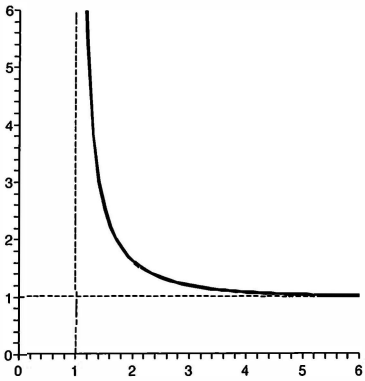
et $\sum_{n\geqslant 1} \ln(n)/n^a$ est une série de Bertrand convergente. La série d'applications $\sum_n f_n'$ converge donc normalement (et par conséquent uniformément) sur tout intervalle fermé borné inclus dans I. Le théorème 3.10 permet d'affirmer que la fonction ζ de Riemann est dérivable sur $]1,+\infty[$ et que

$$\forall x \in]1,+\infty[\qquad \zeta'(x) = -\sum_{n=1}^{+\infty} \frac{\ln(n)}{n^x} < 0.$$

La fonction ζ est donc strictement décroissante sur $]1,+\infty[$.

On peut visualiser la représentation graphique de la fonction ζ de Riemann à l'aide de MAPLE.

```
> des1:=plot([Zeta(x),1],x=0..6,y=0..6):
> des2:=plot([1,t,t=0..6]):
> display(des1,des2);
```



(15) Voir la définition 1.5 p. 26.

Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU

Exercice 3.7 On considère la série de fonctions $\sum_n f_n$ de terme général $f_n : x \mapsto \frac{1}{n^2} e^{-nx}$.
1 - Étudier la convergence de cette série et déterminer l'ensemble de définition $\mathcal{D}$ de la somme ϕ de cette série.
2 - Montrer que ϕ est une application continue sur $\mathcal{D}$.
3 - a) Montrer que pour tout réel $a > 0$, la série de fonctions $\sum_{n\geqslant 1} \frac{1}{n} e^{-nx}$ converge uniformément sur $]a, +\infty[$. Que peut-on dire en ce qui concerne la convergence uniforme sur $]0, +\infty[$ de cette série ?
b) Montrer que ϕ est une application dérivable sur $]0, +\infty[$.

Une curiosité : la fonction de Van der Waerden

Les séries de fonctions fournissent un moyen d'introduire des fonctions qui ont des propriétés sortant de l'ordinaire des fonctions usuelles. Par exemple, la fonction de Van der Waerden est une fonction qui est continue sur $\mathbb{R}$ mais qui n'est dérivable nulle part.

La fonction de Van der Waerden est définie comme la somme de la série de fonctions $\sum_n f_n$ où les fonctions f_n sont les fonctions périodiques de période 1 définies par récurrence sur $[0, 1[$ par :

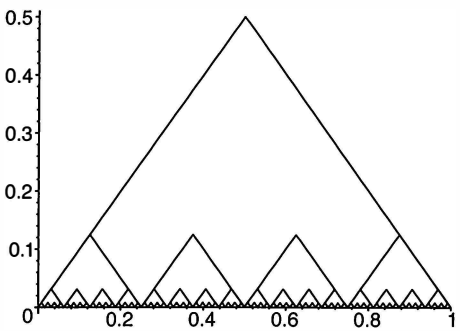
$$f_0(x) = x \quad \text{si } x \in [0, \tfrac{1}{2}] \quad \text{et} \quad f_0(x) = 1 - x \quad \text{si } x \in]\tfrac{1}{2}, 1[$$

et pour tout $n \in \mathbb{N}$:

$$f_n(x) = 4^{-n} f_0(4^n x) \quad \forall x \in [0, 1[.$$

On peut obtenir la représentation graphique des fonctions f_n avec MAPLE de la manière suivante.

```
> f01:= x->piecewise(x>0 and x< 1/2,x,x>1/2 and x<1,1-x):
> f[0]:= x->piecewise(x>0,f01(frac(x)),x<0,f01(frac(x)+1)):
> N:=3: for n from 1 to N do
          f[n] := x->4^(-n)*f[0](4^n*x):
        end:
> plot([seq(f[n](x),n=0..N)],x=0..1);
```



Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU

L'application f_0 est clairement continue sur $[0,1]$ et donc sur $\mathbb{R}$ compte tenu de la périodicité. Il s'ensuit que pour tout $n \in \mathbb{N}$, l'application f_n est continue sur $\mathbb{R}$. Il est facile d'établir que la série $\sum_n f_n$ converge normalement sur $\mathbb{R}$: on a

$$\sup_{x\in\mathbb{R}} |f_n(x)| = \sup_{x\in[0,1]} |f_n(x)| = 4^{-n} \sup_{x\in[0,1]} |f_0(4^n x)| \leqslant 4^{-n},$$

car f_0 est majorée par 1. La série $\sum_{n\geqslant 1} \frac{1}{4^n}$ est une série géométrique de raison $\frac{1}{4}$; il s'agit donc d'une série convergente. La convergence normale de la série $\sum_n f_n$ est alors une conséquence de la proposition 3.8. On en déduit que la fonction de Van der Waerden $\mathfrak{V} : x \mapsto \sum_{n=0}^{+\infty} f_n(x)$ est continue sur $\mathbb{R}$.

Montrons que la fonction de Van der Waerden $\mathfrak{V}$ n'est dérivable en aucun point de $\mathbb{R}$. Elle sera dérivable en $x_0 \in \mathbb{R}$ si le taux d'accroissement

$$\Delta_{x_0}(h) = \frac{\mathfrak{V}(x_0+h) - \mathfrak{V}(x_0)}{h}$$

admet une limite lorsque h tend vers 0. Pour montrer que $\mathfrak{V}$ n'est pas dérivable en x_0, il suffit[16] de trouver une suite $(h_n)_n$ qui converge vers 0 mais telle que la suite de terme général $\delta_n = \Delta_{x_0}(h_n)$ diverge. Considérons une suite $(h_n)_n$ de terme général $h_n = 4^{-n}\varepsilon_n$ où ε_n vaut ou bien 1 ou bien -1 (on précisera la valeur ultérieurement). Cette suite converge vers 0 et

$$\delta_n = \Delta_{x_0}(h_n) = \frac{\mathfrak{V}(x_0+4^{-n}\varepsilon_n) - \mathfrak{V}(x_0)}{4^{-n}\varepsilon_n} = \sum_{k=0}^{+\infty} \frac{f_k(x_0+4^{-n}\varepsilon_n) - f_k(x_0)}{4^{-n}\varepsilon_n}$$

$$= \sum_{k=0}^{+\infty} \frac{f_0(4^k x_0 + 4^{k-n}\varepsilon_n) - f_0(4^k x_0)}{4^{k-n}\varepsilon_n}.$$

Remarquons que si $k \geqslant n$ alors $4^{k-n}\varepsilon_n$ est un entier relatif et par conséquent puisque f_0 est 1-périodique, $f_0(4^k x_0 + 4^{k-n}\varepsilon_n) = f_0(4^k x_0)$. On en déduit que

$$\delta_n = \sum_{k=0}^{n-1} \frac{f_0(4^k x_0 + 4^{k-n}\varepsilon_n) - f_0(4^k x_0)}{4^{k-n}\varepsilon_n}.$$

Pour tout $k \in \{0,\dots,n-1\}$, notons $y_k = 4^k x_0$ et $p_k = \mathrm{E}(y_k)$ sa partie entière.

De deux choses l'une :

- si $y_{n-1} \in [p_{n-1}, p_{n-1}+\frac{1}{2}[$, l'un des deux réels $y_{n-1}-\frac{1}{4}$ ou $y_{n-1}+\frac{1}{4}$ appartient lui aussi à $[p_{n-1}, p_{n-1}+\frac{1}{2}[$;
- si $y_{n-1} \in [p_{n-1}+\frac{1}{2}, p_{n-1}+1[$, l'un des deux réels $y_{n-1}-\frac{1}{4}$ ou $y_{n-1}+\frac{1}{4}$ appartient lui aussi à $[p_{n-1}+\frac{1}{2}, p_{n-1}+1[$.

On choisit la valeur de ε_n de sorte que y_{n-1} et $y_{n-1}+\frac{1}{4}\varepsilon_n$ appartiennent au même intervalle $[p_{n-1}, p_{n-1}+\frac{1}{2}[$ ou $[p_{n-1}+\frac{1}{2}, p_{n-1}+1[$. Ceci implique que pour tout $k \in \{0,\dots,n-2\}$, y_k et $y_k + 4^{k-n}\varepsilon_n$ appartiennent eux aussi

[16] Voir la proposition 3.11 p. 575 du *Cours de première année* et la remarque 2 de la p. 576.

Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU

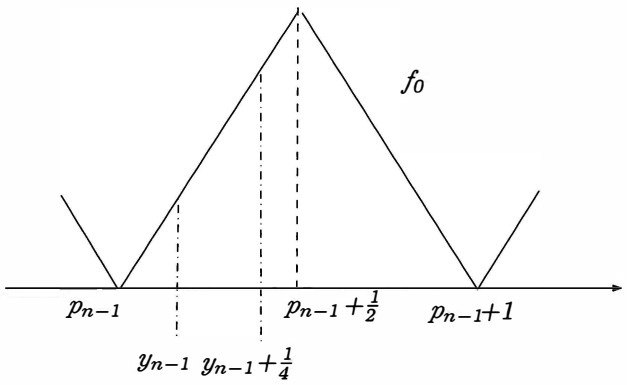


Fig. 1 Représentation de la fonction f_0 sur l'intervalle $[p_{n-1}, p_{n-1}+1]$.

nécessairement à un même intervalle : $[p_k, p_k+\frac{1}{2}[$ ou $[p_k+\frac{1}{2}, p_k+1[$. En effet, aucun rationnel z multiple de $\frac{1}{2}$ ne peut être compris strictement entre y_k et $y_k+4^{k-n}\varepsilon_n$ car

$$y_k < z < y_k + 4^{k-n}\varepsilon_n \quad \Longrightarrow \quad y_{n-1} < 4^{n-1-k}z < y_{n-1} + \frac{1}{4}.$$

Si z est un multiple de $\frac{1}{2}$, pour tout $k \in \{0, \ldots, n-2\}$ le rationnel $4^{n-1-k}z$ est lui aussi multiple de $\frac{1}{2}$ ce qui contredit le fait qu'avec le choix effectué pour la valeur de ε_n les deux réels y_{n-1} et $y_{n-1}+\frac{1}{4}$ appartiennent au même intervalle $[p_{n-1}, p_{n-1}+\frac{1}{2}[$ ou $[p_{n-1}+\frac{1}{2}, p_{n-1}+1[$ qui ne comprennent pas de multiple de $\frac{1}{2}$ dans leur intérieur.

De ceci, on déduit que

- $f_0(4^k x_0 + 4^{k-n}\varepsilon_n) - f_0(4^k x_0) = 4^{k-n}\varepsilon_n$ si $y_k \in [p_k, p_k+\frac{1}{2}[$;
- $f_0(4^k x_0 + 4^{k-n}\varepsilon_n) - f_0(4^k x_0) = -4^{k-n}\varepsilon_n$ si $y_k \in [p_k+\frac{1}{2}, p_k+1[$.

Dans tous les cas,

$$\left| \frac{f_0(4^k x_0 + 4^{k-n}\varepsilon_n) - f_0(4^k x_0)}{4^{k-n}\varepsilon_n} \right| = 1.$$

La suite $(\delta_n)_n$ ne peut converger car on a $|\delta_{n+1} - \delta_n| = 1$. La fonction de Van der Waerden n'est donc dérivable en aucun réel x_0.

3.4 Exercices de synthèse

Exercice 3.8

L'objectif de cet exercice est de déterminer une approximation de $\int_0^1 t^{at^b}\,dt$ pour a et b deux réels strictement positifs, type d'intégrale que MAPLE ne parvient pas à déterminer explicitement.

Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU

1 - Justifier que l'application $x \in]0,1] \mapsto x^{ax^b}$ est prolongeable en une application f continue sur $[0,1]$.

2 - On considère la série d'applications $\sum_n f_n$ définie par

$$f_n : x \in [0,1] \mapsto \begin{cases} \frac{1}{n!}\left(ax^b \ln(x)\right)^n & \text{si } x \in]0,1] \\ 0 & \text{si } x = 0 \end{cases}$$

a) Justifier que $\sup_{x \in]0,1]} |x^b \ln(x)| = \sup_{y \in]-\infty,0]} |y\,e^{by}|$ et calculer la valeur de ce supremum.

b) Montrer que $\sum_n f_n$ converge uniformément sur $[0,1]$ et a pour somme f.

3 - Montrer que $\displaystyle\int_0^1 t^{at^b}\,dt = \sum_{n=0}^{+\infty} \frac{1}{n!} \int_0^1 \left(at^b \ln(t)\right)^n\,dt$.

4 - Pour $\alpha \in \mathbb{R}_+^*$ et $n \in \mathbb{N}$, on note $I_{\alpha,n} = \displaystyle\int_0^1 t^\alpha (\ln t)^n\,dt$.

a) Montrer que pour tout $n \in \mathbb{N}^*$ on a $I_{\alpha,n} = -\frac{n}{\alpha+1} I_{\alpha,n-1}$.

b) En déduire la valeur de $I_{\alpha,n}$ (en fonction de n et de α).

5 - Pour tout $n \in \mathbb{N}$, on note $S_n = \displaystyle\sum_{k=0}^{n} \frac{(-1)^k a^k}{(bk+1)^{k+1}}$.

a) Montrer que $\displaystyle\int_0^1 t^{at^b}\,dt = \lim_{n \to +\infty} S_n$.

b) Montrer que pour tout $n \in \mathbb{N}^*$, $\left|\displaystyle\int_0^1 t^{at^b}\,dt - S_n\right| \leqslant \frac{1}{a}\left(\frac{a}{b(n+1)+1}\right)^{n+2}$

c) Interpréter ce résultat et conclure.

Exercice 3.9

Cet exercice utilise les résultats établis dans l'exercice 2.10 p. 96 où on étudie la suite d'applications $(\psi_n)_n$ de terme général $\psi_n : x \in \mathbb{R} \mapsto \int_1^n \frac{1}{t^x}\,dt$. Rappelons que l'on a montré au cours de cet exercice que pour tout $n \in \mathbb{N}^*$, l'application ψ_n est de classe $\mathcal{C}^1$ sur $\mathbb{R}$ et pour $x \in \mathbb{R} \setminus \{1\}$,

$$\psi_n'(x) = \frac{1}{x-1}\,\frac{1}{n^{x-1}}\left(\ln(n) + \frac{1}{x-1}\right) \quad \text{et} \quad \psi_n'(1) = -\frac{1}{2}\ln^2(n).$$

Par ailleurs on a établi que la suite de fonctions $(\psi_n)_n$ converge simplement sur $]1,+\infty[$ vers $\psi : x \in]1,+\infty[\mapsto \frac{1}{x-1}$.

On s'intéresse dans cet exercice à certaines propriétés de la fonction Zeta de Riemann, notée ζ, qui est la fonction de la variable réelle définie par $\zeta : x \mapsto \displaystyle\sum_{n=1}^{+\infty} \frac{1}{n^x}$.

1 - a) Déterminer l'ensemble de définition D de la fonction ζ.

Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU

b) Montrer que ζ est une application strictement décroissante sur D.

c) Y a-t-il convergence normale de la série $\sum_{n\geqslant 1}\frac{1}{n^x}$ sur D ?

2 - On considère la série de fonctions $\sum_{n\geqslant 1}\phi_n$ de terme général $\phi_n : x\in]0,+\infty[\mapsto \displaystyle\int_n^{n+1}\left(\frac{1}{n^x}-\frac{1}{t^x}\right)\,\mathrm{d}t$.

a) Calculer $\phi_n(1)$ et montrer que la série numérique $\sum_{n\geqslant 1}\phi_n(1)$ converge.

b) Montrer que pour tout $n\in\mathbb{N}^*$ et $x\in]0,+\infty[$ on a

$$\phi_n(x)=n^{-x}-\psi_{n+1}(x)+\psi_n(x).$$

En déduire que ϕ_n est de classe $\mathcal{C}^1$ sur $]0,+\infty[$ et donner l'expression de $\phi_n'(x)$ pour tout $x\in]0,+\infty[$.

c) En utilisant une primitivation par parties, montrer que pour tout $x\neq 1$ et $t\in[1,+\infty[$

$$\int\frac{\ln(t)}{t^x}\,\mathrm{d}t=-\frac{1}{(x-1)^2\,t^{x-1}}-\frac{\ln(t)}{(x-1)\,t^{x-1}}+C\quad(C\in\mathbb{R}).$$

En déduire que pour tout $x\in]0,+\infty[$ on a

$$\phi_n'(x)=-\int_n^{n+1}\left(\frac{\ln(n)}{n^x}-\frac{\ln(t)}{t^x}\right)\,\mathrm{d}t.$$

d) Montrer que la série $\sum_{n\geqslant 1}\phi_n'$ converge normalement sur tout intervalle $[a,b]$ où $0<a<b$. (Indication : on pourra utiliser l'expression de ϕ_n' obtenue à la question précédente et appliquer le théorème des accroissements finis à la fonction $y\mapsto y^{-x}\ln(y)$ pour $x\in[a,b]$ fixé.)

e) En utilisant les résultats du cours, conclure que la série de fonctions $\sum_{n\geqslant 1}\phi_n$ converge sur $]0,+\infty[$ et a pour somme une application Φ qui est de classe $\mathcal{C}^1$ sur $]0,+\infty[$.

3 - a) Montrer que pour tout $x\in D$ on a $\zeta(x)=\dfrac{1}{x-1}+\Phi(x)$.

b) En déduire que ζ est une application de classe $\mathcal{C}^1$ sur D.

c) Montrer que $\zeta(x)\underset{1^+}{\sim}\dfrac{1}{x-1}$.

d) Montrer que pour $n\in\mathbb{N}^*$ et $x\in\mathbb{R}_+^*$ on a $0\leqslant\phi_n(x)\leqslant n^{-x}-(n+1)^{-x}$; en déduire que $0\leqslant\Phi(x)\leqslant 1$ puis que $\Phi(x)\geqslant 1-\dfrac{1}{x-1}\left(1-2^{1-x}\right)$. En conclure que $\lim\limits_{x\to+\infty}\zeta(x)=1$.

4 - Donner le tableau de variation de la fonction ζ sur D ainsi que sa représentation graphique.

Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU

3.5 Solutions des exercices

Solution de l'exercice 3.1

1 - Pour $x \in \mathbb{R}$ fixé, la série numérique $\sum_n \cos(x)^n$ est une série géométrique de raison $\cos(x)$. Elle converge si et seulement si $|\cos(x)| < 1$, c'est-à-dire si $x \in \mathbb{R} \setminus \pi\mathbb{Z}$. La série de fonctions $\sum_n \cos(x)^n$ converge donc simplement sur $\mathbb{R} \setminus \pi\mathbb{Z}$. Pour $x \in \mathbb{R} \setminus \pi\mathbb{Z}$ on a

$$\sum_{n=0}^{+\infty} \cos(x)^n = \frac{1}{1-\cos(x)}.$$

2 - Pour $x \in \mathbb{R}$ fixé, la série numérique $\sum_n \exp(-nx)$ est une série géométrique de raison e^{-x}. Elle converge si et seulement si $|\mathrm{e}^{-x}| < 1$, c'est-à-dire si $x \in \mathbb{R}_+^*$. On en déduit que la série de fonctions $\sum_n \exp(-nx)$ converge simplement sur $\mathbb{R}_+^*$. Pour $x \in \mathbb{R}_+^*$ on a

$$\sum_{n=0}^{+\infty} \exp(-nx) = \frac{1}{1-\mathrm{e}^{-x}} = \frac{\mathrm{e}^x}{\mathrm{e}^x - 1}.$$

3 - Pour $x \in \mathbb{R}$ fixé on a $\sum_n x(1-x)^n = x\sum_n (1-x)^n$ et la série $\sum_n (1-x)^n$ est une série géométrique de raison $1-x$. Elle converge si et seulement si $|1-x| < 1$, c'est-à-dire si $x \in]0,2[$. La série de fonctions $\sum_n x(1-x)^n$ converge donc simplement sur $]0,2[$. Pour tout $x \in]0,2[$ on a

$$\sum_{n=0}^{+\infty} x(1-x)^n = x\sum_{n=0}^{+\infty} (1-x)^n = x\,\frac{1}{1-(1-x)} = 1.$$

Solution de l'exercice 3.2

1 - Pour $x \in \mathbb{R}$ fixé, appliquons le critère de D'Alembert à la série numérique de terme général $u_n = f_n(x)$. Pour tout $n \in \mathbb{N}$, on a

$$\left|\frac{u_{n+1}}{u_n}\right| = \frac{1}{1+x^2} \quad \text{et} \quad \ell = \lim_{n\to+\infty} \left|\frac{u_{n+1}}{u_n}\right| = \frac{1}{1+x^2}.$$

On en déduit que si $x \neq 0$ alors $\ell < 1$ et la série numérique $\sum_n u_n$ converge absolument. Par ailleurs, si $x = 0$ alors le critère de D'Alembert ne permet pas de conclure mais dans ce cas, $f_n(x) = f_n(0) = 0$ donc la série $\sum_n u_n$ est la série nulle qui converge. On en déduit que la série de fonctions $\sum_n f_n$ converge simplement (et même absolument) sur $\mathbb{R}$.

2 - Pour tout $x \in \mathbb{R}^*$ on a $|\frac{1}{1+x^2}| < 1$ et en utilisant les propriétés des séries géométriques, on obtient

$$\rho_n(x) = \sum_{k=n+1}^{+\infty} f_k(x) = x\sum_{k=n+1}^{+\infty} \left(\frac{1}{1+x^2}\right)^k = x\left(\frac{1}{1+x^2}\right)^{n+1} \sum_{k=0}^{+\infty} \left(\frac{1}{1+x^2}\right)^k$$
$$= x\left(\frac{1}{1+x^2}\right)^{n+1} \frac{1}{1-\frac{1}{1+x^2}} = \frac{1}{x(1+x^2)^n}.$$

Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU

Si $x = 0$ alors $\rho_n(x) = 0$. Pour tout $x \in \mathbb{R}^*$ on a $\rho_n(x) > 0$ et

$$\rho_n'(x) = -\frac{(2n+1)x^2+1}{x^2(1+x^2)^{n+1}} < 0.$$

La fonction ρ_n est donc positive et décroissante sur $\mathbb{R}^+$.

Comme $\lim_{x\to 0^+} \rho_n(x) = +\infty$, on en déduit que $|\rho_n|$ n'est pas majorée sur $\mathbb{R}^+$ et donc que la suite $(\rho_n)_n$ ne converge pas uniformément vers l'application nulle sur $\mathbb{R}^+$. D'après le corollaire 3.1, on en déduit que la série de fonctions $\sum_n f_n$ ne converge pas uniformément sur $[0, +\infty[$.

Toutefois pour tout $a > 0$, on a $\sup_{x\in[a,+\infty[} |\rho_n(x)| = \rho_n(a) = \frac{1}{a(1+a^2)^n}$ et $\lim_{x\to 0^+} \rho_n(a) = 0$ donc la suite $(\rho_n)_n$ converge uniformément vers l'application nulle sur $[a, +\infty[$.

On a ainsi établi que la série de fonctions $\sum_n f_n$ converge uniformément sur $[a, +\infty[$ pour tout réel $a > 0$ mais ne converge pas uniformément sur $[0, +\infty[$.

Solution de l'exercice 3.3

Nous avons établi au cours des exemples de la p. 117 que la série d'applications $\sum_n f_n$ converge simplement sur $\mathbb{R}_+^*$.

Pour tout $n \in \mathbb{N}^*$, l'application $f_n : x \in \mathbb{R}^+ \mapsto ne^{-nx}$ est continue, positive et strictement décroissante. On a donc $u_n = \sup_{x\in]0,+\infty[} |f_n(x)| = f_n(0) = n$. Comme la série $\sum_n u_n$ diverge grossièrement, la série $\sum_n f_n$ ne converge pas normalement sur $]0, +\infty[$.

Pour tout $r \in \mathbb{R}_+^*$ fixé, on a

$$u_n = \sup_{x\in[r,+\infty[} |f_n(x)| = f_n(r) = ne^{-nr}.$$

La série $\sum_n u_n$ converge car d'après le critère de D'Alembert, on a

$$\lim_{n\to+\infty} \left|\frac{u_{n+1}}{u_n}\right| = \lim_{n\to+\infty} \frac{n+1}{n}\, e^{-r} = e^{-r} < 1.$$

On en déduit que la série $\sum_n f_n$ converge normalement sur $[r, +\infty[$ pour tout réel $r > 0$.

Solution de l'exercice 3.4

Pour tout $x \in [a, b]$ on a $|f_n(x)| \leqslant |f_n(x) - f_n(a)| + |f_n(a)|$. D'après le théorème des accroissements finis, il existe un réel $c \in]a, b[$ tel que

$$|f_n(x) - f_n(a)| = (x-a)|f_n'(c)|.$$

Comme f_n' est continue sur $[a, b]$, elle est bornée (et atteint ses bornes). On en déduit que pour tout $x \in [a, b]$

$$|f_n(x)| \leqslant (b-a)\ \|f_n'\|_{\infty,[a,b]} + |f_n(a)|.$$

Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU

Ainsi,

$$\sup_{x\in[a,b]} |f_n(x)| \leqslant (b-a)\ \|f_n'\|_{\infty,[a,b]} + |f_n(a)|.$$

Comme la série de fonctions $\sum_n f_n'$ converge normalement sur $[a,b]$, la série numérique $\sum_n \|f_n'\|_{\infty,[a,b]}$ converge. Par ailleurs, puisque la série $\sum_n f_n(a)$ converge absolument, on en déduit que la série $\sum_n \sup_{x\in[a,b]} |f_n(x)|$ converge c'est-à-dire que la série $\sum_n f_n$ converge normalement sur $[a,b]$.

Solution de l'exercice 3.5

1 - Les fonctions f_n sont définies sur $\mathbb{R}$ pour tout $n \in \mathbb{N}^*$. Pour $x \in \mathbb{R}$ fixé, la série numérique $\sum_n \frac{(-1)^n}{x^2+n}$ est une série alternée. Pour $n \in \mathbb{N}^*$, notons $\alpha_n = \frac{1}{x^2+n}$. On a $\alpha_n > 0$ et $\lim_{n\to+\infty} \alpha_n = 0$. De plus, $\alpha_{n+1}/\alpha_n = \frac{x^2+n}{x^2+n+1} < 1$ donc la suite $(\alpha_n)_n$ converge vers 0. Le critère spécial pour les séries alternées (théorème 1.3 p. 42) permet d'en conclure que la série $\sum_n \frac{(-1)^n}{x^2+n}$ converge pour tout $x \in \mathbb{R}$. La série de fonctions $\sum_n f_n$ converge donc simplement sur $\mathbb{R}$.

2 - a) Utilisons le critère spécial de convergence uniforme pour les séries alternées pour montrer que la série de fonctions $\sum_n f_n$ converge uniformément sur $\mathbb{R}$. Pour $n \in \mathbb{N}^*$, posons $g_n : x \in \mathbb{R} \mapsto \frac{1}{x^2+n}$. Pour tout $x \in \mathbb{R}$ fixé, la suite numérique $(g_n(x))_n = (\alpha_n)_n$ est une suite à termes positifs, décroissante. De plus, la suite d'applications $(g_n)_n$ converge uniformément sur X vers l'application nulle puisque

$$\sup_{x\in\mathbb{R}} \left|\frac{1}{x^2+n}\right| = \frac{1}{n} \quad \text{et} \quad \lim_{n\to+\infty} \frac{1}{n} = 0.$$

On en déduit par le théorème 3.5 que la série d'applications $\sum_n f_n$ converge uniformément sur $\mathbb{R}$.

b) Montrons qu'il n'existe aucun intervalle de $\mathbb{R}$ où la convergence de la série $\sum_n f_n$ est normale. Considérons un intervalle I de $\mathbb{R}$ d'extrémités a et b avec $a < b$.

- Si $0 \in I$ alors $\sup_{x\in I} |f_n(x)| = \sup_{x\in I} \left|\frac{1}{x^2+n}\right| = \frac{1}{n}$; comme la série $\sum_n \frac{1}{n}$ diverge, on en déduit que la série $\sum_n f_n$ ne converge pas normalement sur I.
- Si $0 \notin I$ alors $\sup_{x\in I} |f_n(x)| = \sup_{x\in I} \left|\frac{1}{x^2+n}\right| = \frac{1}{c^2+n}$ où $c = \min(|a|,|b|)$; comme la série $\sum_n \frac{1}{c^2+n}$ diverge (on a $\frac{1}{c^2+n} \underset{+\infty}{\sim} \frac{1}{n}$), on en déduit que la série $\sum_n f_n$ ne converge pas normalement sur I.

Ainsi, la série $\sum_n f_n$ ne converge normalement sur aucun intervalle I de $\mathbb{R}$.

Solution de l'exercice 3.6

1 - Pour $x \in \mathbb{R}^*$ fixé, on a $\sum_n \frac{x}{(1+x^2)^n} = x \sum_n \left(\frac{1}{1+x^2}\right)^n$ et $|\frac{1}{1+x^2}| < 1$. La série numérique $\sum_n \frac{x}{(1+x^2)^n}$ est donc une série convergente. De plus,

$$\sum_{n=0}^{+\infty} \frac{x}{(1+x^2)^n} = x \sum_{n=0}^{+\infty} \left(\frac{1}{1+x^2}\right)^n = x \frac{1}{1-\frac{1}{1+x^2}} = \frac{1+x^2}{x}.$$

Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU

Par ailleurs, si $x = 0$ on a $f_n(x) = 0$ pour tout $n \in \mathbb{N}$ donc la série $\sum_n f_n$ est la série nulle et a pour somme 0.

On en conclut que la série $\sum_n \frac{x}{(1+x^2)^n}$ converge simplement sur $\mathbb{R}$ et a pour somme l'application

$$\phi : x \in \mathbb{R} \longmapsto \begin{cases} \dfrac{1+x^2}{x} & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases}$$

2 - Pour tout $n \in \mathbb{N}$, l'application f_n est continue sur $\mathbb{R}$ car il s'agit d'une fonction rationnelle sans pôle réel. Par contre, l'application ϕ n'est pas continue en 0 car $\lim_{x \to 0^+} \phi(x) = +\infty$. Si la série $\sum_n f_n$ convergeait uniformément sur $\mathbb{R}$ alors, d'après le théorème 3.7, sa somme ϕ serait une application continue sur $\mathbb{R}$. Comme ce n'est pas le cas, on peut affirmer que la série $\sum_n f_n$ ne converge pas uniformément sur $\mathbb{R}$.

Solution de l'exercice 3.7

1 - Les fonctions f_n sont définies sur $\mathbb{R}$, mais pour tout $x < 0$ fixé, $f_n(x)$ tend vers $+\infty$ quand n tend vers $+\infty$. On en déduit que la série de fonctions $\sum_{n \geqslant 1} f_n$ diverge grossièrement sur $]-\infty, 0[$.

Pour tout réel $x \geqslant 0$ et pour tout $n \in \mathbb{N}^*$ on a $-nx \leqslant 0$ donc $0 < \mathrm{e}^{-nx} \leqslant 1$. On en déduit que :

$$\forall n \in \mathbb{N}^* \quad \forall x \in \mathbb{R}^+ \quad 0 < f_n(x) \leqslant \frac{1}{n^2}.$$

Comme la série de Riemann $\sum_{n \geqslant 1} 1/n^2$ converge, on en déduit que la série $\sum_{n \geqslant 1} f_n$ converge normalement sur $\mathbb{R}^+$. On a en particulier convergence simple sur $\mathbb{R}^+$. L'ensemble de définition de ϕ est par conséquent $\mathcal{D} = \mathbb{R}^+$.

2 - Pour tout $n \in \mathbb{N}^*$, l'application $f_n : x \in \mathbb{R}^+ \mapsto \frac{1}{n^2}\mathrm{e}^{-nx}$ est clairement continue. Comme la série $\sum_{n \geqslant 1} f_n$ converge normalement sur $\mathbb{R}^+$, elle converge uniformément sur $\mathbb{R}^+$. D'après le théorème 3.7, la somme ϕ de la série de fonctions $\sum_{n \geqslant 1} f_n$ est une application continue sur $\mathbb{R}^+$.

3 - a) Soit a un réel strictement positif. Considérons la série d'applications de terme général $g_n : x \in]a, +\infty[\mapsto \frac{1}{n}\mathrm{e}^{-nx}$. Comme la fonction exponentielle est positive et strictement croissante sur $\mathbb{R}$, les applications g_n sont positives et strictement décroissantes sur $]a, +\infty[$. On a donc

$$\sup_{x \in]a, +\infty[} |g_n(x)| = g_n(a) = \frac{1}{n\,\mathrm{e}^{na}}.$$

Il est aisé de vérifier, par exemple, en utilisant le critère de D'Alembert, que la série numérique $\sum_{n \geqslant 1} \frac{1}{n\,\mathrm{e}^{na}}$ converge. On a

$$\left|\frac{u_{n+1}}{u_n}\right| = \frac{n\,\mathrm{e}^{na}}{(n+1)\,\mathrm{e}^{(n+1)\,a}} \frac{n}{n+1}\,\mathrm{e}^{-a}$$

Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU

et cette quantité tend vers e^{-a} qui est un réel strictement inférieur à 1 puisque $a > 0$. On en déduit que la série de fonctions $\sum_{n\geqslant 1} \frac{1}{n} e^{-nx}$ converge normalement et par conséquent uniformément sur $]a, +\infty[$.

Il n'y a pas convergence normale de cette série sur $]0, +\infty[$ car

$$\sup_{x\in]0,+\infty[} |g_n(x)| = g_n(0) = \frac{1}{n}$$

et la série $\sum_{n\geqslant 1} \frac{1}{n}$ diverge. On ne peut toutefois pas affirmer qu'il n'y a pas convergence uniforme sur $]0, +\infty[$.

b) L'application ϕ est dérivable sur $]0, +\infty[$ si pour tout réel $x_0 \in]0, +\infty[$ elle est dérivable en x_0. Soit x_0 un réel strictement positif fixé et $a = \frac{1}{2}x_0$; on a $x_0 \in]a, +\infty[$. On a montré que

- pour tout $n \in \mathbb{N}$, f_n est dérivable sur $]a, +\infty[$ et $f_n' = -g_n$;
- la série de fonctions $\sum_{n\geqslant 1} g_n$ converge uniformément sur $]a, +\infty[$;
- la série de fonctions $\sum_{n\geqslant 1} f_n$ converge simplement sur $]a, +\infty[$.

D'après le théorème 3.10, la somme ϕ de la série de fonctions $\sum_{n\geqslant 1} f_n$ est dérivable sur $]a, +\infty[$. Elle est donc dérivable en x_0 et

$$\phi'(x_0) = -\sum_{k=1}^{+\infty} \frac{1}{k} e^{-kx_0}.$$

On ne peut bien entendu rien conclure à ce niveau en ce qui concerne la dérivabilité de ϕ en 0 sans qu'une étude spécifique ne soit menée.

Solution de l'exercice 3.8

1 - En utilisant l'écriture exponentielle des fonctions puissances réelles, on obtient pour $x \in]0, 1]$

$$f(x) = \exp(a \ln(x) x^b) = \exp(a \ln(x) \exp(b \ln(x))).$$

La fonction logarithme étant continue sur $]0, 1]$ et la fonction exponentielle étant continue sur $\mathbb{R}$, f est continue sur $]0, 1]$. De plus, a et b étant strictement positifs on a

$$\lim_{x\to 0^+} a \ln(x) \exp(x \ln(b)) = \lim_{y\to -\infty} ay \exp(by) = \frac{a}{b} \lim_{z\to -\infty} z \exp(z) = 0.$$

Puisque la fonction exponentielle est continue en 0 et prend la valeur 1, on en déduit que f est prolongeable par continuité à droite en 0 en posant $f(0) = 1$.

Comme f est continue sur $[0, 1]$, elle est Riemann-intégrable sur $[0, 1]$ et le réel $\int_0^1 t^{at^b}\, dt$ existe bien pour a et b réels strictement positifs.

2 - a) On a $\sup_{x\in]0,1]} |x^b \ln(x)| = \sup_{x\in]0,1]} |\ln(x) \exp(b \ln(x))|$ et comme la fonction logarithme est une bijection de $]0, 1]$ dans $]-\infty, 0]$, on a $\sup_{x\in]0,1]} |x^b \ln(x)| = \sup_{y\in]-\infty,0]} |y e^{by}|$.

Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU

Pour déterminer la valeur de ce supremum, étudions l'application $\psi : x \in]-\infty, 0] \mapsto x\,\mathrm{e}^{bx}$. Cette application est dérivable et pour tout $x \in]-\infty, 0]$ on a $\psi'(x) = \mathrm{e}^{bx}(1 - bx)$. On a donc $\psi'(x) = 0$ si et seulement si $y = -1/b$. On en déduit le tableau de variation suivant pour ψ :

x	$-\infty$		$-\frac{1}{b}$		0
$\psi'(x)$		$-$	0	$+$	
$\psi(x)$	0	↘	$-\frac{1}{\mathrm{e}b}$	↗	0

On en conclut que $\sup_{x\in]0,1]} |x^b \ln(x)| = \frac{1}{\mathrm{e}b}$.

b) Pour montrer que la série $\sum_n f_n$ converge uniformément sur $[0, 1]$, nous allons montrer qu'elle converge normalement sur $[0, 1]$ en montrant que $u_n = \sup_{x\in[0,1]} |f_n(x)|$ est le terme général d'une série numérique convergente. On a

$$u_n = \sup_{x\in]0,1]} \left|\frac{1}{n!}\big(ax^b \ln(x)\big)^n\right| = \frac{a^n}{n!}\left(\sup_{x\in]0,1]} |x^b \ln(x)|\right)^n = \frac{1}{\mathrm{e}^n\, n!}\left(\frac{a}{b}\right)^n.$$

La série numérique $\sum_n u_n$ est une série à termes strictement positifs pour laquelle

$$\frac{u_{n+1}}{u_n} = \frac{a}{\mathrm{e}b}\,\frac{1}{n+1}$$

tend vers 0. La critère de D'Alembert nous assure que la série numérique $\sum_n u_n$ converge. On en conclut que la série $\sum_n f_n$ converge normalement et donc uniformément sur $[0, 1]$.

Pour tout $x \in]0, 1]$, on a

$$\sum_{n=0}^{+\infty} f_n(x) = \sum_{n=0}^{+\infty} \frac{1}{n!}\big(ax^b \ln(x)\big)^n = \exp(ax^b \ln(x))$$

car $\sum_{n=0}^{+\infty} \frac{y^n}{n!} = \mathrm{e}^y$ pour tout $y \in \mathbb{R}$. De plus, pour $x = 0$ on a

$$\sum_{n=0}^{+\infty} f_n(0) = f_0(0) + \sum_{n=1}^{+\infty} \underbrace{f_n(0)}_{=0} = f_0(0) = 1.$$

On en conclut que pour tout $x \in [0, 1]$ on a $\sum_{n=0}^{+\infty} f_n(x) = f(x)$.

3 - Utilisons le théorème 3.8 p. 136.

– Pour tout $n \in \mathbb{N}$, l'application f_n est continue sur $]0, 1]$ (car la fonction puissance b et la fonction logarithme sont continues sur $]0, 1]$). Pour $x \in]0, 1]$ on a

$$f_n(x) = \frac{1}{n!}\big(ax^b \ln(x)\big)^n = \frac{a^n}{n!}\left(\ln(x)\mathrm{e}^{b\ln(x)}\right)^n$$

Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU

et

$$\lim_{x\to 0+} \ln(x)\mathrm{e}^{b\ln(x)} = \lim_{y\to-\infty} y\mathrm{e}^{by} = 0$$

d'où $\lim_{x\to 0+} f_n(x) = 0 = f_n(0)$. On en déduit que f_n est continue sur $[0,1]$ et par conséquent que f_n est Riemann-intégrable sur $[0,1]$.

– La série $\sum_n f_n$ converge uniformément sur $[0,1]$.

D'après le théorème 3.8, f est Riemann-intégrable sur $[0,1]$ et

$$\int_0^1 f(x)\,\mathrm{d}x = \sum_{n=0}^{+\infty}\int_0^1 f_n(x)\,\mathrm{d}x = \sum_{n=0}^{+\infty}\int_0^1 \frac{1}{n!}\big(ax^b\ln(x)\big)^n\,\mathrm{d}x.$$

4 - a) Pour montrer que $I_{\alpha,n} = -\frac{n}{\alpha+1}I_{\alpha,n-1}$ pour tout $n\in\mathbb{N}^*$, on intègre par parties :

$$I_{\alpha,n} = \underbrace{\Big[\frac{1}{\alpha+1}t^{\alpha+1}\ln(t)^n\Big]_0^1}_{=0} - \frac{n}{\alpha+1}\int_0^1 t^\alpha\,\ln(t)^{n-1}\,\mathrm{d}t = I_{\alpha,n-1}.$$

b) Un raisonnement par récurrence, que nous omettons de détailler, permet d'en déduire que $I_{\alpha,n} = (-1)^n\,\frac{n!}{(\alpha+1)^n}\,I_{\alpha,0}$.

Comme $I_{\alpha,0} = \int_0^1 t^\alpha\,\mathrm{d}t = \frac{1}{\alpha+1}$ on a finalement $I_{\alpha,n} = (-1)^n\,\frac{n!}{(\alpha+1)^{n+1}}$.

5 - a) On a montré que $\int_0^1 t^{at^b}\,\mathrm{d}t = \sum_{n=0}^{+\infty}\frac{a^n}{n!}\int_0^1\big(t^b\ln(t)\big)^n\,\mathrm{d}t$. On en déduit que

$$\begin{aligned}\int_0^1 t^{at^b}\,\mathrm{d}t &= \sum_{n=0}^{+\infty}\frac{a^n}{n!}I_{nb,n} = \sum_{n=0}^{+\infty}\frac{a^n}{n!}(-1)^n\,\frac{n!}{(nb+1)^{n+1}}\\ &= \lim_{N\to+\infty}\sum_{n=0}^{N}(-1)^n\,\frac{a^n}{(nb+1)^{n+1}} = \lim_{N\to+\infty} S_N.\end{aligned}$$

b) La série $\sum_n(-1)^n\,\frac{a^n}{(nb+1)^{n+1}}$ est une série numérique alternée convergente (sa somme vaut $\int_0^1 t^{at^b}\,\mathrm{d}t$). D'après la proposition 1.15 p. 44, son reste d'ordre n qui est donné par

$$R_n = \sum_{k=n+1}^{+\infty}(-1)^k\,\frac{a^k}{(kb+1)^{k+1}} = \int_0^1 t^{at^b}\,\mathrm{d}t - S_n$$

est borné par le terme de rang $n+1$ de la série ; on a

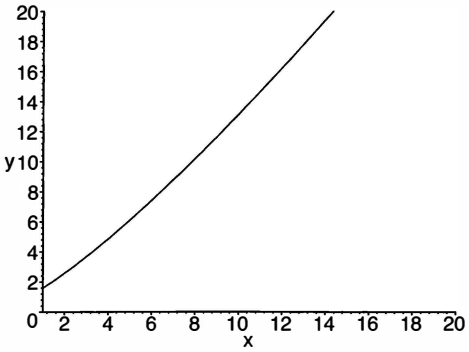
$$|R_n| \leqslant \frac{a^{n+1}}{((n+1)b+1)^{n+2}} = \frac{1}{a}\left(\frac{a}{(n+1)b+1}\right)^{n+2}$$

c) Une valeur approchée de $I = \int_0^1 t^{at^b}\,\mathrm{d}t$ peut être obtenue en calculant la somme $S_n = \sum_{k=0}^{n}(-1)^k\,\frac{a^k}{(kb+1)^{k+1}}$ pour une valeur de n « assez grande ». Pour $n\in\mathbb{N}$ donné, l'erreur commise en approchant I par S_n est inférieure à $E_n = \frac{1}{a}\left(\frac{a}{(n+1)b+1}\right)^{n+2}$.

Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU

Illustrons à l'aide de MAPLE la méthode étudiée dans cet exercice en considérant le calcul d'une valeur approchée de l'intégrale $\int_0^1 t^{\pi t^\pi}\,dt$. On peut représenter graphiquement le comportement de l'erreur E_n en fonction de la valeur de n. Afin d'avoir une figure lisible, on représente $-\log_{10}(E_n)$ en fonction de n.

```
> a:=Pi:b:=Pi:
> E := n-> (1/a)*(a/((n+1)*b+1))^(n+2):plot(-log10(E(x)),x=1..20);
```



Par exemple, pour $n = 8$, on lit sur la figure une valeur voisine de 10, ce qui signifie que l'erreur est de l'ordre de 10^{-10} ; S_8 approche la valeur de I avec 10 chiffres significatifs.

```
> S:=n-> sum((-1)^k*a^k/((b*k+1)^(k+1)),k=0..n):
> evalf(S(8));
                                .8399666383
```

On peut comparer cette valeur à celle obtenue par MAPLE en utilisant une quadrature numérique.

```
> evalf(int(t^(a*t^b),t=0..1));
                                .8399666382
```

Solution de l'exercice 3.9

1 - a) La série $\sum_n \frac{1}{n^x}$ est une série de Riemann ; elle converge si et seulement si $x > 1$. L'ensemble de définition de la fonction ζ est donc $]1, +\infty[$.

b) Soient $(x, y) \in]1, +\infty[^2$ tel que $x < y$; montrons que $\zeta(x) > \zeta(y)$. On a

$$\zeta(x) - \zeta(y) = \sum_{n=1}^{+\infty} \left(\frac{1}{n^x} - \frac{1}{n^y} \right) = \sum_{n=1}^{+\infty} \frac{n^y - n^x}{n^x n^y}.$$

Comme $x < y$, on a $n^x < n^y$ et donc $\frac{n^y - n^x}{n^x n^y} > 0$. La série $\sum_n \frac{n^y - n^x}{n^x n^y}$ est une série à termes strictement positifs, convergente (c'est la différence de deux séries convergentes). On en déduit que sa somme $\zeta(x) - \zeta(y)$ est un réel strictement positif. On a donc $\zeta(x) > \zeta(y)$. On en conclut que la fonction ζ est strictement croissante.

Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU

c) On a $\sup_{x\in]1,+\infty[}|\frac{1}{n^x}| = \frac{1}{n}$. Comme la série $\sum_n \frac{1}{n}$ diverge, la série de fonctions $\sum_n \frac{1}{n^x}$ ne converge pas normalement sur $]1,+\infty[$.

2 - a) Calculons $\phi_n(1)$. On a

$$\phi_n(1) = \int_n^{n+1} \left(\frac{1}{n} - \frac{1}{t}\right) \mathrm{d}t = \frac{1}{n} - \int_n^{n+1} \frac{1}{t}\,\mathrm{d}t = \frac{1}{n} - \ln\left(\frac{n+1}{n}\right).$$

On a $\ln(\frac{n+1}{n}) = \ln(1+\frac{1}{n}) = \frac{1}{n} + \frac{1}{2n^2} + \mathcal{O}_{+\infty}(\frac{1}{n^2})$ donc $\phi_n(1) \underset{+\infty}{\sim} \frac{1}{2n^2}$. On en déduit que la série $\sum_n \phi_n(1)$ est une série numérique à termes négatifs à partir d'un certain rang et comme la série numérique $\sum_n \frac{1}{2n^2}$ converge, $\sum_n \phi_n(1)$ converge aussi.

b) Pour $n \in \mathbb{N}^*$ et $x \in]0,+\infty[$ on a

$$\begin{aligned}\phi_n(x) &= \int_n^{n+1} \left(\frac{1}{n^x} - \frac{1}{t^x}\right) \mathrm{d}t = \frac{1}{n^x} - \int_1^{n+1} \frac{1}{t^x}\,\mathrm{d}t + \int_1^n \frac{1}{t^x}\,\mathrm{d}t \\ &= n^{-x} - \psi_{n+1}(x) + \psi_n(x).\end{aligned}$$

Comme ψ_n est de classe $\mathcal{C}^1$ sur $\mathbb{R}$ et que $x \mapsto \frac{1}{n^x}$ est de classe $\mathcal{C}^1$ sur $\mathbb{R}$, on en déduit que ϕ_n de classe $\mathcal{C}^1$ sur $\mathbb{R}$ et donc sur $]0,+\infty[$. D'après le préambule pour $x \in]0,1[\cup]1+\infty[$ on a

$$\begin{aligned}\phi_n'(x) &= \left(\frac{1}{n^x}\right)' - \psi_{n+1}'(x) + \psi_n'(x) \\ &= -\frac{\ln(n)}{n^x} - \frac{1}{x-1}\,\frac{1}{(n+1)^{x-1}}\left(\ln(n+1) + \frac{1}{x-1}\right) \\ &\quad + \frac{1}{x-1}\,\frac{1}{n^{x-1}}\left(\ln(n) + \frac{1}{x-1}\right).\end{aligned}$$

et

$$\phi_n'(1) = -\frac{\ln(n)}{n} + \frac{1}{2}\ln^2(n+1) - \frac{1}{2}\ln^2(n) = -\frac{\ln(n)}{n} + \frac{1}{2}\ln(1+\frac{1}{n}).$$

c) Pour $x \in]0,1[\cup]1,+\infty[$ et $t \in]1,+\infty[$, on obtient en primitivant par parties :

$$\begin{aligned}\int \frac{\ln(t)}{t^x}\,\mathrm{d}t &= \frac{1}{1-x}\,\frac{\ln(t)}{t^{x-1}} - \frac{1}{1-x}\int \frac{1}{t^x}\,\mathrm{d}t \\ &= -\frac{\ln(t)}{(x-1)\,t^{x-1}} - \frac{1}{(x-1)^2\,t^{x-1}} + C \quad (C \in \mathbb{R}).\end{aligned}$$

Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU

On a donc pour $x \in]0,1[\cup]1,+\infty[$ et $n \in \mathbb{N}^*$,

$$\begin{aligned}\int_n^{n+1}\left(\frac{\ln(n)}{n^x}-\frac{\ln(t)}{t^x}\right)\,\mathrm{d}t &= \frac{\ln(n)}{n^x}-\int_n^{n+1}\frac{\ln(t)}{t^x}\,\mathrm{d}t\\ &= \frac{\ln(n)}{n^x}-\left[-\frac{\ln(t)}{(x-1)\,t^{x-1}}-\frac{1}{(x-1)^2\,t^{x-1}}\right]_n^{n+1}\\ &= \frac{\ln(n)}{n^x}+\frac{1}{x-1}\,\frac{1}{(n+1)^{x-1}}\left(\ln(n+1)+\frac{1}{x-1}\right)\\ &\quad -\frac{1}{x-1}\,\frac{1}{n^{x-1}}\left(\ln(n)+\frac{1}{x-1}\right)\\ &= -\phi_n'(x).\end{aligned}$$

Par ailleurs, pour $x=1$,

$$\begin{aligned}\int_n^{n+1}\left(\frac{\ln(n)}{n}-\frac{\ln(t)}{t}\right)\,\mathrm{d}t &= \frac{\ln(n)}{n}-\int_n^{n+1}\frac{\ln(t)}{t}\,\mathrm{d}t=\frac{\ln(n)}{n}-\left[\frac{1}{2}\ln(t)^2\right]_n^{n+1}\\ &= \frac{\ln(n)}{n}-\frac{1}{2}\ln(n+1)^2+\frac{1}{2}\ln(n)^2=\phi_n'(1).\end{aligned}$$

La relation est donc vraie pour tout $x \in]0,+\infty[$.

d) La série de fonctions $\sum_{n\geqslant 1}\phi_n'$ converge normalement sur tout intervalle $[a,b]$ où $0<a<b$ si la série numérique de terme général $\sup_{x\in[a,b]}|\phi_n'(x)|$ converge. On a

$$|\phi_n'(x)|=\left|\int_n^{n+1}\frac{\ln(n)}{n^x}-\frac{\ln(t)}{t^x}\,\mathrm{d}t\right|.$$

Appliquons le théorème des accroissements finis[17] à l'application $g : y \in \mathbb{R}_+^* \mapsto \ln(y)/y^x$ entre $n \in \mathbb{N}^*$ et $t > n$ pour $x \in [a,b]$ fixé (g est dérivable sur $\mathbb{R}_+^*$) : il existe un réel $c \in]n,t[$ tel que

$$g(t)-g(n)=(t-n)g'(c)=(t-n)\,\frac{1-x\ln(c)}{c^{x+1}}.$$

On en déduit que pour tout $x \in [a,b]$, on a

$$|\phi_n'(x)|=\left|\int_n^{n+1}(t-n)\,\frac{1-x\ln(c)}{c^{x+1}}\,\mathrm{d}t\right|\leqslant\int_n^{n+1}(t-n)\,\frac{|1-x\ln(c)|}{c^{x+1}}\,\mathrm{d}t.$$

Pour $t \in [n,n+1]$ on a $0 \leqslant t-n \leqslant 1$ et $n<c<n+1$. On a donc

$$|1-x\ln(c)|=x\ln(c)-1\leqslant x\ln(c)\leqslant x\ln(n+1).$$

On en déduit que

$$|\phi_n'(x)|\leqslant\int_n^{n+1}\frac{1}{n^{x+1}}\,x\ln(n+1)\,\mathrm{d}t=\frac{x}{n^{x+1}}\ln(n+1)$$

[17] Voir le théorème 16.2 p.731 du *Cours de première année*.

Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU

et $\sup_{x\in[a,b]} |\phi_n'(x)| \leqslant \dfrac{b}{n^{a+1}} \ln(n+1)$. Or $\dfrac{1}{n^{a+1}} \ln(n+1) \underset{+\infty}{\sim} \dfrac{1}{n^{a+1}} \ln(n)$ et la série de Bertrand $\sum_n \frac{1}{n^\alpha (\ln(n))^\beta}$ converge si et seulement si $(\alpha > 1)$ ou $(\alpha = 1$ et $\beta > 1)$. On en déduit que puisque $a > 0$, la série $\sum_n \frac{b}{n^{a+1}} \ln(n+1)$ converge. La série de fonctions $\sum_n \phi_n'$ converge donc normalement sur $[a,b]$ pour tout $(a,b) \in \mathbb{R}^2$ avec $0 < a < b$.

e) On applique le théorème 3.9 p. 139 à la série de fonctions $\sum_n \phi_n$ sur $[a,b]$ pour $(a,b) \in \mathbb{R}^2$ avec $0 < a < 1 < b$. On a établi que :

- pour tout $n \in \mathbb{N}^*$ l'application ϕ_n est de classe $\mathcal{C}^1$ sur $[a,b]$;
- la série $\sum_n \phi_n'$ converge normalement donc uniformément sur $[a,b]$;
- la série numérique de terme général $\phi_n(1)$ converge.

On en conclut par le théorème 3.9 que la série $\sum_n \phi_n$ converge simplement sur $[a,b]$ vers une application $\Phi : x \in [a,b] \mapsto \sum_{n=0}^{+\infty} \phi_n(x)$ qui est de classe $\mathcal{C}^1$ sur $[a,b]$ et qui vérifie $\Phi'(x) = \sum_{n=0}^{+\infty} \phi_n'(x)$.

Il reste à justifier que Φ est de classe $\mathcal{C}^1$ sur $]0,+\infty[$, c'est-à-dire que pour tout $x \in]0,+\infty[$, l'application Φ est continue en x, dérivable en x, à dérivée continue en x. Soit $x \in]0,+\infty[$ fixé.

- Si $x \in]0,1[$, posons $a = \frac{1}{2}x$ et $b = 1 + x$; on a $0 < a < 1 < b$ et $x \in [a,b]$.
- Si $x \in [1,+\infty[$, posons $a = \frac{1}{2}$ et $b = 1 + x$; on a $0 < a < 1 < b$ et $x \in [a,b]$.

Comme Φ est de classe $\mathcal{C}^1$ sur $[a,b]$ et que $x \in [a,b]$, l'application Φ est continue en x, dérivable en x, à dérivée continue en x. On en conclut que Φ est de classe $\mathcal{C}^1$ sur $]0,+\infty[$.

3 - a) Pour tout $x \in]1,+\infty[$, on a $\Phi(x) = \lim_{n\to+\infty} \sum_{k=1}^n \phi_k(x)$ et

$$\sum_{k=1}^n \phi_k(x) = \sum_{k=1}^n \int_k^{k+1} \left(\frac{1}{k^x} - \frac{1}{t^x}\right) \mathrm{d}t = \sum_{k=1}^n \frac{1}{k^x} - \sum_{k=1}^n \int_k^{k+1} \frac{1}{t^x}\,\mathrm{d}t$$
$$= \sum_{k=1}^n \frac{1}{k^x} - \int_1^{n+1} \frac{1}{t^x}\,\mathrm{d}t = \sum_{k=1}^n \frac{1}{k^x} - \frac{1}{x-1}\left(1 - \frac{1}{n^{1-x}}\right).$$

En passant à la limite quand n tend vers $+\infty$, on en déduit que

$$\Phi(x) = \zeta(x) - \frac{1}{x-1}, \quad \textit{i.e.} \quad \zeta(x) = \Phi(x) + \frac{1}{x-1}.$$

Remarques

1. Bien que Φ soit définie sur $]0,+\infty[$, la relation

$$\Phi(x) = \zeta(x) - \frac{1}{x-1} = \zeta(x) - \int_1^{+\infty} \frac{1}{t^x}\,\mathrm{d}t$$

n'est valable que sur $]1,+\infty[$. Sur $]0,1]$ la série $\sum_n \frac{1}{n^x}$ diverge tout comme l'intégrale généralisée $\int_1^{+\infty} \frac{1}{t^x}\,\mathrm{d}t$ mais leur différence est bien définie.

Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU

2. La relation $\zeta(x) = \Phi(x) + \frac{1}{x-1}$ sert à prolonger à $]0,+\infty[$ la fonction ζ de Riemann puisque les deux fonctions Φ et $x \mapsto \frac{1}{x-1}$ sont bien définies sur $]0,1]$.

3. On aurait pu étudier la fonction ζ directement, sans avoir recours à l'étude de la série $\sum_n \phi_n$. La motivation de ce qui a été fait est justifiée par la remarque précédente. ◇

b) La fonction ζ est de classe $\mathcal{C}^1$ sur $]1,+\infty[$ car pour tout $x \in]1,+\infty[$, $\zeta(x) = \Phi(x) + \frac{1}{x-1}$ et les deux fonctions Φ et $x \mapsto \frac{1}{x-1}$ sont de classe $\mathcal{C}^1$ sur $]1,+\infty[$.

c) Comme Φ est de classe $\mathcal{C}^1$ sur $]0,+\infty[$, elle est continue en 1. Comme par ailleurs, la fonction $x \mapsto \frac{1}{x-1}$ n'est pas bornée en 1, on a $\Phi(x) = \mathcal{O}_1\left(\frac{1}{x-1}\right)$ et par conséquent $\zeta(x) \underset{1^+}{\sim} \frac{1}{x-1}$.

d) Comme l'application $t \in [1,+\infty[\mapsto \frac{1}{t^x}$ est décroissante, pour tout $n \in \mathbb{N}^*$,

$$\int_n^{n+1} \frac{1}{(n+1)^x}\,\mathrm{d}t = \frac{1}{(n+1)^x} \leqslant \int_n^{n+1} \frac{1}{t^x}\,\mathrm{d}t \leqslant \int_n^{n+1} \frac{1}{n^x}\,\mathrm{d}t = \frac{1}{n^x}.$$

On en déduit que la quantité $\phi_n(x) = \frac{1}{n^x} - \int_n^{n+1} \frac{1}{t^x}\,\mathrm{d}t$ vérifie

$$0 \leqslant \phi_n(x) \leqslant \frac{1}{n^x} - \frac{1}{(n+1)^x}.$$

Ainsi, pour tout $x \in \mathbb{R}_+^*$ et pour tout $n \in \mathbb{N}^*$ on a

$$0 \leqslant \sum_{k=1}^{n} \phi_k(x) \leqslant \sum_{k=1}^{n} \left(\frac{1}{k^x} - \frac{1}{(k+1)^x}\right) = 1 - \frac{1}{(n+1)^x}.$$

En passant à la limite quand n tend vers $+\infty$, on obtient par le théorème d'encadrement : $0 \leqslant \Phi(x) \leqslant 1$.

Comme pour tout $x \in \mathbb{R}_+^*$ et pour tout $n \in \mathbb{N}^*$ on a $\phi_n(x) \geqslant 0$, on en déduit que

$$\Phi(x) = \sum_{n=1}^{+\infty} \phi_n(x) \geqslant \phi_1(x) = \int_1^2 (1 - \frac{1}{t^x})\,\mathrm{d}t = 1 + \frac{1}{x-1}\left(2^{1-x} - 1\right).$$

On en conclut que pour tout $x \in \mathbb{R}_+^*$ on a : $1 + \frac{1}{x-1}\left(2^{1-x} - 1\right) \leqslant \Phi(x) \leqslant 1$.
En passant à la limite quand x tend vers $+\infty$, on obtient par le théorème d'encadrement : $\lim_{x\to+\infty} \Phi(x) = 1$. Comme $\lim_{x\to+\infty} \frac{1}{x-1} = 0$, on en conclut que $\lim\limits_{x\to+\infty} \zeta(x) = 1$.

4 - L'étude qui précède permet de dresser le tableau de variation suivant pour la fonction ζ :

x	1		$+\infty$
$\zeta'(x)$		$-$	
$\zeta(x)$	$+\infty$	↘	1

La représentation graphique de la fonction ζ de Riemann est donnée p. 141.

Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU

CHAPITRE 4

Introduction aux fonctions complexes de la variable complexe

Une fonction de $\mathbb{R}$ dans $\mathbb{C}$ n'est autre que la donnée simultanée de deux fonctions de $\mathbb{R}$ dans $\mathbb{R}$ (la partie réelle et la partie imaginaire) et les propriétés de telles fonctions sont des généralisations assez immédiates de celles des fonctions de $\mathbb{R}$ dans $\mathbb{R}$. Il en va tout autrement des fonctions de $\mathbb{C}$ dans $\mathbb{C}$ qui possèdent de nombreuses propriétés très particulières par rapport aux fonctions de $\mathbb{R}$ dans $\mathbb{R}$ en liaison avec la notion de dérivabilité.

On dit qu'une application f de $\mathbb{C}$ dans $\mathbb{C}$ est $\mathbb{C}$-dérivable en $z_0 \in \mathbb{C}$ si la limite[1] $\lim_{z \to z_0} \frac{f(z)-f(z_0)}{z-z_0}$ existe et est finie. Il s'agit là de la généralisation naturelle[2] de la définition de la dérivabilité pour les applications de $\mathbb{R}$ dans $\mathbb{R}$. Toutefois, les applications $\mathbb{C}$-dérivables en chaque point z_0 d'un ouvert U de $\mathbb{C}$ (on parle de fonction *holomorphes* sur U) ont des propriétés spécifiques que ne possèdent pas les fonctions de la variable réelle dérivables. Citons à titre d'exemple le fait qu'une application holomorphe sur un ouvert U de $\mathbb{C}$ (donc une fois $\mathbb{C}$-dérivable en chaque point de U) est indéfiniment $\mathbb{C}$-dérivable sur U (autrement dit, admet des dérivées successives à tout ordre en tout point U). On montre aussi qu'une application f de $\mathbb{C}$ dans $\mathbb{C}$ qui est holomorphe et qui vérifie $f(\frac{1}{n}) = 0$ pour tout $n \in \mathbb{N}^*$ est nulle sur $\mathbb{C}$. On peut également citer comme propriété des fonctions holomorphes le *principe du maximum* : si f est une fonction holomorphe sur un ouvert connexe U de $\mathbb{C}$ et si $|f|$ (le module de f) admet un maximum local en un point $z_0 \in U$ alors f est constante sur U. Les trois propriétés que nous venons de citer sont bien entendu fausses dans le cas des fonctions de la variable réelle dérivables.

Ce chapitre est entièrement consacré à l'étude des fonctions complexes de la variable complexe. Après avoir introduit quelques notions de topologie du plan complexe, nous étudierons les notions de limite, de continuité et de dérivabilité pour les fonctions complexes de la variable complexe. Nous apporterons également quelques compléments aux chapitres précédents consacrés aux suites et séries de fonctions dans le cas de la variable complexe.

---

[1] Les différentes notions concernant les fonctions de $\mathbb{C}$ dans $\mathbb{C}$ évoquées ici seront abordées en détail ultérieurement dans le chapitre.

[2] Voir la définition 16.1 p. 713 du *Cours de première année*.

Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU

4.1 Fonctions complexes d'une variable réelle

Considérons une application f d'un sous-ensemble X de $\mathbb{R}$ à valeurs dans $\mathbb{C}$ qui à un réel x associe le complexe $f(x)$. Puisque $f(x)$ dépend de la variable x, sa partie réelle et sa partie imaginaire dépendent aussi de x. Ainsi, étant donné une fonction f de X dans $\mathbb{C}$, il existe deux fonctions f_r et f_i de X dans $\mathbb{R}$ (déterminées de manière unique) telles que pour tout $x \in X$

$$f(x) = f_r(x) + \mathrm{i} f_i(x).$$

On note $f_r = \mathcal{R}\mathrm{e}(f)$ et $f_i = \mathcal{I}\mathrm{m}(f)$.

L'ensemble $\mathcal{A}(X, \mathbb{C})$ des applications de X dans $\mathbb{C}$ est un espace vectoriel sur $\mathbb{C}$.

4.1.1 Limite et continuité

On dit que f admet pour limite le complexe ℓ en $x_0 \in X$ si :

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists \eta \in \mathbb{R}_+^* \quad \forall x \in X \quad \big(|x - x_0| \leqslant \eta \implies |f(x) - \ell| \leqslant \varepsilon\big)$$

où $|x - x_0|$ désigne la valeur absolue de $x - x_0$ et

$$|f(x) - \ell| = \sqrt{(\mathcal{R}\mathrm{e}(f(x) - \ell))^2 + (\mathcal{I}\mathrm{m}(f(x) - \ell))^2}$$

désigne le module de $f(x) - \ell$.

Dire que f admet pour limite ℓ en x_0 équivaut à dire que les deux fonctions réelles de la variable réelle $\mathcal{R}\mathrm{e}(f)$ et $\mathcal{I}\mathrm{m}(f)$ ont pour limite en x_0 respectivement $\mathcal{R}\mathrm{e}(\ell)$ et $\mathcal{I}\mathrm{m}(\ell)$. En effet, d'une part

$$\begin{aligned} & |\mathcal{R}\mathrm{e}(f(x)) - \mathcal{R}\mathrm{e}(\ell)| = |\mathcal{R}\mathrm{e}(f(x) - \ell)| \leqslant |f(x) - \ell| \\ \text{et} \quad & |\mathcal{I}\mathrm{m}(f(x)) - \mathcal{I}\mathrm{m}(\ell)| = |\mathcal{I}\mathrm{m}(f(x) - \ell)| \leqslant |f(x) - \ell| \end{aligned}$$

donc si f admet pour limite ℓ en x_0 alors

$$\lim_{x \to x_0} |\mathcal{R}\mathrm{e}(f(x)) - \mathcal{R}\mathrm{e}(\ell)| = \lim_{x \to x_0} |\mathcal{I}\mathrm{m}(f(x)) - \mathcal{I}\mathrm{m}(\ell)| = 0$$

autrement dit $\lim\limits_{x \to x_0} \mathcal{R}\mathrm{e}(f(x)) = \mathcal{R}\mathrm{e}(\ell)$ et $\lim\limits_{x \to x_0} \mathcal{I}\mathrm{m}(f(x)) = \mathcal{I}\mathrm{m}(\ell)$. D'autre part, on a

$$\begin{aligned} |f(x) - \ell| &= |\mathcal{R}\mathrm{e}(f(x)) - \mathcal{R}\mathrm{e}(\ell) + \mathrm{i}\big(\mathcal{I}\mathrm{m}(f(x)) - \mathcal{I}\mathrm{m}(\ell)\big)| \\ &\leqslant |\mathcal{R}\mathrm{e}(f(x)) - \mathcal{R}\mathrm{e}(\ell)| + |\mathcal{I}\mathrm{m}(f(x)) - \mathcal{I}\mathrm{m}(\ell)| \end{aligned}$$

donc si les deux fonctions $\mathcal{R}\mathrm{e}(f)$ et $\mathcal{I}\mathrm{m}(f)$ ont pour limite en x_0 respectivement $\mathcal{R}\mathrm{e}(\ell)$ et $\mathcal{I}\mathrm{m}(\ell)$ alors $\lim\limits_{x \to x_0} |f(x) - \ell| = 0$, *i.e.* f admet pour limite ℓ en x_0.

On dit que f est continue en $x_0 \in X$ si :

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists \eta \in \mathbb{R}_+^* \quad \forall x \in X \quad \big(|x - x_0| \leqslant \eta \implies |f(x) - f(x_0)| \leqslant \varepsilon\big),$$

où $|f(x) - f(x_0)|$ désigne le module de $f(x) - f(x_0)$.

Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU

Dire que f est continue en x_0 équivaut à dire que les deux fonctions réelles de la variable réelle $\mathcal{R}\mathrm{e}(f)$ et $\mathcal{I}\mathrm{m}(f)$ sont continues en x_0.

On dit que f est continue sur X si pour tout $x \in X$, l'application f est continue en x. On note $\mathcal{C}^0(X, \mathbb{C})$ l'ensemble des applications de X dans $\mathbb{C}$ continues sur X. Il s'agit d'un espace vectoriel sur $\mathbb{C}$.

On dit que f est continue sur l'intervalle fermé borné $[a, b]$ si les applications réelles de la variable réelle $\mathcal{R}\mathrm{e}(f)$ et $\mathcal{I}\mathrm{m}(f)$ sont continues sur $[a, b]$, *i.e.* si pour tout $x \in]a, b[$ elles sont continues en x et si, de plus, elles sont continues à droite en a et à gauche en b.

4.1.2 Dérivabilité

On dit qu'une application f d'un sous-ensemble X de $\mathbb{R}$ à valeurs dans $\mathbb{C}$ est dérivable en $x_0 \in X$ si la quantité

$$\Delta f_{x_0}(x) = \frac{f(x) - f(x_0)}{x - x_0} = \frac{\mathcal{R}\mathrm{e}(f)(x) - \mathcal{R}\mathrm{e}(f)(x_0)}{x - x_0} + \mathrm{i}\,\frac{\mathcal{I}\mathrm{m}(f)(x) - \mathcal{I}\mathrm{m}(f)(x_0)}{x - x_0}$$

admet une limite lorsque x tend vers x_0. Cette limite notée $f'(x_0)$ est appelée *(nombre) dérivée de f en x_0*.

On dit que f est dérivable sur l'ensemble $A \subset X$ si, pour tout $x \in A$, f est dérivable en x. On appelle alors dérivée de f et on note f' l'application qui à $x \in A$ associe le complexe $f'(x)$.

On montre que l'application f est dérivable en $x \in X$ si et seulement si les deux applications réelles d'une variable réelle $\mathcal{R}\mathrm{e}(f)$ et $\mathcal{I}\mathrm{m}(f)$ sont dérivables en x ; on a alors

$$f'(x) = \mathcal{R}\mathrm{e}(f)'(x) + \mathrm{i}\,\mathcal{I}\mathrm{m}(f)'(x) = \mathcal{R}\mathrm{e}(f'(x)) + \mathrm{i}\,\mathcal{I}\mathrm{m}(f'(x)).$$

Exemples

1. L'application $f : x \in \mathbb{R} \mapsto \cos(x) + \mathrm{i}\sin(x) \in \mathbb{C}$ est continue et dérivable sur $\mathbb{R}$. Pour tout $x \in \mathbb{R}$ on a $f'(x) = -\sin(x) + \mathrm{i}\cos(x)$. On remarquera qu'en introduisant la notation exponentielle complexe, cette dernière relation s'écrit :

$$\boxed{\left(\mathrm{e}^{\mathrm{i}x}\right)' = \mathrm{i}\,\mathrm{e}^{\mathrm{i}x}.}$$

2. L'application $f : x \in \mathbb{R} \mapsto \alpha x + \beta \in \mathbb{C}$, où $(\alpha, \beta) \in \mathbb{C}^2$, est continue et dérivable sur $\mathbb{R}$. Pour tout $x \in \mathbb{R}$ on a $f'(x) = \alpha$.

Comme dans le cas des fonctions de $\mathbb{R}$ dans $\mathbb{R}$, on peut définir par récurrence la dérivée d'ordre k ($k \in \mathbb{N}^*$) d'une fonction f de $\mathbb{R}$ dans $\mathbb{C}$, que l'on note $f^{(k)}$, de la manière suivante : $f^{(1)} = f'$ et, pour tout $\ell \in \mathbb{N}$ avec $2 \leqslant \ell \leqslant k$,

$$f^{(\ell)} = (f^{(\ell-1)})'.$$

On dit que f est de classe $\mathcal{C}^k$ sur X si f admet une dérivée d'ordre k ($k \in \mathbb{N}^*$) sur X qui est continue. On note $\mathcal{C}^k(X, \mathbb{C})$ l'ensemble des applications de X dans $\mathbb{C}$ de classe $\mathcal{C}^k$ sur X. Dire que f est de classe $\mathcal{C}^k$ sur X équivaut à dire que $\mathcal{I}\mathrm{m}(f)$ et $\mathcal{R}\mathrm{e}(f)$ sont de classe $\mathcal{C}^k$ sur X.

Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU

4.1.3 Intégrabilité

On dit qu'une application f définie sur un intervalle $[a, b]$ de $\mathbb{R}$, à valeurs complexes, est Riemann-intégrable sur $[a, b]$ si les applications réelles de la variable réelle $\mathcal{R}\mathrm{e}(f)$ et $\mathcal{I}\mathrm{m}(f)$ sont toutes les deux Riemann-intégrables sur $[a, b]$. On appelle alors intégrale de f sur $[a, b]$ et on note $\int_a^b f(x)\,\mathrm{d}x$ le complexe

$$\int_a^b f(x)\,\mathrm{d}x = \int_a^b \mathcal{R}\mathrm{e}(f)(x)\,\mathrm{d}x + \mathrm{i}\int_a^b \mathcal{I}\mathrm{m}(f)(x)\,\mathrm{d}x.$$

Exemple Les fonctions sinus et cosinus étant continues sur $[0, \frac{\pi}{4}]$, elles sont Riemann-intégrables sur cet intervalle. La fonction $f : x \in \mathbb{R} \mapsto \mathrm{e}^{\mathrm{i}x}$ est par conséquent Riemann-intégrable sur $[0, \frac{\pi}{4}]$ et

$$\int_0^{\frac{\pi}{4}} \mathrm{e}^{\mathrm{i}x}\,\mathrm{d}x = \int_0^{\frac{\pi}{4}} \cos(x)\,\mathrm{d}x + \mathrm{i}\int_0^{\frac{\pi}{4}} \sin(x)\,\mathrm{d}x = \frac{1}{\sqrt{2}} + \mathrm{i}(1 - \frac{1}{\sqrt{2}}).$$

4.1.4 Applications bornées

On dit qu'une application f définie sur sous-ensemble X de $\mathbb{R}$, à valeurs complexes, est *bornée* sur X si

$$\exists K \in \mathbb{R}^+ \quad \forall x \in X \quad |f(x)| \leqslant K.$$

On en déduit qu'une application f d'un sous-ensemble X de $\mathbb{R}$ dans $\mathbb{C}$ n'est pas bornée sur X si

$$\forall K \in \mathbb{R}^+ \quad \exists x \in X \quad |f(x)| > K.$$

Exemples

1. L'application $f : x \in \mathbb{R} \mapsto \cos(x) + \mathrm{i}\sin(x) \in \mathbb{C}$ est bornée sur $\mathbb{R}$ puisque

$$\forall x \in \mathbb{R} \quad |f(x)| = \sqrt{\cos^2(x) + \sin^2(x)} = 1.$$

2. L'application $f : x \in \mathbb{R} \mapsto \alpha x + \beta \in \mathbb{C}$, où $(\alpha, \beta) \in \mathbb{C}^* \times \mathbb{C}$, n'est pas bornée. En effet, pour tout $K \in \mathbb{R}^+$, on peut trouver un réel x tel que $|f(x)| > K$; il suffit de prendre $x = (K + 1 + |\beta|)/|\alpha|$ car on a alors

$$|f(x)| = |\alpha x + \beta| \geqslant \big||\alpha x| - |\beta|\big| = K + 1 > K.$$

Toutefois, la restriction de f à un intervalle I borné de $\mathbb{R}$ est une application bornée. En notant a et b les extrémités de I, on a en effet, pour tout $x \in I$,

$$|f(x)| = |\alpha x + \beta| \leqslant |\alpha|\,|x| + |\beta| \leqslant |\alpha|\,\max\{|a|, |b|\} + |\beta|.$$

Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU

4.1.5 Inégalité des accroissements finis

Dans le cas d'une fonction de $\mathbb{R}$ dans $\mathbb{C}$, le théorème des accroissement finis[3] n'est plus valable sous la forme d'une égalité. Considérons par exemple l'application $f : x \in [0, 2\pi] \mapsto e^{ix} \in \mathbb{C}$ qui est continue sur $[0, 2\pi]$ et dérivable sur $]0, 2\pi[$. Il n'existe pas de réel $c \in]0, 2\pi[$ tel que $f(2\pi) - f(0) = 2\pi f'(c)$ car $f(2\pi) - f(0) = 0$ alors que $2\pi f'(c) = 2\pi i e^{ic} = 2\pi i\,(\cos(c) + i\sin(c))$ n'est jamais nul. On a toutefois le résultat suivant.

THÉORÈME 4.1 (Inégalité des accroissements finis)
Soit f une application définie sur l'intervalle $[a, b]$ et à valeurs dans $\mathbb{C}$. Si f est continue sur $[a, b]$, dérivable sur $]a, b[$ et si f' est bornée sur $]a, b[$ alors

$$|f(b) - f(a)| \leqslant (b - a) \sup_{x \in]a,b[} |f'(x)|.$$

Démonstration Notons $f_r = \mathcal{R}e(f)$ et $f_i = \mathcal{I}m(f)$ la partie réelle et la partie imaginaire de f respectivement. Les applications f_r et f_i sont continues sur $[a, b]$ et dérivables sur $]a, b[$.

$\trianglerighteq$ Si $f(b) - f(a)$ est réel, alors $|f(b) - f(a)| = |f_r(b) - f_r(a)|$ et en appliquant le théorème des accroissements finis[3] à la fonction réelle de la variable réelle f_r, on établit l'existence d'un réel $c \in]a, b[$ tel que

$$|f(b) - f(a)| = |f_r(b) - f_r(a)| = (b - a)|f_r'(c)|.$$

Comme

$$|f_r'(c)| \leqslant \sup_{x \in]a,b[} |f_r'(x)| = \sup_{x \in]a,b[} |\mathcal{R}e(f'(x))| \leqslant \sup_{x \in]a,b[} |f'(x)|,$$

le résultat est démontré sous l'hypothèse que $f(b) - f(a)$ est réel.

$\trianglerighteq$ Si $f(b) - f(a)$ n'est pas réel, il existe $(r, \theta) \in \mathbb{R}_+^* \times]-\pi, 0[\cup]0, \pi[$ tel que $f(b) - f(a) = re^{i\theta}$. Considérons l'application $g_\theta : x \in [a, b] \mapsto e^{-i\theta} f(x)$. Elle possède les propriétés suivantes.

- L'application g_θ est continue sur $[a, b]$, dérivable sur $]a, b[$ et l'application $g_\theta' : x \in]a, b[\mapsto e^{-i\theta} f'(x)$ est bornée sur $]a, b[$; comme $|e^{-i\theta}| = 1$, on a
$$\sup_{x \in]a,b[} |g_\theta'(x)| = \sup_{x \in]a,b[} |f'(x)|$$
- De plus, $g_\theta(b) - g_\theta(a) = e^{-i\theta}\,(f(b) - f(a)) = r$ ce qui montre que $g_\theta(b) - g_\theta(a)$ est réel. Notons aussi que $|g_\theta(b) - g_\theta(a)| = |f(b) - f(a)|$.

En appliquant la première partie de la démonstration à l'application g, on obtient :

$$|f(b) - f(a)| = |g_\theta(b) - g_\theta(a)| \leqslant (b - a) \sup_{x \in]a,b[} |g_\theta'(x)| = (b - a) \sup_{x \in]a,b[} |f'(x)|.$$

Le résultat est démontré. □

[3] Voir le théorème 16.2 p. 731 du *Cours de première année*.

Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU

4.2 Topologie du plan complexe

Nous allons étendre au plan complexe un certain nombre de définitions qui ont été introduites dans le *Cours de première année*[4] sur la topologie de la droite réelle. Cette partie concernant la topologie du plan complexe constitue par ailleurs une introduction à la topologie dans un espace métrique qui sera présentée au chapitre 7.

4.2.1 Ouverts et fermés de $\mathbb{C}$

DÉFINITION 4.1 *Soient z_0 un complexe et r un réel strictement positif.*

✗ *On appelle* boule ouverte *(ou disque ouvert) de centre z_0 et de rayon r l'ensemble*

$$D(z_0, r) = \{z \in \mathbb{C} \ ; \ |z - z_0| < r\}.$$

✗ *On appelle* boule fermée *(ou disque fermé) de centre z_0 et de rayon r l'ensemble*

$$\overline{D}(z_0, r) = \{z \in \mathbb{C} \ ; \ |z - z_0| \leqslant r\}.$$

On pourra consulter la figure 1 pour une représentation graphique du disque ouvert $D(z_0, r)$ et du disque fermé $\overline{D}(z_0, r)$. On a coutume de figurer par un trait plein la partie de la frontière de l'ensemble dont les points appartiennent à l'ensemble et par un trait pointillé la partie de la frontière correspondant à des points n'appartenant pas à l'ensemble.

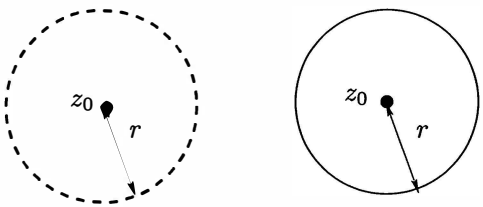


Fig. 1 Représentation graphique de la boule ouverte $D(z_0, r)$ (à gauche) et de la boule fermée $\overline{D}(z_0, r)$ (à droite).

Remarque On peut étendre la définition 4.1 au cas où $r = 0$. La boule ouverte $D(z_0, 0)$ est l'ensemble vide et la boule fermée $\overline{D}(z_0, 0)$ est le singleton $\{z_0\}$. ◇

On sera attentif à ne pas confondre au niveau des notations la boule fermée de centre z_0 et de rayon r notée $\overline{D}(z_0, r)$ et le complémentaire dans $\mathbb{C}$ de la boule ouverte de centre z_0 et de rayon r que nous noterons $\complement D(z_0, r)$.

[4] Voir la section 3.3 p. 109 du *Cours de première année*.

Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU

DÉFINITION 4.2 (Voisinage d'un complexe)
On dit que le sous-ensemble Z de $\mathbb{C}$ est un voisinage du complexe z_0 si Z contient une boule ouverte de centre z_0, i.e. si

$$\exists r \in \mathbb{R}_+^* \quad D(z_0, r) \subset Z.$$

Exemples

1. Par définition, toute boule (ouverte ou fermée) de centre z_0 est un voisinage de z_0.

2. L'ensemble $\mathbb{C}$ est voisinage de tout complexe z_0 puisque $D(z_0, 1) \subset \mathbb{C}$.

3. L'ensemble $Z = \{z \in \mathbb{C} \ ; \ \mathcal{R}e(z) \geqslant 0\}$ est un voisinage de $z_0 = 1 + \mathrm{i}$ (on vérifie sans peine que la boule ouverte de centre $1 + \mathrm{i}$ et de rayon $\frac{1}{2}$ est incluse dans Z). L'ensemble Z contient le complexe 0, mais ce n'est pas un voisinage de 0 car toute boule ouverte $D(0, r)$ contient le complexe $z = -\frac{r}{2}(1 + \mathrm{i})$ qui est de partie réelle nulle donc qui n'appartient pas à Z.

DÉFINITION 4.3 (Ensemble ouvert et ensemble fermé)

✗ *Un sous-ensemble non vide Z de $\mathbb{C}$ est qualifié d'*ensemble ouvert *si pour tout élément z de Z, il existe une boule ouverte de centre z contenue dans Z, i.e. si*

$$\forall z \in Z \quad \exists r \in \mathbb{R}_+^* \quad D(z, r) \subset Z.$$

✗ *Un sous-ensemble Z de $\mathbb{C}$ est appelé* ensemble fermé *si son complémentaire dans $\mathbb{C}$ est un ensemble ouvert.*

Remarques

1. Il résulte des définitions 4.2 et 4.3 qu'un ensemble est ouvert s'il est un voisinage de chacun de ses points.

2. Par convention, l'ensemble vide est un ouvert de $\mathbb{C}$. Son complémentaire dans $\mathbb{C}$ est donc un fermé de $\mathbb{C}$, autrement dit $\mathbb{C}$ est un fermé de $\mathbb{C}$. Nous avons vu que $\mathbb{C}$ est voisinage de chacun de ses points ; c'est donc un ouvert de $\mathbb{C}$. Son complémentaire dans $\mathbb{C}$ est par conséquent un fermé de $\mathbb{C}$, autrement dit l'ensemble vide est un fermé de $\mathbb{C}$. On retiendra que l'ensemble vide et $\mathbb{C}$ sont à la fois ouverts et fermés dans $\mathbb{C}$.

3. Un ensemble peut être ni ouvert ni fermé. Contrairement à son sens dans le langage courant, « fermé » n'est pas le contraire d' « ouvert ». De plus, en dehors de l'ensemble vide et de $\mathbb{C}$, aucun sous-ensemble de $\mathbb{C}$ n'est à la fois un ouvert et un fermé de $\mathbb{C}$. ◇

Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU

Exemples

1. Montrons que toute boule fermée de $\mathbb{C}$ est un ensemble fermé (ce qui justifie *a posteriori la terminologie introduite à la définition 4.1). Considérons la boule fermée* $\overline{D}(z_0, r_0)$ *où* $z_0 \in \mathbb{C}$ *et* $r_0 \in \mathbb{R}_+^*$. *Pour montrer que* $\overline{D}(z_0, r_0)$ *est un ensemble fermé, il faut montrer que son complémentaire* $\complement\overline{D}(z_0, r_0)$ *est un ensemble ouvert de* $\mathbb{C}$. *Désignant par* z_1 *un élément quelconque de* $\complement\overline{D}(z_0, r_0)$, *il s'agit de montrer que l'on peut trouver une boule ouverte de centre* z_1 *incluse dans* $\complement\overline{D}(z_0, r_0)$. *Un petit dessin, voir la figure 2, permet de repérer comme candidat potentiel la boule de rayon* $r_1 = |z_1 - z_0| - r_0$.

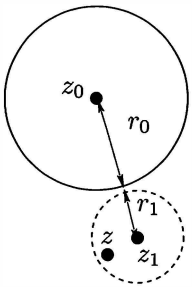


Fig. 2 Boule ouverte de centre z_1 incluse dans $\complement\overline{D}(z_0, r_0)$.

Pour tout $z \in D(z_1, r_1)$, *l'inégalité triangulaire nous donne*

$$|z_1 - z_0| \leqslant |z_1 - z| + |z - z_0| < r_1 + |z - z_0| = |z_1 - z_0| - r_0 + |z - z_0|.$$

On en déduit que $|z - z_0| > r_0$ *ce qui signifie que tous les éléments* z *de* $D(z_1, r_1)$ *appartiennent bien à* $\complement\overline{D}(z_0, r_0)$. *L'ensemble* $\complement\overline{D}(z_0, r_0)$ *est donc bien un ensemble ouvert, ce qui implique que la boule fermée* $\overline{D}(z_0, r_0)$ *est un ensemble fermé.*

2. *On montre également que toute boule ouverte de* $\mathbb{C}$ *est un ensemble ouvert, voir l'exercice 4.1.*

3. *L'ensemble* $Z = \{z \in \mathbb{C} \ ; \ |z| > r\}$ *(où* $r \in \mathbb{R}_+^*$*) est le complémentaire de la boule fermée* $\overline{D}(0, r)$ *; c'est donc un ensemble ouvert.*

4. *L'ensemble* $Z = \{z \in \mathbb{C} \ ; \ \mathcal{I}m(z) > 0\}$ *est un ensemble ouvert puisque pour tout* $z \in Z$ *la boule ouverte* $B(z, \frac{1}{2}\mathcal{I}m(z))$ *est incluse dans* Z. *Par conséquent l'ensemble* $\{z \in \mathbb{C} \ ; \ \mathcal{I}m(z) \leqslant 0\}$ *est un ensemble fermé.*

5. *Pour tout* $z_0 \in \mathbb{C}$, *le singleton* $\{z_0\}$ *est un ensemble fermé puisque l'ensemble* $\complement\{z_0\} = \mathbb{C} \setminus \{z_0\}$ *est ouvert. En effet, pour tout* $z \in \mathbb{C} \setminus \{z_0\}$ *la boule ouverte* $D(z, |z - z_0|)$ *est incluse dans* $\mathbb{C} \setminus \{z_0\}$.

Exercice 4.1 Montrer que toute boule ouverte de $\mathbb{C}$ est un ensemble ouvert.

Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU

L'intervalle $]0, 1[$ est un ensemble ouvert dans $\mathbb{R}$. Mais le « même ensemble » considéré dans $\mathbb{C}$, *i.e.* l'ensemble $\{z \in \mathbb{C} \ ; \ \mathcal{R}\mathrm{e}(z) \in]0, 1[\text{ et } \mathcal{I}\mathrm{m}(z) = 0\}$, n'est pas un sous-ensemble ouvert dans $\mathbb{C}$. Il est facile de voir par exemple que dans $\mathbb{C}$, il n'existe aucune boule ouverte de centre $\frac{1}{2}$ qui soit incluse dans cet ensemble. Ce n'est pas non plus un ensemble fermé dans $\mathbb{C}$.

Proposition 4.1 (Union et intersection d'ouverts ou de fermés)

✗ *L'union d'un nombre quelconque*[5] *d'ensembles ouverts est un ensemble ouvert. L'intersection d'un nombre* fini *d'ensembles ouverts est un ensemble ouvert.*

✗ *L'intersection d'un nombre quelconque d'ensembles fermés est un ensemble fermé. L'union d'un nombre* fini *d'ensembles fermés est un ensemble fermé.*

Démonstration Ce résultat sera démontré dans un cadre plus général au chap. 7. La seconde partie de la proposition concernant les ensembles fermés s'obtient à partir de la première partie de la proposition en utilisant les lois de Morgan[6]. □

Exemples

1. L'ensemble $\{z \in \mathbb{C} \ ; \ \mathcal{I}\mathrm{m}(z) = 0\}$ est un ensemble fermé car son complémentaire dans $\mathbb{C}$ est la réunion des deux ouverts $\{z \in \mathbb{C} \ ; \ \mathcal{I}\mathrm{m}(z) > 0\}$ et $\{z \in \mathbb{C} \ ; \ \mathcal{I}\mathrm{m}(z) < 0\}$.

2. La couronne $\{z \in \mathbb{C} \ ; \ r_1 < |z| < r_2\}$ où $(r_1, r_2) \in \mathbb{R}_+^* \times \mathbb{R}_+^*$ avec $r_1 < r_2$ est un ensemble ouvert. En effet, son complémentaire dans $\mathbb{C}$ qui est l'ensemble $\overline{D}(0, r_1) \cup \complement D(0, r_2)$ est la réunion de deux ensembles fermés.

3. Le cercle $C(z_0, r) = \{z \in \mathbb{C} \ ; \ |z - z_0| = r\}$ où $z_0 \in \mathbb{C}$ et $r \in \mathbb{R}_+^*$ est un ensemble fermé car son complémentaire dans $\mathbb{C}$ qui est $D(z_0, r) \cup \complement \overline{D}(z_0, r)$ est la réunion de deux ensembles ouverts.

On sera attentif au fait que l'intersection d'un nombre infini d'ensembles ouverts n'est pas nécessairement un ensemble ouvert, de même que la réunion d'un nombre infini d'ensembles fermés n'est pas nécessairement un ensemble fermé. Par exemple, l'intersection pour tous les entiers naturels n non nuls des boules ouvertes $D(z_0, \frac{1}{n})$ est le singleton $\{z_0\}$ qui est un ensemble fermé et par conséquent[7] qui n'est pas ouvert.

[5] Fini, dénombrable ou infini non dénombrable.

[6] Voir la proposition 2.3 p. 30 du *Cours de première année*.

[7] Rappelons que les seuls sous-ensembles de $\mathbb{C}$ qui sont à la fois ouverts et fermés dans $\mathbb{C}$ sont $\mathbb{C}$ et l'ensemble vide, voir p. 165.

Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU

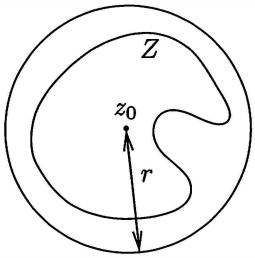


Fig. 3 Exemple d'ensemble Z borné dans $\mathbb{C}$.

DÉFINITION 4.4 *On dit qu'un sous-ensemble Z de $\mathbb{C}$ est borné s'il existe un élément z_0 dans $\mathbb{C}$ et un réel strictement positif r tel que Z soit inclus dans la boule fermée $\overline{B}(z_0, r)$, i.e. si*[8]

$$\exists z_0 \in \mathbb{C} \quad \exists r \in \mathbb{R}^+ \quad \forall z \in Z \quad |z - z_0| \leqslant r.$$

Remarques

1. D'après la définition 4.4, un sous-ensemble Z de $\mathbb{C}$ n'est pas borné si

$$\forall z_0 \in \mathbb{C} \quad \forall r \in \mathbb{R}^+ \quad \exists z \in Z \quad |z - z_0| > r.$$

2. Dans $\mathbb{C}$, un ensemble qui est fermé et borné est qualifié de compact. ◇

Exemples

1. Il résulte de la définition 4.4 que toute boule, ouverte ou fermée, dans $\mathbb{C}$ est un ensemble borné.

2. Un ensemble fini $Z = \{\zeta_1, \ldots, \zeta_n\}$ est borné. En effet, en prenant $z_0 = 0$ et $r = \max\{|\zeta_1|, \ldots, |\zeta_n|\}$, pour tout $z \in Z$ on a $|z| \leqslant \max\{|\zeta_1|, \ldots, |\zeta_n|\} = r$.

3. L'ensemble $\mathbb{C}$ n'est pas borné car pour tout $z_0 \in \mathbb{C}$ et pour tout $r \in \mathbb{R}^+$, le complexe $z = z_0 + r + 1$ est tel que $|z - z_0| = |r + 1| > r$.

4.2.2 Intérieur et adhérence d'un ensemble

DÉFINITION 4.5 (Intérieur)
Soient Z un sous-ensemble non vide de $\mathbb{C}$ et $z_0 \in \mathbb{C}$. On dit que z_0 est un point intérieur *à Z s'il existe une boule ouverte de centre z_0 qui soit incluse dans Z, i.e. si*

$$\exists r \in \mathbb{R}_+^* \quad D(z_0, r) \subset Z.$$

L'ensemble des points intérieurs à Z est noté $\mathrm{Int}\,(Z)$ *ou* $\mathring{Z}$ *et est appelé* intérieur *de Z.*

[8] On obtient une définition similaire en considérant une boule ouverte au lieu d'une boule fermée. Par ailleurs, l'inégalité triangulaire permet de vérifier que la propriété d'être bornée ne dépend pas du choix de z_0.

Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU

Exemples

1. L'intérieur de la boule fermée $\overline{D}(z_0, r)$ est la boule ouverte $D(z_0, r)$.

2. L'intérieur de l'ensemble fermé $\{z \in \mathbb{C} \ ; \ \mathcal{I}\mathrm{m}(z) \leqslant 0\}$ est l'ensemble ouvert $\{z \in \mathbb{C} \ ; \ \mathcal{I}\mathrm{m}(z) < 0\}$.

Remarques

1. Il résulte de la définition 4.5 qu'un point intérieur à Z appartient obligatoirement à Z. On a donc $\mathrm{Int}\,(Z) \subset Z$.

2. D'après la définition 4.2 p. 165, l'élément z_0 est un point intérieur à Z si et seulement si Z est un voisinage de z_0.

3. L'intérieur de Z est un ensemble ouvert ; on peut d'ailleurs montrer que[9] c'est le plus grand ouvert contenu dans Z. ◇

DÉFINITION 4.6 (Adhérence)
Soit Z un sous-ensemble non vide de $\mathbb{C}$ et $z_0 \in \mathbb{C}$.

✗ *On dit que z_0 est un* point adhérent *à Z si toute boule ouverte de centre z_0 contient au moins un élément de Z. L'ensemble des points adhérents à Z est noté $\overline{Z}$ et est appelé* adhérence *de Z.*

✗ *On dit que z_0 est un* point d'accumulation *de Z si toute boule ouverte de centre z_0 contient au moins un élément de Z autre que z_0.*

De manière équivalente, z_0 est un point adhérent à Z si tout voisinage de z_0 contient au moins un élément de Z et z_0 est un point d'accumulation de Z si tout voisinage de z_0 contient au moins un élément de Z autre que z_0.

Remarques

1. Un point d'accumulation de Z est un point adhérent de Z mais la réciproque est fausse. Un point adhérent à l'ensemble Z qui n'est pas un point d'accumulation est appelé un *point isolé*. L'adhérence d'un sous-ensemble de $\mathbb{C}$ est donc l'union disjointe de ses points d'accumulation et de ses points isolés.

2. L'adhérence de Z est un ensemble fermé ; on peut montrer[10] que c'est le plus petit ensemble fermé contenant Z. On a donc $Z \subset \overline{Z}$. ◇

Exemples

1. L'adhérence de la boule ouverte $D(z_0, r)$ est la boule fermée $\overline{D}(z_0, r)$, ce qui justifie la notation utilisée dans la définition 4.1.

2. L'adhérence de l'ensemble ouvert $\{z \in \mathbb{C} \ ; \ \mathcal{I}\mathrm{m}(z) < 0\}$ est l'ensemble fermé $\{z \in \mathbb{C} \ ; \ \mathcal{I}\mathrm{m}(z) \leqslant 0\}$.

[9] Cette affirmation sera établie dans un cadre plus général p. 377.
[10] Cette affirmation sera établie dans un cadre plus général p. 379.

Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU

3. L'ensemble $\{z \in \mathbb{C} \ ; \ \mathcal{I}\mathrm{m}(z) = 0\}$ (autrement dit le sous-ensemble de $\mathbb{C}$ constitué par les nombres réels) a pour intérieur l'ensemble vide et puisqu'il s'agit d'un ensemble fermé, son adhérence est l'ensemble lui-même.

4.2.3 Arcs géométriques dans le plan complexe

L'étude des courbes dans le plan sera effectuée au chap. 11. L'objectif de cette partie est de fournir quelques définitions et propriétés concernant les courbes du plan complexe indispensables pour manipuler la notion de limite d'une fonction complexe de la variable complexe.

Définition 4.7 (Arc paramétré dans $\mathbb{C}$)

✗ *On appelle* arc paramétré *(ou* chemin*) dans* $\mathbb{C}$ *toute application* ϕ *d'un intervalle* I *de* $\mathbb{R}$ *à valeurs dans* $\mathbb{C}$ *qui est continue.*

✗ *Si* ϕ *est de classe* C^k *sur* I, $k \in \mathbb{N} \cup \{+\infty\}$, *alors l'arc paramétré est dit de classe* C^k.

Définition 4.8 (Changement de paramètre)

Soient I *et* J *deux intervalles et* $k \in \mathbb{N} \cup \{+\infty\}$. *Deux arcs paramétrés* $\phi \in C^0(I, \mathbb{C})$ *et* $\psi \in C^0(J, \mathbb{C})$ *sont dits* C^k-équivalents *s'il existe une application bijective* θ *de* I *dans* J *qui est de classe* C^k *sur* I *et*[(11)] *dont la bijection réciproque* θ^{-1} *est de classe* C^k *sur* J, *telle que*

$$\psi = \phi \circ \theta.$$

On dit alors que ϕ *se déduit de* ψ *par changement de paramètre* θ.

Exemple Montrons que les deux arcs paramétrés

$$\phi : t \in]0, \pi[\longmapsto \mathrm{e}^{\mathrm{i}t} \quad \text{et} \quad \psi : t \in]-1, 1[\longmapsto t + \mathrm{i}\sqrt{1 - t^2}$$

représentés à la figure 4 sont C^∞-équivalents

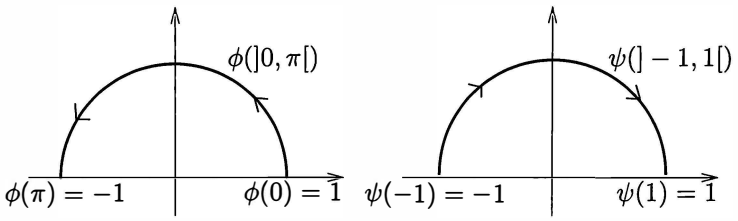


Fig. 4 Représentation de deux arcs paramétrés dans $\mathbb{C}$.

(11) Une telle application est appelée un C^k-difféomorphisme de I sur J, voir p. 664.

Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU

Pour tout $t \in]-1,1[$, on a[12] $\cos(\arccos(t)) = t$ et $\sin(\arccos(t)) = \sqrt{1-t^2}$ de sorte que

$$\psi(t) = t + i\sqrt{1-t^2} = \cos(\arccos(t)) + \sin(\arccos(t)) = \phi(\theta(t))$$

où la fonction θ est définie pour tout $t \in]-1,1[$ par $\theta(t) = \arccos(t)$. La fonction arc-cosinus est une bijection de $]-1,1[$ dans $]0,\pi[$ de classe $\mathcal{C}^\infty$ dont la bijection réciproque, qui est la fonction cosinus, est de classe $\mathcal{C}^\infty$ sur $]0,\pi[$. Les deux arcs paramétrés ϕ et ψ sont donc $\mathcal{C}^\infty$-équivalents.

Remarques

1. La continuité de l'application θ sur I implique[13] la continuité de sa bijection réciproque sur J. Mais le fait que θ soit de classe $\mathcal{C}^k$ sur I (avec $k \in \mathbb{N}^*$) n'implique pas nécessairement que θ^{-1} soit de classe $\mathcal{C}^k$ sur J. Il faut pour cela[14] que la dérivée de θ ne s'annule pas sur I.

2. La relation « les arcs paramétrés ϕ et ψ sont $\mathcal{C}^k$-équivalents » définit une relation d'équivalence[15] sur l'ensemble des arcs paramétrés de $\mathbb{C}$.

3. Deux arcs paramétrés $\mathcal{C}^k$-équivalents ont la même image dans $\mathbb{C}$. ◇

Définition 4.9 (Arcs géométriques)
On appelle arc géométrique de classe $\mathcal{C}^k$ *une classe d'équivalence γ d'arcs paramétrés de classe $\mathcal{C}^k$ pour la relation « être $\mathcal{C}^k$-équivalents ».*

- *Les arcs paramétrés de la classe γ sont appelés* représentations paramétriques admissibles *(ou* paramétrisations*) de l'arc géométrique γ.*
- *L'image commune des arcs paramétrés de la classe γ est appelée* support *de l'arc géométrique γ.*
- *Un arc géométrique est dit* simple *si toute paramétrisation de cet arc est injective.*

Exemples

1. Pour $r \in \mathbb{R}_+^*$ fixé, l'application $\phi : t \in [0,2\pi] \mapsto re^{it}$ définit un arc paramétré de classe $\mathcal{C}^\infty$. Le support de l'arc géométrique γ associé à ϕ est le cercle de centre l'origine et de rayon r. Cet arc géométrique n'est pas simple car $\phi(0) = \phi(2\pi) = r$.

2. L'application $\psi : t \in [-\pi,\pi] \mapsto -re^{it}$ définit un arc paramétré de classe $\mathcal{C}^\infty$ qui est $\mathcal{C}^\infty$-équivalent à ϕ puisque $\psi = \phi \circ \theta$ où $\theta : t \in [-\pi,\pi] \mapsto \pi - t \in [0,2\pi]$

[12] Voir le *Cours de première année* p. 650.
[13] Voir la proposition 14.1 p. 623 du *Cours de première année*.
[14] Voir la proposition 14.2 p. 624 du *Cours de première année*.
[15] Voir la définition 2.28 p. 55 du *Cours de première année*.

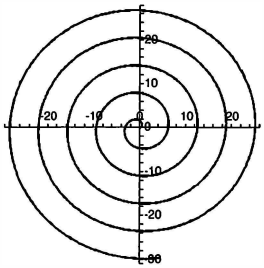
Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU

est un $\mathcal{C}^\infty$-difféomorphisme de $[-\pi, \pi]$ dans $[0, 2\pi]$. Les deux arcs paramétrés ϕ et ψ sont donc associés au même arc géométrique γ.

3. Pour $(z_0, z_1) \in \mathbb{C}^* \times \mathbb{C}^*$ fixé, l'application $\phi : t \in \mathbb{R} \mapsto z_1 + t z_0$ définit un arc paramétré de classe $\mathcal{C}^\infty$. Le support de l'arc géométrique associé à ϕ est la droite passant par le point d'affixe z_1 et de vecteur directeur $(\mathcal{R}\mathrm{e}(z_0), \mathcal{I}\mathrm{m}(z_0))$. Cet arc géométrique est simple.

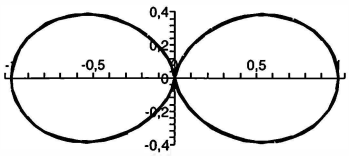
4. Le support de l'arc géométrique associé à $\phi : t \in [0, +\infty[\mapsto t\,\mathrm{e}^{\mathrm{i}t}$, connu sous le nom de *Spirale d'Archimède*[16], peut être aisément représenté avec MAPLE.

```
> with(plots):
> phi:= t-> t*exp(I*t):
> complexplot(phi(t),t=0..30,scaling=constrained);
```



5. Pour $r \in \mathbb{R}_+^*$ fixé, l'application $\phi : t \in [0, 2\pi[\mapsto r\big(\cos^3(t) + \mathrm{i}\cos^2(t)\sin(t)\big)$ définit un arc paramétré de classe $\mathcal{C}^\infty$. Le support de l'arc géométrique γ associé à ϕ se visualise aisément avec MAPLE :

```
> phi:= t-> cos(t)^3+I*sin(t)*cos(t)^2:
> complexplot(phi(t),t=0..2*Pi,scaling=constrained);
```



L'arc géométrique γ n'est pas simple car $\phi(\pi/2) = \phi(3\pi/2) = 0$, autrement dit l'arc « passe deux fois par l'origine ».

---

[16] La spirale d'Archimède est la trajectoire d'un point se déplaçant uniformément sur une droite du plan, cette droite tournant elle-même uniformément autour d'un de ses points. Cette courbe a été étudiée par ARCHIMÈDE DE SYRACUSE (287-212 av. J.-C.). Dans la nature, différents éléments botaniques comme par exemple les fleurons du tournesol ou les primordia des pommes de pin se développent sur une spirale d'Archimède. On pourra consulter l'article suivant pour une présentation détaillée du phénomène : La physique des spirales végétales, S. DOUADY et Y. COUDER, *La Recherche*, n° 250, p. 26 (1993).

Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU

4.3 Fonctions complexes d'une variable complexe

4.3.1 Généralités

On appelle *fonction complexe d'une variable complexe* une relation[17] dont l'ensemble de départ Z et l'ensemble d'arrivée W sont des sous-ensembles de $\mathbb{C}$ et telle que tout élément de Z est en relation avec au plus un élément de W.

On appelle *ensemble de définition* de la fonction complexe d'une variable complexe f de Z dans W, le sous-ensemble $\mathcal{D}_f$ de Z constitué des éléments de Z ayant une image par f ; autrement dit

$$\mathcal{D}_f = \{z \in Z \ ; \ \exists w \in W \ w = f(z)\}.$$

Une fonction complexe d'une variable complexe f de Z dans W est qualifiée d'application si son ensemble de définition est l'ensemble de départ ($Z = \mathcal{D}_f$).

Puisqu'il existe plusieurs façons de représenter un nombre complexe, on peut aussi représenter une application complexe f définie sur un sous-ensemble Z de $\mathbb{C}$ de différentes manières. Si z désigne un élément de Z et w son image par f, on peut représenter z et w sous forme cartésienne :

$$\exists (x,y) \in \mathbb{R}^2 \quad z = x + \mathrm{i}\, y \qquad \text{et} \qquad \exists (u,v) \in \mathbb{R}^2 \quad w = u + \mathrm{i}\, v,$$

ce qui nous donne la représentation $f(z) = u(z) + \mathrm{i}\, v(z)$, les réels u et v dépendant de z.

- La fonction $U : z \in \mathbb{C} \mapsto u(z) \in \mathbb{R}$ est appelée *partie réelle de* f et est notée $\mathcal{R}\mathrm{e}(f)$.
- La fonction $V : z \in \mathbb{C} \mapsto v(z) \in \mathbb{R}$ est appelée *partie imaginaire de* f et est notée $\mathcal{I}\mathrm{m}(f)$.

Exemples

1. L'application $f : z \in \mathbb{C} \mapsto z^2 \in \mathbb{C}$ admet pour partie réelle l'application $U : z \in \mathbb{C} \mapsto \mathcal{R}\mathrm{e}(z)^2 - \mathcal{I}\mathrm{m}(z)^2 \in \mathbb{R}$ et pour partie imaginaire l'application $V : z \in \mathbb{C} \mapsto 2\mathcal{R}\mathrm{e}(z)\mathcal{I}\mathrm{m}(z) \in \mathbb{R}$. En effet, si $z = x + \mathrm{i}y$ alors

$$z^2 = (x + \mathrm{i}y)^2 = x^2 - y^2 + 2\mathrm{i}xy.$$

2. L'application $f : z \in \mathbb{C} \mapsto z + \overline{z} \in \mathbb{C}$ admet pour partie réelle l'application $U : z \in \mathbb{C} \mapsto 2\mathcal{R}\mathrm{e}(z) \in \mathbb{R}$ et pour partie imaginaire l'application nulle car $z + \overline{z} = 2\mathcal{R}\mathrm{e}(z)$.

[17] Voir la définition 2.11 p. 33 du *Cours de première année*.

Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU

Surjection, injection et bijection

On appelle *image* de l'application complexe f de Z dans W le sous-ensemble suivant de W :

$$f(Z) = \{w \in \mathbb{C} \ ; \ \exists z \in Z \quad w = f(z)\}.$$

Plus généralement, si A désigne un sous-ensemble de Z alors l'ensemble

$$f(A) = \{w \in \mathbb{C} \ ; \ \exists z \in A \quad w = f(z)\}$$

est appelé *image de A par f*.

L'application complexe f de Z dans W est une surjection si tout élément de W admet au moins un antécédent par f dans Z, autrement dit si :

$$\forall w \in W \quad \exists z \in Z \quad w = f(z).$$

De manière équivalente, la surjectivité de f de Z dans W se traduit par la relation $f(Z) = W$.

L'application complexe f de Z dans W est une injection si tout élément de W admet au plus un antécédent par f dans Z, autrement dit si :

$$\forall (z, z') \in Z^2 \quad (f(z) = f(z') \implies z = z').$$

De manière équivalente, par contraposition, l'application complexe f de Z dans W est une injection si :

$$\forall (z, z') \in Z^2 \quad (z \neq z' \implies f(z) \neq f(z')).$$

Enfin, l'application complexe f de Z dans W est une bijection s'il s'agit à la fois d'une injection et d'une surjection. Cela signifie que tout élément de W admet un unique antécédent par f dans Z.

Parité et périodicité

✗ Soit f une application de $\mathbb{C}$ dans $\mathbb{C}$.

- On dit que f est *paire* si : $\forall z \in \mathbb{C} \quad f(-z) = f(z)$.
- On dit que f est *impaire* si : $\forall z \in \mathbb{C} \quad f(-z) = -f(z)$.

✗ On dit qu'une application f de $\mathbb{C}$ dans $\mathbb{C}$ admet le complexe non nul T pour *période* si

$$\forall z \in \mathbb{C} \quad f(z + T) = f(z).$$

Une fonction admettant une période est dite *périodique* de période T.

Exemples

1. L'application $z \in \mathbb{C} \mapsto z + \overline{z} \in \mathbb{C}$ est impaire.

2. Pour tout $n \in \mathbb{N}$, l'application $z \in \mathbb{C} \mapsto z^{2n} \in \mathbb{C}$ est paire.

3. L'application $z \in \mathbb{C} \mapsto \mathrm{i} \cos(\mathcal{I}\mathrm{m}(z)) \in \mathbb{C}$ est périodique et admet pour période $2\pi\,\mathrm{i}$.

Comment Créer un Compte 1xBet Officiel : https://youtu.be/Eu3fljkkkSU

Application bornée

On dit qu'une application complexe f définie sur un sous-ensemble Z de $\mathbb{C}$ est *bornée* sur Z si

$$\exists K \in \mathbb{R}^+ \quad \forall z \in Z \quad |f(z)| \leqslant K.$$

Une application complexe f définie sur un sous-ensemble Z de $\mathbb{C}$ n'est pas *bornée* sur Z si

$$\forall K \in \mathbb{R}^+ \quad \exists z \in Z \quad |f(z)| > K.$$

Exemples

1. L'application $f : z \in \mathbb{C}^* \mapsto |z|/z \in \mathbb{C}$ est bornée sur $\mathbb{C}^*$ car pour tout $z \in \mathbb{C}^*$ on a $|f(z)| = 1$.

2. L'application $f : z \in \mathbb{C}^* \mapsto 1/z \in \mathbb{C}$ n'est pas bornée sur $\mathbb{C}^*$ car pour tout $K \in \mathbb{R}^+$ le complexe $z = \frac{1}{K+\mathrm{i}}$ est tel que $|f(z)| = |K + \mathrm{i}| > K$.

4.3.2 Quelques exemples de fonctions de la variable complexe

Fonction racine n-ième

Soient $n \in \mathbb{N}$ avec $n \geqslant 2$ et $w \in \mathbb{C}$. Les complexes z solutions de l'équation $z^n = w$ sont les racines n-ièmes de w. Si w est non nul, il existe n solutions qui sont données par la relation[18] :

$$z_k = \sqrt[n]{|w|}\, \mathrm{e}^{\mathrm{i}\frac{1}{n}(\mathrm{Arg}(w)+2k\pi)} \qquad k \in \{0, \ldots, n-1\}.$$

La fonction puissance $z \mapsto z^n$ n'est donc pas une bijection de $\mathbb{C}$ dans $\mathbb{C}$. Par contre, la relation ci-dessus indique qu'il s'agit d'une bijection de l'ensemble $\mathcal{D}_n = \{z \in \mathbb{C}^* \ ; \ \mathrm{Arg}(z) \in]-\frac{\pi}{n}, \frac{\pi}{n}]\}$ dans $\mathbb{C}$.

On appelle *détermination principale de la fonction racine n-ième complexe* l'application :

$$z \in \mathbb{C}^* \longmapsto \sqrt[n]{|z|}\, \mathrm{e}^{\mathrm{i}\frac{1}{n}\mathrm{Arg}(z)} \in \mathcal{D}_n.$$

Fonctions de la variable complexe et géométrie

Intéressons-nous à quelques exemples de fonctions de la variable complexe intervenant en géométrie du plan.

Translation. L'application $\mathcal{T} : z \in \mathbb{C} \mapsto z + z_0 \in \mathbb{C}$ où z_0 désigne un nombre complexe fixé est une bijection appelée *translation*. Un point M d'affixe z a pour image par $\mathcal{T}$ un point M' qui est le translaté de M par la translation de vecteur $\tau = (\mathcal{R}\mathrm{e}(z_0), \mathcal{I}\mathrm{m}(z_0))$, autrement dit un point M' tel que $\mathbf{MM'} = \tau$.

La bijection réciproque de $\mathcal{T}$ est la translation $\mathcal{T}^{-1} : z \in \mathbb{C} \mapsto z - z_0 \in \mathbb{C}$.

Rotation. L'application $\mathcal{R} : z \in \mathbb{C} \mapsto z\,\mathrm{e}^{\mathrm{i}\alpha} \in \mathbb{C}$ où α désigne un nombre réel fixé est une bijection appelée *rotation*. Un point M d'affixe z a pour image par $\mathcal{R}$ un point M' qui s'obtient par la rotation de centre l'origine et d'angle α.

---

[18] Voir la proposition 4.9 p. 149 du *Cours de première année*.

La bijection réciproque de $\mathcal{R}$ est la rotation $\mathcal{R}^{-1} : z \in \mathbb{C} \mapsto z\,\mathrm{e}^{-\mathrm{i}\alpha} \in \mathbb{C}$.

Dilatation. L'application $\mathcal{D} : z \in \mathbb{C} \mapsto kz \in \mathbb{C}$ où k désigne un nombre réel strictement positif fixé est une bijection appelée *dilatation.* Un point M d'affixe z a pour image par $\mathcal{D}$ un point M' tel que $\mathbf{OM'} = k\,\mathbf{OM}$.

La bijection réciproque de $\mathcal{D}$ est la dilatation $\mathcal{D}^{-1} : z \in \mathbb{C} \mapsto k^{-1} z$.

Inversion. L'application $\mathcal{I} : z \in \mathbb{C}^* \mapsto 1/z \in \mathbb{C}^*$ est une bijection appelée *inversion.* Une inversion est sa propre bijection réciproque.

Homographie. L'application $\mathcal{H} : z \in \mathbb{C} \setminus \{-\frac{d}{c}\} \mapsto \dfrac{az+b}{cz+d} \in \mathbb{C}^*$ où a, b, c, d désignent quatre complexes tels que $ad - bc \neq 0$ est appelée homographie.

4.3.3 Représentation graphique d'une fonction de la variable complexe

La représentation graphique d'une fonction complexe de la variable complexe n'est pas aisée car l'ensemble de départ et l'ensemble d'arrivée étant dans le plan complexe, pour disposer d'une représentation graphique analogue à celle que l'on connaît pour les fonctions réelle de la variable réelle, il faudrait pouvoir visualiser un espace de dimension 4. Les représentations graphiques des fonctions complexes de la variable complexe sont donc « partielles » au sens où il faut choisir les aspects à privilégier pour la fonction (partie réelle, imaginaire, module, argument...).

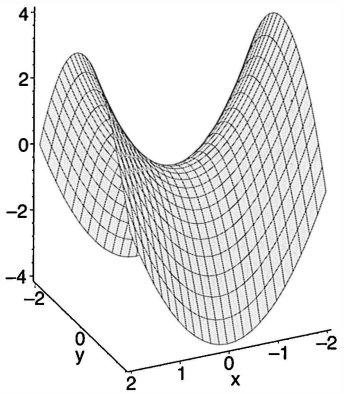
Nous allons illustrer les différentes façons de représenter une fonction complexe de la variable complexe en considérant la fonction $f : z \in \mathbb{C} \mapsto z^2 \in \mathbb{C}$ et en ayant recours à MAPLE.

On peut représenter une fonction complexe de la variable complexe en représentant séparément la partie réelle et la partie imaginaire. On visualise ainsi la fonction par deux surfaces dans l'espace. Si on associe au complexe z le couple (x, y) de $\mathbb{R}^2$ formé de sa partie réelle et de sa partie imaginaire, on représente graphiquement les deux fonctions $f_r : (x, y) \in \mathbb{R}^2 \mapsto \mathcal{R}\mathrm{e}(f(x + \mathrm{i}y))$ et $f_i : (x, y) \in \mathbb{R}^2 \mapsto \mathcal{I}\mathrm{m}(f(x + \mathrm{i}y))$.

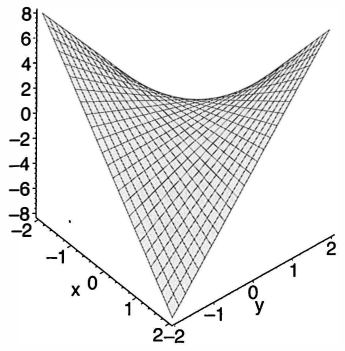
Dans le cas de notre exemple on a $z^2 = (x + \mathrm{i}y)^2 = x^2 - y^2 + 2\mathrm{i}xy$ et on représente simultanément $f_r : (x, y) \in \mathbb{R}^2 \mapsto x^2 - y^2$ et $f_i : (x, y) \in \mathbb{R}^2 \mapsto 2xy$.

```
> f:=z->z^2:
> f_r:=unapply(Re(subs(z=x+I*y,f(z))),x,y):
> f_i:=unapply(Im(subs(z=x+I*y,f(z))),x,y):
> plot3d(f_r(x,y),x=-2..2,y=-2..2,axes=framed);
> plot3d(f_i(x,y),x=-2..2,y=-2..2,axes=framed);
```

Partie réelle



Partie imaginaire

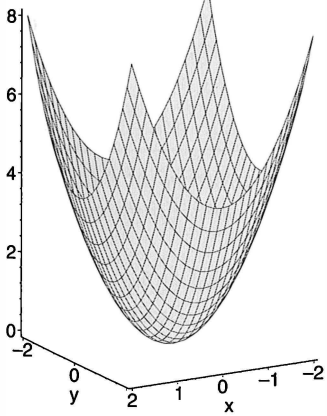


On peut aussi représenter une fonction complexe de la variable complexe en représentant séparément le module et l'argument. On visualise là encore la fonction par deux surfaces dans l'espace. Si on identifie un complexe z au couple (x, y) de $\mathbb{R}^2$ formé de sa partie réelle et de sa partie imaginaire, on représente graphiquement les deux fonctions $f_m : (x, y) \in \mathbb{R}^2 \mapsto |f(x + \mathrm{i}y)|$ et $f_a : (x, y) \in \mathbb{R}^2 \mapsto \mathrm{Arg}(f(x + \mathrm{i}y))$.

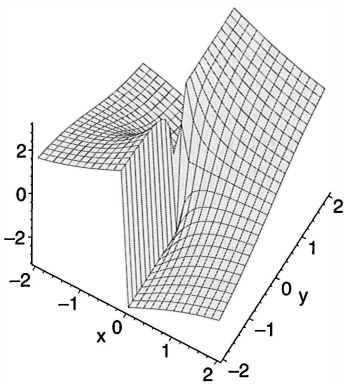
Dans le cas de notre exemple on a $z^2 = |z|^2 \mathrm{e}^{2\mathrm{i}\mathrm{Arg}(z)}$ et on représente simultanément $f_m : (x, y) \in \mathbb{R}^2 \mapsto x^2 + y^2$ et $f_a : (x, y) \in \mathbb{R}^2 \mapsto 2\mathrm{Arg}(x + \mathrm{i}y)$.

```
> f_m:=unapply(abs(subs(z=x+I*y,f(z))),x,y):
> f_a:=unapply(argument(subs(z=x+I*y,f(z))),x,y):
> plot3d(f_m(x,y),x=-2..2,y=-2..2,axes=framed);
> plot3d(f_a(x,y),x=-2..2,y=-2..2,axes=framed);
```

Module



Argument



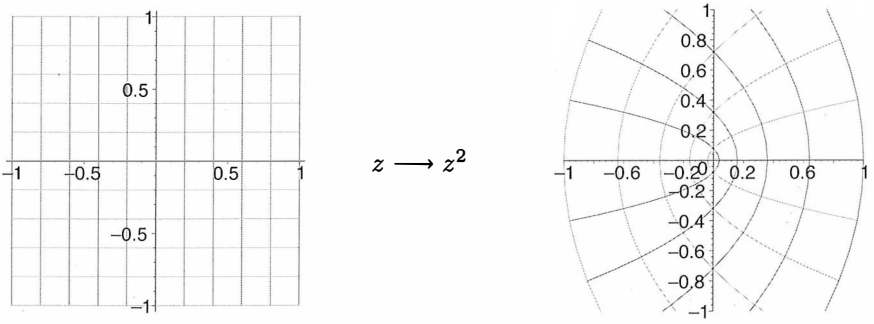
Sous MAPLE, la commande `complexplot3d` trace la représentation graphique du module $f_m : (x, y) \in \mathbb{R}^2 \mapsto |f(x + \mathrm{i}y)|$ en assignant à chaque point $(x, y, f_m(x, y))$ de la surface correspondant à la représentation graphique de f_m

une couleur qui est fonction de l'argument $f_a(x,y) = 2\mathrm{Arg}(x+\mathrm{i}y)$. Cette commande s'utilise de la manière suivante.

```
> complexplot3d(f(z),z=-2-2*I..2+2*I,axes=framed);
```

MAPLE dispose également de la commande `conformal` qui permet de visualiser l'image par f d'un maillage régulier de lignes horizontales et de lignes verticales du plan complexe.

```
> conformal(z,z =-1-I..1+I,scaling=constrained);
> conformal(z^2,z =-1-I..1+I,scaling=constrained,
            view=[-1..1,-1..1]);
```



4.4 Limite d'une fonction de la variable complexe

4.4.1 Définition de la limite

La définition de la limite pour une fonction complexe d'une variable complexe prend une forme très voisine de celle connue pour une fonction réelle d'une variable réelle, voir la définition 13.2 p. 569 du *Cours de première année*. Elle s'énonce de la manière suivante (on sera attentif au fait qu'ici $|\cdot|$ désigne le module d'un nombre complexe).

DÉFINITION 4.10 *Soient U un ouvert non vide de $\mathbb{C}$ et z_0 un point adhérent à U. On dit que l'application f de U dans $\mathbb{C}$ admet $\ell \in \mathbb{C}$ pour limite en z_0 si*

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists \eta \in \mathbb{R}_+^* \quad \forall z \in U \quad \Big(|z - z_0| \leqslant \eta \Longrightarrow |f(z) - \ell| \leqslant \varepsilon\Big).$$

Interprétation

✗ L'application f de U dans $\mathbb{C}$ admet pour limite en z_0 le complexe ℓ si, pour tout réel ε strictement positif, on peut trouver une boule de centre z_0 telle que pour tout complexe z de U appartenant à cette boule, la distance entre les points d'affixe $f(z)$ et ℓ est plus petite que ε, voir la figure 5.

✗Une autre façon d'interpréter la définition 4.10 est la suivante : pour toute boule centrée en ℓ et de rayon ε, on peut trouver une boule $D(z_0, \eta)$ de centre z_0 et de rayon η dont les points qui appartiennent aussi à U, ont tous une image par f appartenant à la boule $D(\ell, \varepsilon)$, *i.e.* telle que $f(D(z_0, \eta) \cap U) \subset D(\ell, \varepsilon)$.

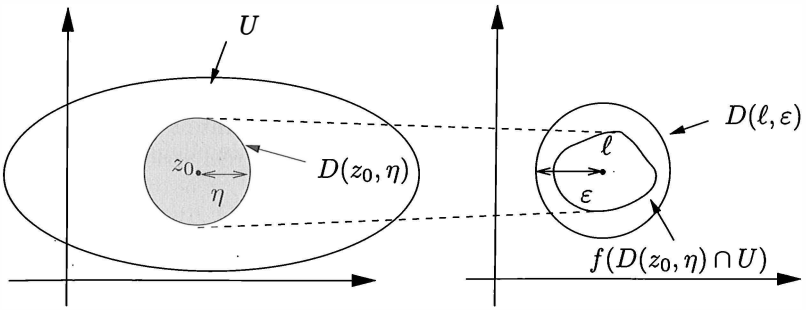


Fig. 5 Interprétation graphique de la définition de la limite.

Remarques

1. L'application f de U dans $\mathbb{C}$ n'admet pas de limite en z_0 si

$$\forall \ell \in \mathbb{C} \quad \exists \varepsilon \in \mathbb{R}_+^* \quad \forall \eta \in \mathbb{R}_+^* \quad \exists z \in U \quad \big(|z - z_0| \leqslant \eta \text{ et } |f(z) - \ell| > \varepsilon\big).$$

Cette assertion est la négation de l'assertion quantifiée donnée à la définition 4.10. Elle s'interprète de la manière suivante : f n'a pas de limite en z_0 si pour tout complexe ℓ, on peut trouver un réel ε strictement positif tel que quel que soit le rayon η de la boule centrée en z_0, au moins un élément de cette boule possède une image par f qui est distante de ℓ de plus que ε.

2. Dans la définition 4.10, on peut prendre des inégalités larges ou des inégalités strictes ; on obtient des définitions équivalentes pour la limite en z_0 de f. Par ailleurs, on peut étendre la notion de limite à une application f définie sur un ensemble U non nécessairement ouvert.

3. Il résulte des relations

$$|\mathcal{R}\mathrm{e}(f(z)) - \mathcal{R}\mathrm{e}(\ell)| = |\mathcal{R}\mathrm{e}(f(z) - \ell)| \leqslant |f(z) - \ell|$$
$$\text{et} \quad |\mathcal{I}\mathrm{m}(f(z)) - \mathcal{I}\mathrm{m}(\ell)| = |\mathcal{I}\mathrm{m}(f(z) - \ell)| \leqslant |f(z) - \ell|$$

que si l'application f de U dans $\mathbb{C}$ admet pour limite en z_0 le complexe ℓ alors les applications $\mathcal{R}\mathrm{e}(f)$ et $\mathcal{I}\mathrm{m}(f)$ de U dans $\mathbb{R}$ admettent respectivement pour limite en z_0 les réels $\mathcal{R}\mathrm{e}(\ell)$ et $\mathcal{I}\mathrm{m}(\ell)$. Réciproquement, puisque

$$|f(z) - \ell|^2 \leqslant (\mathcal{R}\mathrm{e}(f(z)) - \mathcal{R}\mathrm{e}(\ell))^2 + (\mathcal{I}\mathrm{m}(f(z)) - \mathcal{I}\mathrm{m}(\ell))^2,$$

si les applications $\mathcal{R}\mathrm{e}(f)$ et $\mathcal{I}\mathrm{m}(f)$ admettent pour limites respectives en z_0 les réels a et b alors l'application f admet pour limite en z_0 le complexe $a + \mathrm{i}\, b$. ⋄

Exemples

1. L'application $f : z \in \mathbb{C} \mapsto z\,e^{i\alpha} \in \mathbb{C}$ où α désigne un nombre réel fixé (cette application représente une rotation d'angle α de centre O) admet pour limite $\ell = 0$ en $z_0 = 0$. En effet, pour tout réel ε strictement positif, il existe un réel strictement positif η (prendre par exemple $\eta = \varepsilon$) tel que si $|z| \leqslant \eta$ alors

$$|f(z) - 0| = |ze^{i\alpha}| = |z| \leqslant \eta = \varepsilon.$$

2. L'application $f : z \in \mathbb{C}^* \mapsto z^2/|z| \in \mathbb{C}$ admet pour limite $\ell = 0$ en $z_0 = 0$. En effet, pour tout réel ε strictement positif, il existe un réel strictement positif η (prendre par exemple $\eta = \varepsilon$) tel que si $|z| \leqslant \eta$ alors

$$|f(z) - 0| = \left|\frac{z^2}{|z|}\right| = |z| \leqslant \eta = \varepsilon.$$

3. Rappelons que la fonction Arg : $z \in \mathbb{C} \mapsto \mathrm{Arg}(z) \in]-\pi, \pi]$ qui à un complexe associe son argument principal n'est pas définie en $z_0 = 0$. Toutefois, l'application $f : z \in \mathbb{C}^* \mapsto z\,\mathrm{Arg}(z)$ admet pour limite 0 en $z_0 = 0$ car pour tout réel ε strictement positif fixé, il existe un réel strictement positif η (prendre par exemple $\eta = \varepsilon/\pi$) tel que si $|z| \leqslant \eta$ alors

$$|f(z) - 0| = |z\ \mathrm{Arg}(z)| \leqslant \pi|z| \leqslant \pi\eta = \varepsilon.$$

Exercice 4.2 En utilisant la définition 4.10, montrer que :
1 - l'application $z \in \mathbb{C} \mapsto \mathcal{R}\mathrm{e}(z)$ admet pour limite $\mathcal{R}\mathrm{e}(z_0)$ en $z_0 \in \mathbb{C}$;
2 - l'application $z \in \mathbb{C} \mapsto \mathcal{I}\mathrm{m}(z)$ admet pour limite $\mathcal{I}\mathrm{m}(z_0)$ en $z_0 \in \mathbb{C}$;
3 - l'application $z \in \mathbb{C} \mapsto \overline{z}$ admet pour limite $\overline{z_0}$ en $z_0 \in \mathbb{C}$;
4 - l'application $z \in \mathbb{C} \mapsto |z|$ admet pour limite $|z_0|$ en $z_0 \in \mathbb{C}$.

PROPOSITION 4.2 *Soient U un ouvert non vide de $\mathbb{C}$ et z_0 un point adhérent à U. Si l'application f de U dans $\mathbb{C}$ admet une limite en z_0 alors cette limite est unique.*

Démonstration Raisonnons par l'absurde et supposons que l'application f admette en z_0 deux limites ℓ_1 et ℓ_2 distinctes. D'après la définition 4.10 considérée avec $\varepsilon = \frac{1}{3}|\ell_1 - \ell_2|$, on a d'une part,

$$\exists \eta_1 \in \mathbb{R}_+^* \quad \forall z \in U \quad \Big(|z - z_0| \leqslant \eta_1 \Longrightarrow |f(z) - \ell_1| \leqslant \tfrac{1}{3}|\ell_1 - \ell_2|\Big)$$

et d'autre part,

$$\exists \eta_2 \in \mathbb{R}_+^* \quad \forall z \in U \quad \Big(|z - z_0| \leqslant \eta_2 \Longrightarrow |f(z) - \ell_2| \leqslant \tfrac{1}{3}|\ell_1 - \ell_2|\Big).$$

Notons $\eta = \min(\eta_1, \eta_2)$; en utilisant la première inégalité triangulaire, on obtient pour tout $z \in B(z_0, \eta) \cap U$:

$$|\ell_1 - \ell_2| \leqslant |f(z) - \ell_1| + |f(z) - \ell_2| \leqslant \tfrac{2}{3}|\ell_1 - \ell_2|.$$

La relation $|\ell_1 - \ell_2| \leqslant \frac{2}{3}|\ell_1 - \ell_2|$ est impossible si $\ell_1 \neq \ell_2$. On aboutit à une contradiction. Cela signifie que l'hypothèse de deux limites distinctes pour f en z_0 est absurde. On a ainsi démontré que si l'application f admettait une limite en z_0 alors celle-ci était nécessairement unique. □

La limite (lorsqu'elle existe) d'une application complexe f, définie sur un ouvert U de $\mathbb{C}$, en $z_0 \in \overline{U}$ étant unique, on peut introduire un symbole pour désigner cette limite. Comme dans le cas des fonctions de la variable réelle, on désigne par

$$\lim_{\substack{z \to z_0 \\ z \in U}} f(z)$$

la limite de l'application f en $z_0 \in \overline{U}$. Lorsqu'il n'y a pas d'ambiguïté possible sur l'ensemble de départ de l'application f, on écrit plus simplement $\lim\limits_{z \to z_0} f(z)$.

Remarque Il résulte de la définition 4.10 que

$$\lim_{z \to z_0} f(z) = \ell \quad \Longleftrightarrow \quad \lim_{w \to 0} f(z_0 + w) = \ell,$$

et, selon les situations, il peut être plus avantageux de privilégier l'une ou l'autre des deux formes. ◇

4.4.2 Étude de la limite

Contrairement au cas des fonctions réelles de la variable réelle, nous allons voir que dans le cas des fonctions complexes de la variable complexe, il est en général beaucoup plus facile (d'un point de vue pratique) de montrer qu'une fonction n'a pas de limite en un point que de montrer qu'elle en a une.

Un moyen pratique d'établir qu'une fonction n'a pas de limite

Soient f une application complexe définie sur un ouvert U de $\mathbb{C}$ admettant pour limite en $z_0 \in \overline{U}$ le complexe ℓ et γ un arc géométrique de classe $\mathcal{C}^1$ représenté par un arc paramétré $\phi \in \mathcal{C}^1(I, \mathbb{C})$. Supposons que le support de γ contienne le point d'affixe z_0 et que $z_0 = \phi(t_0)$ pour un certain t_0 appartenant à I. Pour tout $z \in \gamma$, il existe $t \in I$ tel que $z = \phi(t)$ et on a :

$$|z - z_0| = |\phi(t) - \phi(t_0)| = \sqrt{\mathcal{R}\mathrm{e}(\phi(t) - \phi(t_0))^2 + \mathcal{I}\mathrm{m}(\phi(t) - \phi(t_0))^2}.$$

Or $|\mathcal{R}\mathrm{e}(\phi(t) - \phi(t_0)| \leqslant |\phi(t) - \phi(t_0)|$, $|\mathcal{I}\mathrm{m}(\phi(t) - \phi(t_0))| \leqslant |\phi(t) - \phi(t_0))|$ et par l'inégalité des accroissements finis (théorème 4.1 p. 163),

$$|\phi(t) - \phi(t_0)| \leqslant M\,|t - t_0|$$

où $M = \sup_{s \in I} |\phi'(s)|$. On a par conséquent : $|z - z_0| \leqslant \sqrt{2} M |t - t_0|$.

Pour tout $\varepsilon \in \mathbb{R}_+^*$, il existe $\eta \in \mathbb{R}_+^*$ tel que, pour tout $z \in U$,

$$|z - z_0| \leqslant \eta \quad \Longrightarrow \quad |f(z) - \ell| \leqslant \varepsilon.$$

On en déduit que pour tout $t \in I$,

$$\left(\phi(t) \in U \text{ et } |t - t_0| \leqslant \tfrac{\eta}{\sqrt{2}M}\right) \quad \Longrightarrow \quad |f(\phi(t)) - \ell| \leqslant \varepsilon.$$

Cela signifie que si f admet pour limite le complexe ℓ en z_0 alors nécessairement la limite de $f \circ \phi$ en t_0 est égale à ℓ.

On dispose ainsi d'un moyen pratique de montrer qu'une fonction f de la variable complexe n'a pas de limite en z_0.

1. On peut chercher un arc géométrique de paramétrisation ϕ ayant z_0 dans son support (*i.e.* tel qu'il existe $t_0 \in \mathbb{R}$ pour lequel $z_0 = \phi(t_0)$) tel que la fonction $f \circ \phi$ n'ait pas de limite en t_0.
2. On peut chercher deux arcs géométriques distincts de paramétrisations respectives ϕ et ψ ayant z_0 dans leurs supports ($\exists t_0 \in \mathbb{R}$ $z_0 = \phi(t_0)$ et $\exists s_0 \in \mathbb{R}$ $z_0 = \psi(s_0)$) tels que la limite de la fonction $f \circ \phi$ en t_0 et la limite de la fonction $f \circ \psi$ en s_0 soient distinctes.

Exemples

1. L'application $f : z \in \mathbb{C}^* \mapsto \operatorname{Arg}(z)$ n'a pas de limite en 0. Considérons l'arc géométrique dont le support est l'axe des imaginaires et dont une paramétrisation est $\phi : t \in \mathbb{R} \mapsto it \in \mathbb{C}$. L'application $f \circ \phi : t \in \mathbb{R}^* \mapsto \operatorname{Arg}(it)$ n'a pas de limite en $t_0 = 0$, car $(f \circ \phi)(t) = -\frac{\pi}{2}$ pour $t < 0$ et $(f \circ \phi)(t) = \frac{\pi}{2}$ pour $t > 0$. On en déduit que f n'a pas de limite en $z_0 = \phi(t_0) = 0$.

2. Intéressons-nous à la limite en 0 de l'application $f : z \in \mathbb{C}^* \mapsto \frac{\mathcal{R}\mathrm{e}(z)\,\mathcal{I}\mathrm{m}(z)}{|z|^2}$. Considérons un premier arc géométrique dont le support est l'axe des réels et dont une paramétrisation est $\phi_1 : t \in \mathbb{R} \mapsto t \in \mathbb{C}$. L'application $f \circ \phi_1 : t \in \mathbb{R}^* \mapsto 0$ admet pour limite 0 en $t_0 = 0$. Considérons un second arc géométrique dont le support est la première bissectrice et dont une paramétrisation est $\phi_2 : s \in \mathbb{R} \mapsto s + \mathrm{i}s \in \mathbb{C}$. L'application $f \circ \phi_2 : s \in \mathbb{R}^* \mapsto \frac{1}{2}$ admet pour limite $\frac{1}{2}$ en $s_0 = 0$. On en conclut que f n'a pas de limite en 0.

Remarque D'un point de vue pratique, lorsque l'on souhaite montrer par cette méthode qu'une application n'a pas de limite en z_0, il est commode et souvent suffisant de considérer simultanément tous les arcs géométriques qui sont des droites passant par z_0. Ces arcs admettent pour paramétrisation l'application $\phi : t \in \mathbb{R} \mapsto z_0 + t\,\mathrm{e}^{\mathrm{i}\theta}$ avec θ fixé dans $[0, 2\pi[$. En utilisant cette méthode, on montre aisément que l'application $f : z \in \mathbb{C}^* \mapsto z^2/|z|^2$ n'a pas de limite en 0. Si l'on considère les droites passant par l'origine de paramétrisation $\phi : t \in \mathbb{R} \mapsto t\,\mathrm{e}^{\mathrm{i}\theta}$, on obtient

$$f(\phi(t)) = \frac{t^2 \mathrm{e}^{2\mathrm{i}\theta}}{|t|^2} = \mathrm{e}^{2\mathrm{i}\theta}.$$

La limite de cette quantité lorsque t tend vers 0 vaut $\mathrm{e}^{2\mathrm{i}\theta}$ et dépend par conséquent de l'arc paramétré choisi. ◇

Exercice 4.3 Soit $z_0 \in \mathbb{C}$; montrer que $f : z \in \mathbb{C} \setminus \{z_0\} \mapsto \dfrac{|z| - |z_0|}{z - z_0}$ n'a pas de limite en z_0.

Comment établir qu'une fonction a une limite

Reprenons la définition : une application complexe f de U dans $\mathbb{C}$ admet pour limite $\ell \in \mathbb{C}$ en $z_0 \in \overline{U}$ si

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists \eta \in \mathbb{R}_+^* \quad \forall z \in U \quad \big(|z - z_0| \leqslant \eta \implies |f(z) - \ell| \leqslant \varepsilon\big).$$

Ainsi, pour montrer que f admet ℓ pour limite en z_0, il faut et il suffit de montrer que l'application $g : z \in U \mapsto |f(z) - \ell| \in \mathbb{R}^+$ admet pour limite 0 en z_0.

Il résulte du théorème d'encadrement que si h est une application définie au voisinage de z_0 qui admet 0 pour limite en z_0 et telle qu'au voisinage de z_0 on ait $0 \leqslant g(z) \leqslant h(z)$ alors g admet pour limite 0 en z_0 (et f admet pour limite ℓ).

D'un point de vue pratique on peut donc procéder ainsi pour établir qu'une fonction f a une limite en z_0.

1. On cherche la valeur de ℓ en considérant un arc géométrique ayant z_0 dans son support et on étudie la limite de f en restriction à cet arc.
2. On majore au voisinage de z_0 la fonction à valeurs réelles $g : z \mapsto |f(z) - \ell|$ par une fonction dont on sait que la limite en z_0 est 0.

Illustrons la méthode sur un exemple.

Exemple Étudions la limite en $z_0 = \mathrm{i}$ de l'application

$$f : z \in \mathbb{C} \setminus \{-\mathrm{i}, \mathrm{i}\} \longmapsto \frac{(|z| - 1)^2}{z^2 + 1}.$$

En considérant l'arc géométrique formé par l'axe des imaginaires et en étudiant la limite de f en i en restriction à cet arc, on obtient que si la limite existe sa valeur est nécessairement 0. Pour tout $z \in \mathbb{C} \setminus \{\mathrm{i}, -\mathrm{i}\}$ on a

$$g(z) = |f(z) - 0| = \left|\frac{(|z| - 1)^2}{z^2 + 1}\right| = \frac{||z| - 1|^2}{|z^2 + 1|}.$$

Or $z^2 + 1 = (z - \mathrm{i})(z + \mathrm{i})$ et d'après la seconde inégalité triangulaire[19] on a : $||z| - 1| = ||z| - |\mathrm{i}|| \leqslant |z - \mathrm{i}|$. On en déduit que

$$|g(z)| \leqslant \frac{|z - \mathrm{i}|^2}{|(z - \mathrm{i})(z + \mathrm{i})|} = \frac{|z - \mathrm{i}|}{|z + \mathrm{i}|}.$$

La fonction $h : z \mapsto |z - \mathrm{i}|/|z + \mathrm{i}|$ admet pour limite 0 en i donc la limite en i de f est 0.

[19] Voir la proposition 4.4 p. 134 du *Cours de première année*.

4.4.3 Propriétés de la limite

Proposition 4.3 *Soient U un ouvert non vide de $\mathbb{C}$ et z_0 un point adhérent à U. Si l'application f de U dans $\mathbb{C}$ admet une limite en z_0 alors il existe un voisinage de z_0 sur lequel f est bornée.*

Démonstration Supposons que f admette le complexe ℓ pour limite en z_0. D'après la définition 4.10 considérée avec $\varepsilon = 1$, il existe un réel strictement positif η_1 tel que pour tout $z \in U$,

$$|z - z_0| \leqslant \eta_1 \Longrightarrow |f(z) - \ell| \leqslant 1.$$

On en déduit, en utilisant l'inégalité triangulaire, que pour $z \in B(z_0, \eta) \cap U$ on a

$$|f(z)| = |f(z) - \ell + \ell| \leqslant |f(z) - \ell| + |\ell| \leqslant 1 + |\ell|.$$

L'application f est donc bornée sur $B(z_0, \eta) \cap U$. □

La proposition 4.3 peut être utilisée afin de montrer qu'une application n'a pas de limite. En effet, il résulte de la proposition 4.3, par contraposition, que s'il n'existe pas de voisinage de z_0 sur lequel l'application f de U dans $\mathbb{C}$ est bornée alors elle n'admet pas de limite en z_0. Autrement dit, si

$$\forall r \in \mathbb{R}_+^* \quad \forall M \in \mathbb{R}^+ \quad \exists z \in U \cap D(z_0, r) \quad |f(z)| > M,$$

alors f n'admet pas de limite en z_0.

Exemple L'application $f : z \in \mathbb{C}^* \mapsto 1/z$ n'est bornée sur aucun voisinage de $z_0 = 0$ puisque pour tout $r \in \mathbb{R}_+^*$ et pour tout $M \in \mathbb{R}^+$, le complexe $z = \frac{1}{2}\min(\frac{1}{M}, r)$ vérifie $z \in D(0, r)$ et $|f(z)| > M$. On en déduit que f n'a pas de limite en 0.

Proposition 4.4 *Soient U un ouvert non vide de $\mathbb{C}$ et z_0 un point adhérent à U. L'application f de U dans $\mathbb{C}$ admet pour limite en z_0 le complexe ℓ si et seulement si pour toute suite complexe $(u_n)_n$ d'éléments de U convergeant vers z_0 la suite de terme général $f(u_n)$ converge vers ℓ.*

Démonstration $\trianglerighteq$ Supposons que f admette pour limite ℓ en z_0 et considérons une suite $(u_n)_n$ convergeant vers z_0 et un réel strictement positif ε. Puisque f admet pour limite ℓ en z_0, d'après la définition 4.10,

$$\exists \eta \in \mathbb{R}_+^* \quad \forall z \in U \quad \big(|z - z_0| \leqslant \eta \Longrightarrow |f(z) - \ell| \leqslant \varepsilon\big)$$

et puisque[20] la suite $(u_n)_n$ converge vers z_0

$$\exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad (n \geqslant N \Longrightarrow |u_n - z_0| \leqslant \eta).$$

[20] Voir la définition 5.1, p. 168 du *Cours de première année* ; on prend pour première variable (ε) de l'assertion quantifiée le réel η défini auparavant.

Pour $n \geqslant N$ on a $|u_n - z_0| \leqslant \eta$ et par conséquent $|f(u_n) - \ell| \leqslant \varepsilon$. Ainsi,

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad (n \geqslant N \Longrightarrow |f(u_n) - \ell| \leqslant \varepsilon),$$

autrement dit la suite de terme général $f(u_n)$ converge vers ℓ.

$\trianglerighteq$ Pour montrer la réciproque, raisonnons par l'absurde. Supposons que pour toute suite $(u_n)_n$ convergeant vers z_0 la suite de terme général $f(u_n)$ converge vers ℓ et que f n'admette pas pour limite ℓ en z_0 et montrons que l'on aboutit à une contradiction. La négation de l'assertion « f admet pour limite ℓ en z_0 » s'écrit :

$$\exists \varepsilon \in \mathbb{R}_+^* \quad \forall \eta \in \mathbb{R}_+^* \quad \exists \zeta_\eta \in U \quad \big(|\zeta_\eta - z_0| \leqslant \eta \ \text{ et } \ |f(\zeta_\eta) - \ell| > \varepsilon\big).$$

Pour cette valeur de ε, en prenant pour η des valeurs de la forme $1/n$ où $n \in \mathbb{N}^*$, on obtient :

$$\forall n \in \mathbb{N} \quad \exists \zeta_n \in U \quad \big(|\zeta_n - z_0| \leqslant \tfrac{1}{n} \quad \text{et} \quad |f(\zeta_n) - \ell| > \varepsilon\big). \tag{1}$$

On déduit de la condition $|\zeta_n - z_0| \leqslant \frac{1}{n}$ que la suite $(\zeta_n)_n$ converge vers z_0. On déduit alors de l'assertion (1) en prenant $n = N$ l'assertion suivante :

$$\exists \varepsilon \in \mathbb{R}_+^* \quad \forall N \in \mathbb{N} \quad \exists n \in \mathbb{N} \quad \big(n \geqslant N \quad \text{et} \quad |f(\zeta_n) - \ell| > \varepsilon\big).$$

Cette assertion constitue la négation[21] de l'assertion quantifiée exprimant que la suite de terme général $f(\zeta_n)$ converge vers ℓ. On dispose donc d'une suite $(\zeta_n)_n$ qui converge vers z_0 mais pour laquelle la suite de terme général $f(\zeta_n)$ ne converge pas vers ℓ. C'est en contradiction avec nos hypothèses. □

Remarque Il est difficile en pratique d'utiliser la proposition 4.4 pour établir qu'une application donnée a une limite en un point z_0 ; cela nécessiterait de considérer toutes les suites complexes convergeant vers z_0 et il y en a beaucoup ! Par contre, la proposition 4.4 fournit un moyen pratique pour montrer qu'une application donnée n'a pas de limite en un point z_0. Il suffit d'exhiber deux suites $(u_n)_n$ et $(v_n)_n$ convergeant toutes deux vers z_0 mais pour lesquelles les deux suites de terme général $f(u_n)$ et $f(v_n)$ convergent vers deux limites distinctes. De la même manière, si on trouve une suite $(u_n)_n$ qui converge vers z_0 mais pour laquelle la suite de terme général $f(u_n)$ diverge, on peut conclure que f n'a pas de limite en z_0. ◇

Les exemples suivants viennent illustrer cette remarque.

Exemples

1. Montrons en utilisant la proposition 4.4 que quel que soit le complexe z_0 l'application

$$f : z \in \mathbb{C} \setminus \{z_0\} \longmapsto \frac{\mathcal{I}\mathrm{m}(z) - \mathcal{I}\mathrm{m}(z_0)}{z - z_0}$$

[21] Pour la définition la divergence d'une suite, voir le *Cours de première année* p. 168.

n'a pas de limite en z_0. Considérons la suite $(u_n)_n$ de terme général $u_n = z_0 + \frac{\mathrm{i}}{n}$. Cette suite converge vers z_0 et on a pour tout $n \in \mathbb{N}^*$,

$$f(u_n) = \frac{\mathcal{I}\mathrm{m}(u_n) - \mathcal{I}\mathrm{m}(z_0)}{u_n - z_0} = -\mathrm{i}.$$

Par conséquent la suite de terme général $f(u_n)$ converge vers $-\mathrm{i}$.

Considérons à présent la suite $(v_n)_n$ de terme général $v_n = z_0 + \frac{1}{n}$. Cette suite converge vers z_0 et on a pour tout $n \in \mathbb{N}^*$,

$$f(v_n) = \frac{\mathcal{I}\mathrm{m}(v_n) - \mathcal{I}\mathrm{m}(z_0)}{v_n - z_0} = 0.$$

Par conséquent, la suite de terme général $f(v_n)$ converge vers 0. On peut affirmer d'après la proposition 4.4 que f n'a pas de limite en z_0.

2. Montrons que quel que soit le complexe z_0 l'application

$$f : z \in \mathbb{C} \setminus \{z_0\} \longmapsto \frac{\overline{z} - \overline{z_0}}{z - z_0}$$

n'a pas de limite en z_0. Considérons la suite $(u_n)_n$ de terme général $u_n = \frac{1}{n}\mathrm{e}^{\mathrm{i}n} + z_0$. Cette suite converge vers z_0 et pour tout $n \in \mathbb{N}^*$ on a,

$$f(u_n) = \frac{\overline{u_n} - \overline{z_0}}{u_n - z_0} = \mathrm{e}^{-2\mathrm{i}n}.$$

Comme $\mathrm{e}^{-2\mathrm{i}n} = \cos(2n) - \mathrm{i}\sin(2n)$ n'a pas de limite quand n tend vers $+\infty$, la suite de terme général $f(u_n)$ diverge. On en déduit que f n'a pas de limite en z_0.

PROPOSITION 4.5 (Propriétés algébriques)
Soient U un ouvert non vide de $\mathbb{C}$ et z_0 un point adhérent à U. Soient f et g deux applications de U dans $\mathbb{C}$ telles que f ait pour limite ℓ en z_0 et g ait pour limite ℓ' en z_0. On a les propriétés suivantes :

1. $\lim\limits_{z \to z_0} |f(z)| = |\ell|$;
2. $\forall \lambda \in \mathbb{C} \quad \lim\limits_{z \to z_0} (\lambda \cdot f)(z) = \lambda \ell$;
3. $\lim\limits_{z \to z_0} (f + g)(z) = \ell + \ell'$;
4. $\lim\limits_{z \to z_0} (f \times g)(z) = \ell\, \ell'$;
5. *si de plus on suppose que $\ell' \neq 0$ alors* $\lim\limits_{z \to z_0} \frac{f}{g}(z) = \frac{\ell}{\ell'}$.

Démonstration $\trianglerighteq$ Supposons que f admette pour limite en z_0 le complexe ℓ ; on a alors

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists \eta_\varepsilon \in \mathbb{R}_+^* \quad \forall z \in U \quad \big(|z - z_0| \leqslant \eta_\varepsilon \implies |f(z) - \ell| \leqslant \varepsilon\big). \qquad (2)$$

D'après la seconde inégalité triangulaire, on a $\big||f(z)| - |\ell|\big| \leqslant |f(z) - \ell|$. On déduit alors de (2) que

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists \eta_\varepsilon \in \mathbb{R}_+^* \quad \forall z \in U \quad \big(|z - z_0| \leqslant \eta_\varepsilon \implies \big||f(z)| - |\ell|\big| \leqslant \varepsilon\big).$$

Autrement dit l'application $|f|$ admet pour limite en z_0 le réel $|\ell|$.

$\trianglerighteq$ Supposons que f admette pour limite ℓ en z_0 et montrons que pour tout $\lambda \in \mathbb{C}$, l'application $\lambda \cdot f$ admet pour limite $\lambda\ell$ en z_0, *i.e.* montrons que

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists \eta \in \mathbb{R}_+^* \quad \forall z \in U \quad \big(|z - z_0| \leqslant \eta \implies |\lambda f(z) - \lambda\ell| \leqslant \varepsilon\big). \qquad (3)$$

Si $\lambda = 0$ alors $|\lambda f(z) - \lambda\ell| = 0 \leqslant \varepsilon$ pour tout $\varepsilon \in \mathbb{R}_+^*$ et l'assertion (3) est évidente. Si $\lambda \in \mathbb{C}^*$, alors pour tout $\varepsilon \in \mathbb{R}_+^*$, on a d'après l'assertion (2)

$$\exists \eta \in \mathbb{R}_+^* \quad \forall z \in U \quad \big(|z - z_0| \leqslant \eta \implies |f(z) - \ell| \leqslant \frac{1}{\lambda}\varepsilon\big).$$

Comme $|\lambda f(z) - \lambda\ell| = |\lambda||f(z) - \ell|$, on en déduit que pour tout $z \in U$ avec $|z - z_0| \leqslant \eta$ on a $|\lambda f(z) - \lambda\ell| \leqslant \varepsilon$ et l'assertion (3) est vraie.

$\trianglerighteq$ Intéressons-nous à la troisième assertion. Soit $\varepsilon \in \mathbb{R}_+^*$ fixé ; puisque f admet pour limite ℓ en z_0 et que g admet pour limite ℓ' en z_0, on a

$$\exists \eta_1 \in \mathbb{R}_+^* \quad \forall z \in U \quad \big(|z - z_0| \leqslant \eta_1 \implies |g(z) - \ell'| \leqslant \tfrac{1}{2}\varepsilon\big),$$

et

$$\exists \eta_2 \in \mathbb{R}_+^* \quad \forall z \in U \quad \big(|z - z_0| \leqslant \eta_2 \implies |f(z) - \ell| \leqslant \tfrac{1}{2}\varepsilon\big).$$

Pour $z \in \overline{D}(z_0, \eta_1)$ on a $\big|f(z) - \ell\big| \leqslant \frac{1}{2}\varepsilon$ et pour $z \in \overline{D}(z_0, \eta_2)$ on a $\big|g(z) - \ell'\big| \leqslant \frac{1}{2}\varepsilon$. Ainsi pour tout réel $z \in U$ tel que $|z - z_0| \leqslant \min\{\eta_1, \eta_2\}$ on a

$$\big|(f(z) + g(z)) - (\ell + \ell')\big| \leqslant \big|f(z) - \ell\big| + \big|g(z) - \ell'\big| \leqslant \tfrac{1}{2}\varepsilon + \tfrac{1}{2}\varepsilon = \varepsilon.$$

On en conclut que si f admet pour limite ℓ en z_0 et g admet pour limite ℓ' en z_0, alors $f + g$ admet pour limite $\ell + \ell'$ en z_0.

$\trianglerighteq$ Les deux dernières assertions plus longues à établir sont admises. □

Proposition 4.6 (Limite de la composée de 2 applications)
Soient U un ouvert non vide de $\mathbb{C}$ et z_0 un point adhérent à U. On considère une application f de U dans $\mathbb{C}$, un ouvert V de $\mathbb{C}$ tel que $f(U) \subset V$ et une application g de V dans $\mathbb{C}$.

Si f admet pour limite ω_0 en z_0 et si g admet pour limite ℓ en ω_0 alors $g \circ f$ admet pour limite ℓ en z_0.

Démonstration Supposons que f admette pour limite ω_0 en z_0, *i.e.* que

$$\forall \varepsilon_1 \in \mathbb{R}_+^* \quad \exists \eta_1 \in \mathbb{R}_+^* \quad \forall z \in U \quad \Big(|z - z_0| \leqslant \eta_1 \Longrightarrow |f(z) - \omega_0| \leqslant \varepsilon_1\Big). \quad (4)$$

et supposons que g admette pour limite ℓ en ω_0, *i.e.* que

$$\forall \varepsilon_2 \in \mathbb{R}_+^* \quad \exists \eta_2 \in \mathbb{R}_+^* \quad \forall \omega \in V \quad \Big(|\omega - \omega_0| \leqslant \eta_2 \Longrightarrow |g(\omega) - \ell| \leqslant \varepsilon_2\Big). \quad (5)$$

Pour montrer que $g \circ f$ admet pour limite ℓ en z_0, il faut montrer que

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists \eta \in \mathbb{R}_+^* \quad \forall z \in U \quad \Big(|z - z_0| \leqslant \eta \Longrightarrow |g(f(z)) - \ell| \leqslant \varepsilon\Big). \quad (6)$$

Soit ε un réel strictement positif fixé, z un élément de U et $\omega = f(z)$ un élément de V. D'après (5) considérée avec $\varepsilon_2 = \varepsilon$, il existe $\eta_2 \in \mathbb{R}_+^*$ tel que si $|\omega - \omega_0| = |f(z) - \omega_0| \leqslant \eta_2$ alors $|g(\omega) - \ell| = |g(f(z)) - \ell| \leqslant \varepsilon$. D'après l'assertion (4) considérée avec $\varepsilon_1 = \eta_2$, il existe $\eta_1 \in \mathbb{R}_+^*$ tel que si $|z - z_0| \leqslant \eta_1$ alors $|f(z) - \omega_0| \leqslant \eta_2$. On en déduit que si $z \in U$ est tel que $|z - z_0| \leqslant \eta_1$ alors $|g(f(z)) - \ell| \leqslant \varepsilon$. L'assertion (6) est démontrée. □

Il a été établi à l'exercice 4.2 que

1. l'application $g_1 : z \in \mathbb{C} \mapsto \mathcal{R}\mathrm{e}(z)$ admet pour limite en ℓ le complexe $\mathcal{R}\mathrm{e}(\ell)$;
2. l'application $g_2 : z \in \mathbb{C} \mapsto \mathcal{I}\mathrm{m}(z)$ admet pour limite en ℓ le complexe $\mathcal{I}\mathrm{m}(\ell)$;
3. l'application $g_3 : z \in \mathbb{C} \mapsto \overline{z}$ admet pour limite en ℓ le complexe $\overline{\ell}$;
4. l'application $g_4 : z \in \mathbb{C} \mapsto |z|$ admet pour limite en ℓ le complexe $|\ell|$.

On déduit donc de la proposition 4.6 le résultat suivant.

Corollaire 4.1 *Soient U un ouvert de $\mathbb{C}$, z_0 un point adhérent à U et f une application de U dans $\mathbb{C}$ admettant pour limite en z_0 le complexe ℓ. On a les propriétés suivantes :*

1. $\lim\limits_{z \to z_0} \mathcal{R}e(f(z)) = \mathcal{R}e(\ell)$;
2. $\lim\limits_{z \to z_0} \mathcal{I}m(f(z)) = \mathcal{I}m(\ell)$;
3. $\lim\limits_{z \to z_0} \overline{f(z)} = \overline{\ell}$;
4. $\lim\limits_{z \to z_0} |f(z)| = |\ell|$.

4.5 Continuité d'une fonction de la variable complexe

4.5.1 Définitions et premières propriétés

DÉFINITION 4.11 *Soient U un ouvert non vide de $\mathbb{C}$, f une application de U dans $\mathbb{C}$ et z_0 un élément de U.*

✗ *On dit que f est* continue en z_0 *si f admet pour limite $f(z_0)$ en z_0, i.e. si :*

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists \eta \in \mathbb{R}_+^* \quad \forall z \in U \quad \big(|z - z_0| \leqslant \eta \Longrightarrow |f(z) - f(z_0)| \leqslant \varepsilon\big). \quad (7)$$

✗ *On dit que f est* continue sur U *si f est continue en z_0 pour tout $z_0 \in U$.*

On note $\mathcal{C}^0(U, \mathbb{C})$ l'ensemble des applications de U dans $\mathbb{C}$ continues sur U.

Exemples

1. Toute application f constante sur $\mathbb{C}$ est continue sur $\mathbb{C}$ puisque pour tous complexes z et z_0, on a $|f(z) - f(z_0)| = 0 \leqslant \varepsilon$ pour tout $\varepsilon \in \mathbb{R}_+^*$.

2. L'application identité est continue sur $\mathbb{C}$ puisque pour tous $z_0 \in \mathbb{C}$ et $\varepsilon \in \mathbb{R}_+^*$, il existe $\eta \in \mathbb{R}_+^*$ (par exemple la valeur $\eta = \varepsilon$ convient) tel que pour tout $z \in \mathbb{C}$ tel que $|z - z_0| \leqslant \eta$ on ait $|f(z) - f(z_0)| = |z - z_0| = \eta \leqslant \varepsilon$.

3. Montrons que l'application $f : z \in \mathbb{C} \mapsto z^2$ est continue sur $\mathbb{C}$. Pour tous complexes z et z_0, on a

$$|f(z) - f(z_0)| = |z^2 - z_0^2| = |z - z_0| \times |z + z_0|.$$

Comme $|z + z_0| \leqslant 2|z_0| + |z + z_0|$, pour que la condition $|z - z_0| \leqslant \eta$ implique que $|f(z) - f(z_0)| \leqslant \varepsilon$ pour tout $\varepsilon \in \mathbb{R}_+^*$, il suffit de choisir $\eta \in \mathbb{R}_+^*$ tel que $\eta(2|z_0| + \eta) \leqslant \varepsilon$ c'est-à-dire $\eta \leqslant \sqrt{|z_0|^2 + \varepsilon} - |z_0|$.

4. Les applications $z \in \mathbb{C} \mapsto \mathcal{R}\mathrm{e}(z)$, $z \in \mathbb{C} \mapsto \mathcal{I}\mathrm{m}(z)$, $z \in \mathbb{C} \mapsto |z|$ et $z \in \mathbb{C} \mapsto \overline{z}$ sont continues sur $\mathbb{C}$, voir l'exercice 4.2.

DÉFINITION 4.12 (Prolongement par continuité)
Soient U un ouvert non vide de $\mathbb{C}$, z_0 un élément de U et f une application de $U \setminus \{z_0\}$ dans $\mathbb{C}$ continue. Si f admet pour limite en z_0 le complexe ℓ, alors l'application $\widetilde{f}$ définie sur U par

$$\widetilde{f}(z) = \begin{cases} f(z) & \text{si } z \in U \setminus \{z_0\} \\ \ell & \text{si } z = z_0 \end{cases}$$

est une application continue sur U appelée prolongement par continuité *de l'application f en z_0.*

Exemple Nous établirons ultérieurement que $f : z \in \mathbb{C}^* \mapsto z^2/|z|$ est une application continue sur $\mathbb{C}^*$. Nous avons montré p. 180 qu'elle admet pour limite 0 en 0. Elle est donc est prolongeable par continuité en 0 et admet pour prolongement par continuité l'application :

$$\widetilde{f} : z \in \mathbb{C} \longmapsto \begin{cases} z^2/|z| & \text{si } z \neq 0 \\ 0 & \text{si } z = 0 \end{cases}.$$

Remarque Il est fréquent de confondre au niveau des notations l'application f et son prolongement par continuité en notant également f ce dernier. ◇

On a le résultat suivant qui découle de la proposition 4.3.

PROPOSITION 4.7 *Soient U un ouvert non vide de $\mathbb{C}$, f une application de U dans $\mathbb{C}$ et z_0 un élément de U. Si f est continue en z_0 alors il existe un voisinage de z_0 sur lequel f est bornée.*

Exemples

1. L'application $z \in \mathbb{C}^* \mapsto 1/z \in \mathbb{C}$ n'est pas continue en 0 car elle n'est pas bornée en 0, voir p. 184.

2. L'application $z \in \mathbb{C}^* \mapsto |z|/z \in \mathbb{C}$ est bornée sur $\mathbb{R}$, voir p. 175, mais elle n'est pas continue en 0 car elle n'a pas de limite en 0, voir l'exercice 4.3 p. 183.

PROPOSITION 4.8 *Soient U un ensemble ouvert non vide de $\mathbb{C}$, f une application de U dans $\mathbb{C}$ et z_0 un élément de U. Si f est continue en z_0 et si $f(z_0) \neq 0$ alors il existe un voisinage de z_0 sur lequel f ne s'annule pas.*

Démonstration L'application f est continue en z_0 donc, d'après la définition 4.11,

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists \eta \in \mathbb{R}_+^* \quad \forall z \in U \quad \big(|z - z_0| \leqslant \eta \implies |f(z) - f(z_0)| \leqslant \varepsilon\big).$$

Prenons $\varepsilon = \frac{1}{2}|f(z_0)|$. On a $\varepsilon > 0$ car $f(z_0) \neq 0$ et il existe $\eta \in \mathbb{R}_+^*$ tel que :

$$\forall z \in U \quad \big(|z - z_0| \leqslant \eta \implies |f(z) - f(z_0)| \leqslant \tfrac{1}{2}|f(z_0)|\big).$$

L'ensemble $\mathcal{V} = U \cap D(z_0, \eta)$ est un voisinage de z_0 et pour $z \in \mathcal{V}$, on obtient en utilisant la seconde inégalité triangulaire :

$$|f(z_0)| - |f(z)| \leqslant \big||f(z_0)| - |f(z)|\big| \leqslant |f(z_0) - f(z)| \leqslant \tfrac{1}{2}|f(z_0)|.$$

On en déduit que pour $z \in \mathcal{V}$ on a $|f(z)| \geqslant \frac{1}{2}|f(z_0)| > 0$, autrement dit que f ne s'annule pas sur $\mathcal{V}$ qui est un voisinage de z_0. □

On a le résultat suivant qui est un corollaire de la proposition 4.4. Il est surtout utile pour montrer qu'une application n'est pas continue en un point.

PROPOSITION 4.9 *Soient U un ouvert non vide de $\mathbb{C}$, f une application de U dans $\mathbb{C}$ et z_0 un élément de U. L'application f est continue en z_0 si et seulement si pour toute suite $(u_n)_n$ d'éléments de U convergeant vers z_0, la suite de terme général $f(u_n)$ converge vers $f(z_0)$.*

4.5.2 Opérations algébriques sur les applications continues

La proposition suivant découle de la proposition 4.5.

PROPOSITION 4.10 *Soient U un ouvert non vide de $\mathbb{C}$, λ un complexe et f et g deux applications de U dans $\mathbb{C}$. Si f et g sont continues en $z_0 \in U$ (resp. sur U) alors :*

1. *$|f|$ est continue en z_0 (resp. sur U) ;*
2. *$f+g$ est continue en z_0 (resp. sur U) ;*
3. *$f \times g$ est continue en z_0 (resp. sur U) ;*
4. *$\lambda \cdot f$ est continue en z_0 (resp. sur U) ;*
5. *si de plus $g(z_0) \neq 0$ (resp. g ne s'annule pas sur U) alors f/g est continue en z_0 (resp. sur U).*

Exemples

1. On a démontré précédemment que l'application identité $z \in \mathbb{C} \mapsto z$ était continue sur $\mathbb{C}$. En utilisant la troisième assertion de la proposition 4.10, on peut montrer par récurrence que pour tout entier n avec $n \geqslant 2$, l'application $z \in \mathbb{C} \mapsto z^n$ est continue sur $\mathbb{C}$. La troisième et la quatrième assertion de la proposition 4.10 permettent alors d'en déduire que toute fonction polynomiale associée à un polynôme de $\mathbb{C}[X]$ est continue sur $\mathbb{C}$.

2. La cinquième des assertions permet d'établir qu'une fonction rationnelle[(22)] de $\mathbb{C}(X)$ est continue en tout complexe qui n'est pas un pôle de la fraction rationnelle.

3. L'application $z \in \mathbb{C}^* \mapsto z^2/|z|$ est continue sur $\mathbb{C}^*$ car les deux applications $f : z \in \mathbb{C} \mapsto z^2$ et $g : z \in \mathbb{C} \mapsto |z|$ sont continues sur $\mathbb{C}$ et g ne s'annule qu'en 0.

Le résultat suivant concernant la composition d'applications continues résulte de la proposition 4.6.

(22) Voir le chap. 7 du *Cours de première année*.

Proposition 4.11 *Soient U un ouvert non vide de $\mathbb{C}$, f une application de U dans $\mathbb{C}$, Z un ouvert de $\mathbb{C}$ tel que $f(U) \subset Z$ et g une application de Z dans $\mathbb{C}$. Si f est continue en $z_0 \in U$ (resp. sur U) et si g est continue en $f(z_0)$ (resp. sur U) alors $g \circ f$ est continue en z_0 (resp. sur U).*

Exemples

1. Les applications $z \in \mathbb{C} \mapsto \mathcal{R}\mathrm{e}(z)$, $z \in \mathbb{C} \mapsto \mathcal{I}\mathrm{m}(z)$, $z \in \mathbb{C} \mapsto |z|$ et $z \in \mathbb{C} \mapsto \overline{z}$ étant continues sur $\mathbb{C}$ (voir l'exercice 4.2), on déduit de la proposition 4.11 que si f est une application continue en $z_0 \in \mathbb{C}$ alors les applications $z \in \mathbb{C} \mapsto \mathcal{R}\mathrm{e}(f(z))$, $z \in \mathbb{C} \mapsto \mathcal{I}\mathrm{m}(f(z))$, $z \in \mathbb{C} \mapsto |f(z)|$ et $z \in \mathbb{C} \mapsto \overline{f(z)}$ sont continues en z_0.

2. Montrons que la fonction Arg qui à un complexe non nul associe son argument principal est continue en $z_0 = 1$. Pour tout $h \in B(0,1)$ on a

$$\mathrm{Arg}(1+h) \in]-\tfrac{\pi}{2}, \tfrac{\pi}{2}[\quad \text{et} \quad \mathrm{Arg}(1+h) = \arctan\left(\frac{\mathcal{I}\mathrm{m}(h)}{1+\mathcal{R}\mathrm{e}(h)}\right).$$

Comme les fonctions $\mathcal{R}\mathrm{e}$ et $\mathcal{I}\mathrm{m}$ sont continues en 0 avec $\mathcal{R}\mathrm{e}(0) = 0$ et $\mathcal{I}\mathrm{m}(0) = 0$, on en déduit que $\lim_{\substack{h \to 0 \\ h \in \mathbb{C}}} \mathrm{Arg}(1+h) = 0$. Finalement, comme $\mathrm{Arg}(1) = 0$ on peut en conclure que Arg est continue en 1.

3. Pour montrer que la fonction Arg n'est pas continue en $z_0 = -1$, on peut considérer l'arc paramétré $\phi : t \in [0, \frac{\pi}{2}] \mapsto -\mathrm{i}\mathrm{e}^{-\mathrm{i}t}$. On a $\mathrm{Arg}(\phi(t)) = -\frac{\pi}{2} - t$, $z_0 = \phi(\frac{\pi}{2})$ et $\lim_{t \to \frac{\pi}{2}} \mathrm{Arg}(\phi(t)) = -\pi$ alors que $\mathrm{Arg}(-1) = \pi$.

On peut sur le même principe vérifier que la fonction Arg n'est continue en aucun complexe $z_0 \in \{z \in \mathbb{C} \ ; \ \mathcal{R}\mathrm{e}(z) \leqslant 0 \ \text{ et } \ \mathcal{I}\mathrm{m}(z) = 0\}$.

4. Montrons à présent que la fonction Arg est continue sur l'ensemble $\mathbb{C}^\ddagger = \mathbb{C} \setminus \{z \in \mathbb{C} \ ; \ \mathcal{R}\mathrm{e}(z) \leqslant 0 \ \text{ et } \ \mathcal{I}\mathrm{m}(z) = 0\}$. Soit $z_0 \in \mathbb{C}^\ddagger$; comme $\mathrm{Arg}(z_0) \in]-\pi, \pi]$ et que Arg est continue en 1, on peut trouver un voisinage V de l'origine tel que pour tout $h \in V$,

$$\mathrm{Arg}\left(\frac{z_0+h}{z_0}\right) + \mathrm{Arg}(z_0) \in]-\pi, \pi] \quad \text{et} \quad \mathrm{Arg}\left(\frac{z_0+h}{z_0}\right) + \mathrm{Arg}(z_0) = \mathrm{Arg}(z_0+h).$$

On en déduit que

$$\lim_{\substack{h \to 0 \\ h \in \mathbb{C}}} \big(\mathrm{Arg}(z_0+h) - \mathrm{Arg}(z_0)\big) = \lim_{\substack{h \to 0 \\ h \in \mathbb{C}}} \mathrm{Arg}\left((z_0+h)/z_0\right) = 0$$

car $\lim_{\substack{h \to 0 \\ h \in \mathbb{C}}} (z_0+h)/z_0 = 1$ et Arg est continue en 1 avec $\mathrm{Arg}(1) = 0$. On en conclut que Arg est continue en tout $z_0 \in \mathbb{C}^\ddagger$.

On a également le résultat suivant concernant la continuité de la composée d'applications continues.

PROPOSITION 4.12 *Soient ϕ une application d'un intervalle $]a,b[$ de $\mathbb{R}$ à valeurs dans $\mathbb{C}$ et f une application d'un ouvert U de $\mathbb{C}$ à valeurs dans $\mathbb{C}$ tels que $\phi(]a,b[) \subset U$. Si ϕ est continue en $t_0 \in]a,b[$ (resp. sur $]a,b[$) et si f est continue en $z_0 = \phi(t_0)$ (resp. sur U) alors $f \circ \phi : t \in]a,b[\mapsto f(\phi(t)) \in \mathbb{C}$ est continue en t_0 (resp. sur $]a,b[$).*

Exercice 4.4 En s'inspirant de la preuve de la proposition 4.6, démontrer la proposition 4.12.

4.6 Dérivabilité d'une fonction de la variable complexe

4.6.1 Fonctions holomorphes et fonctions différentiables

La définition suivante généralise de manière naturelle aux fonctions de la variable complexe la notion de dérivée pour une fonction de la variable réelle.

DÉFINITION 4.13 (Fonction holomorphe)

✗ *Soient U un ouvert non vide de $\mathbb{C}$, z_0 un élément de U et f une application de U dans $\mathbb{C}$. On dit que f est* holomorphe *(ou $\mathbb{C}$-dérivable) en z_0 si la quantité Δf_{z_0} définie par*

$$\Delta f_{z_0}(z) = \frac{f(z) - f(z_0)}{z - z_0}$$

admet une limite quand z tend vers z_0. Cette limite, notée $f'(z_0)$, est appelée dérivée de f en z_0.

✗ *On dit que f est holomorphe sur un ouvert U inclus dans U si pour tout $z \in U$, f est holomorphe en z. On appelle alors* dérivée de f *l'application de U dans $\mathbb{C}$ qui à $z \in U$ associe le complexe $f'(z)$. On note f' cette application.*

Il existe plusieurs manières de reformuler la définition 4.13.

- f est holomorphe en z_0 si la quantité $\frac{1}{h}(f(z_0+h)-f(z_0))$ admet une limite ℓ quand le complexe h tend vers 0. Le complexe ℓ est alors la dérivée de f en z_0.
- f est holomorphe en z_0 s'il existe $\ell \in \mathbb{C}$ tel que

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists \eta \in \mathbb{R}_+^* \quad \forall z \in U \quad \left(|z - z_0| \leqslant \eta \implies \left| \frac{f(z) - f(z_0)}{z - z_0} - \ell \right| \leqslant \varepsilon \right).$$

  Le complexe ℓ est alors la dérivée de f en z_0.

– f est holomorphe en z_0 s'il existe $\ell \in \mathbb{C}$ tel que, pour tout complexe h dans un voisinage de 0, on ait :

$$f(z_0 + h) = f(z_0) + h\ell + |h|\,\varepsilon(h)$$

où ε désigne une fonction de $\mathbb{C}$ dans $\mathbb{C}$ telle que $\lim\limits_{\substack{h \to 0 \\ h \in \mathbb{C}}} \varepsilon(h) = 0$. Le complexe ℓ est alors la dérivée de f en z_0.

Remarque Une application de $\mathbb{C}$ dans $\mathbb{C}$ holomorphe sur tout $\mathbb{C}$ est appelée *fonction entière.* ◇

Exemples

1. Une application constante sur $\mathbb{C}$ est holomorphe sur $\mathbb{C}$, à dérivée nulle puisque pour tout $z_0 \in \mathbb{C}$ on a $\Delta_{z_0}(z) = 0$.

2. L'application identité $z \in \mathbb{C} \mapsto z$ est holomorphe sur $\mathbb{C}$ et admet pour dérivée la fonction constante égale à 1 puisque pour tout $z_0 \in \mathbb{C}$ on a $\Delta_{z_0}(z) = \frac{z - z_0}{z - z_0} = 1$.

3. L'application $z \in \mathbb{C} \mapsto z^2$ est holomorphe sur $\mathbb{C}$. Sa dérivée est l'application $z \in \mathbb{C} \mapsto 2z$. En effet, étant donné un nombre complexe z_0, pour tout $z \in \mathbb{C} \setminus \{z_0\}$, on a

$$\Delta_{z_0}(z) = \frac{z^2 - z_0^2}{z - z_0} = z + z_0 \quad \text{d'où} \quad \lim_{\substack{z \to z_0 \\ z \in \mathbb{C}}} \Delta_{z_0}(z) = 2z_0.$$

4. L'application $z \in \mathbb{C}^* \mapsto 1/z$ est holomorphe sur $\mathbb{C}^*$ et sa dérivée est l'application $z \in \mathbb{C}^* \mapsto -1/z^2$. En effet, si z_0 désigne un complexe non nul alors pour tout $z \in \mathbb{C}^* \setminus \{z_0\}$ on a

$$\Delta_{z_0}(z) = \frac{\frac{1}{z} - \frac{1}{z_0}}{z - z_0} = -\frac{1}{zz_0} \quad \text{d'où} \quad \lim_{\substack{z \to z_0 \\ z \in \mathbb{C}^*}} \Delta_{z_0}(z) = -\frac{1}{z_0^2}.$$

5. L'application $z \in \mathbb{C} \mapsto \mathcal{I}\mathrm{m}(z)$ n'est holomorphe en aucun complexe $z_0 \in \mathbb{C}$. En effet, nous avons montré p. 185 que pour tout $z_0 \in \mathbb{C}$ la quantité

$$\Delta_{z_0}(z) = \frac{\mathcal{I}\mathrm{m}(z) - \mathcal{I}\mathrm{m}(z_0)}{z - z_0}$$

n'avait pas de limite lorsque z tendait vers z_0.

Exercice 4.5
1 - Montrer que pour tout $n \in \mathbb{N}^*$, l'application $f : z \in \mathbb{C} \mapsto z^n$ est holomorphe sur $\mathbb{C}$ et déterminer sa dérivée.
2 - Montrer que l'application $z \in \mathbb{C} \mapsto |z|$ est holomorphe nulle part.

Même si la définition 4.13 est une généralisation naturelle de la notion de dérivabilité introduite dans le cas des fonctions réelles de la variable réelle, la notion de dérivabilité pour les fonctions complexes de la variable complexe est assez différente de celle des fonctions réelles de la variable réelle. Pour en avoir une illustration, il suffit de considérer l'application $z \in \mathbb{C} \mapsto |z|$ qui n'est holomorphe en aucun $z \in \mathbb{C}$ alors que l'application $x \in \mathbb{R} \mapsto |x|$ est dérivable en tout $x \in \mathbb{R}^*$.

Dans le but de comprendre les raisons du comportement très différent vis-à-vis de la dérivabilité de ces deux applications, nous allons introduire la notion de différentielle pour une fonction de la variable complexe.

Différentielle d'une fonction de la variable complexe

Une application L de $\mathbb{C}$ dans $\mathbb{C}$ est dite $\mathbb{C}$-*linéaire* si pour tout $(z_1, z_2) \in \mathbb{C}^2$ et pour tout $(\lambda, \mu) \in \mathbb{C}^2$ on a

$$L(\lambda z_1 + \mu z_2) = \lambda L(z_1) + \mu L(z_2).$$

On montre qu'une application L de $\mathbb{C}$ dans $\mathbb{C}$ qui est $\mathbb{C}$-linéaire est de la forme :

$$L(z) = \beta z$$

avec $\beta \in \mathbb{C}$. On définit aussi la notion d'application complexe $\mathbb{R}$-linéaire : une application L de $\mathbb{C}$ dans $\mathbb{C}$ est $\mathbb{R}$-linéaire si pour tout $(z_1, z_2) \in \mathbb{C}^2$ et pour tout $(\lambda, \mu) \in \mathbb{R}^2$, on a

$$L(\lambda z_1 + \mu z_2) = \lambda L(z_1) + \mu L(z_2).$$

On ne confondra pas cette définition avec celle d'application $\mathbb{C}$-linéaire. Il est facile de vérifier que toute application $\mathbb{C}$-linéaire est $\mathbb{R}$-linéaire mais que la réciproque est fausse.

La bijection $z \in \mathbb{C} \mapsto (\mathcal{R}\mathrm{e}(z), \mathcal{I}\mathrm{m}(z)) \in \mathbb{R}^2$ permet d'identifier $\mathbb{C}$ à $\mathbb{R}^2$ et une fonction de $\mathbb{C}$ dans $\mathbb{C}$ peut être vue comme un fonction de $\mathbb{R}^2$ dans $\mathbb{R}^2$. Une application $\mathbb{R}$-linéaire de $\mathbb{C}$ dans $\mathbb{C}$ est alors identifiable à une application linéaire de $\mathbb{R}^2$ dans $\mathbb{R}^2$ de la manière suivante :

$$L : (x, y) \in \mathbb{R}^2 \longmapsto \big(\alpha_1 x + \alpha_2 y, \alpha_3 x + \alpha_4 y\big).$$

Ainsi, une application L de $\mathbb{C}$ dans $\mathbb{C}$ qui est $\mathbb{R}$-linéaire est de la forme :

$$L : z \in \mathbb{C} \mapsto (\alpha_1 \mathcal{R}\mathrm{e}(z) + \alpha_2 \mathcal{I}\mathrm{m}(z)) + \mathrm{i}(\alpha_3 \mathcal{R}\mathrm{e}(z) + \alpha_4 \mathcal{I}\mathrm{m}(z))$$

où $(\alpha_1, \alpha_2, \alpha_3, \alpha_4) \in \mathbb{R}^4$.

Cette distinction entre applications $\mathbb{C}$-linéaires et $\mathbb{R}$-linéaires qui à première lecture peut paraître assez peu importante est pourtant essentielle dans le cadre de la théorie des fonctions holomorphes. Elle permet d'expliquer pourquoi les fonctions holomorphes possèdent des propriétés très particulières que n'ont pas les fonctions de la variable réelle. Citons à titre d'exemple le fait que si une

fonction de la variable complexe est une fois $\mathbb{C}$-dérivable (autrement dit holomorphe) sur un ouvert, alors elle admet des dérivées à tout ordre sur cet ouvert.

Disposant de la notion d'application $\mathbb{R}$-linéaire, nous allons pouvoir introduire la notion de différentielle pour une fonction de $\mathbb{C}$ dans $\mathbb{C}$. Signalons que cette notion a un caractère très général qui sera étudié en détail au chap. 10.

DÉFINITION 4.14 (Fonction différentiable)
Soient U un ouvert non vide de $\mathbb{C}$ et f une application de U dans $\mathbb{C}$. On dit que f est différentiable en $z_0 \in U$ s'il existe une application $\mathbb{R}$-linéaire L de $\mathbb{C}$ dans $\mathbb{C}$ telle que pour tout complexe h dans un voisinage de 0 on ait

$$f(z_0+h)-f(z_0)=L(h)+o_0(|h|).$$

L'application $\mathbb{R}$-linéaire L est notée $\mathrm{d}f_{z_0}$ et est appelée différentielle de f en z_0.

L'application $\mathbb{R}$-linéaire L dans la définition 4.14 est unique. On le démontre en supposant qu'il existe deux applications $\mathbb{R}$-linéaire L_1 et L_2 telles que pour tout complexe h dans un voisinage de 0 on ait

$$f(z_0+h)-f(z_0)=L_1(h)+o_0(|h|) \quad \text{et} \quad f(z_0+h)-f(z_0)=L_2(h)+o_0(|h|).$$

Par différence, l'application $L=L_1-L_2$ vérifie $L(h)=o_0(|h|)$, autrement dit il existe une fonction ε de $\mathbb{C}$ dans $\mathbb{C}$ tel que $L(h)=|h|\varepsilon(h)$ avec $\lim_{h\to 0}\varepsilon(h)=0$. Pour tout $t\in\mathbb{R}^*$, on a

$$\frac{1}{t}L(th)=\frac{1}{t}|th|\,\varepsilon(th)=\frac{|t|}{t}|h|\,\varepsilon(th).$$

Comme l'application $t\in\mathbb{R}^*\mapsto |t|/t$ est bornée, on en déduit que pour tout $h\in\mathbb{R}$ on a $\lim_{t\to 0}\frac{1}{t}L(th)=0$. Or, par linéarité de l'application L, on a $\frac{1}{t}L(th)=L(h)$. On en déduit que $L(h)=0$ pour tout $h\in\mathbb{R}$, autrement dit que $L=0$. On a donc $L_1=L_2$.

Si f est une application de U dans $\mathbb{C}$ qui est différentiable en tout $z\in U$, on appelle *différentielle de f*, et on note $\mathrm{d}f$, l'application de U dans l'ensemble $\mathcal{L}_{\mathbb{R}}(\mathbb{C},\mathbb{C})$ des applications $\mathbb{R}$-linéaire de $\mathbb{C}$ dans $\mathbb{C}$ définie par :

$$\mathrm{d}f : z\in U \longmapsto \mathrm{d}f_z\in\mathcal{L}_{\mathbb{R}}(\mathbb{C},\mathbb{C}).$$

Remarque De manière équivalente, f est différentiable en $z_0\in U$ s'il existe une application $\mathbb{R}$-linéaire L de $\mathbb{C}$ dans $\mathbb{C}$ telle que

$$\lim_{\substack{h\to 0\\ h\in\mathbb{C}^*}}\frac{f(z_0+h)-f(z_0)-L(h)}{|h|}=0.$$

On pourra comparer cette expression à celle de la définition 4.13. ◇

D'un point de vue pratique, comme cela est illustré dans les exemples qui suivent, pour étudier la différentiabilité d'une fonction complexe de la variable complexe f en $z_0 \in \mathbb{C}$, on forme la différence $f(z_0+h)-f(z_0)$ et on cherche en général à isoler dans cette expression une partie correspondant à une application $\mathbb{R}$-linéaire de $\mathbb{C}$ dans $\mathbb{C}$ et un terme complémentaire, non linéaire, dont il faut établir qu'il est négligeable au voisinage de 0 devant $|h|$.

Exemples

1. L'application $f : z \in \mathbb{C} \mapsto \mathcal{I}\mathrm{m}(z)$ est différentiable sur $\mathbb{C}$ puisque pour tous complexes z_0 et h on a

$$f(z_0+h)-f(z_0) = \mathcal{I}\mathrm{m}(h)$$

et l'application $L : h \in \mathbb{C} \mapsto \mathcal{I}\mathrm{m}(h)$ est $\mathbb{R}$-linéaire (on a $(\alpha_1,\alpha_2,\alpha_3,\alpha_4) = (0,0,0,1)$). On remarquera que la différentielle en $z_0 \in \mathbb{C}$ est l'application f elle-même : $\mathrm{d}f : z \in \mathbb{C} \mapsto f$. On dispose là d'un exemple d'application qui est différentiable sur $\mathbb{C}$ mais qui est holomorphe nulle part, voir p. 194.

2. Montrons que l'application $f : z \in \mathbb{C} \mapsto |z|$ est différentiable sur $\mathbb{C}^*$. Soit z_0 un complexe non nul de partie réelle x_0 et de partie imaginaire y_0 ; pour tout complexe $h = h_1 + \mathrm{i}h_2$ non nul on a

$$\begin{aligned} f(z_0+h)-f(z_0) &= \sqrt{(x_0+h_1)^2+(y_0+h_2)^2}-\sqrt{x_0^2+y_0^2} \\ &= \frac{2x_0h_1+2y_0h_2+h_1^2+h_2^2}{\sqrt{(x_0+h_1)^2+(y_0+h_2)^2}+\sqrt{x_0^2+y_0^2}}. \end{aligned}$$

On a une décomposition de la forme $f(z_0+h)-f(z_0) = L(h)+\phi(h)$ en posant $\phi(h) = \dfrac{h_1^2+h_2^2}{|z_0+h|+|z_0|} + \dfrac{2x_0h_1+2y_0h_2}{|z_0+h|+|z_0|} - \dfrac{x_0h_1+y_0h_2}{|z_0|}$ et $L(h) = \dfrac{x_0h_1+y_0h_2}{\sqrt{x_0^2+y_0^2}}$.

On vérifie que $L : h \in \mathbb{C} \mapsto L(h)$ est $\mathbb{R}$-linéaire (on a $\alpha_1 = \mathcal{R}\mathrm{e}(z_0)/|z_0|$, $\alpha_2 = \mathcal{I}\mathrm{m}(z_0)/|z_0|$ et $\alpha_3 = \alpha_4 = 0$) et que $\lim\limits_{h\to 0} \phi(h)/|h| = 0$. Ainsi, pour tout $z \in \mathbb{C}^*$ on a

$$\mathrm{d}f_{z_0} : h \in \mathbb{C} \longmapsto \frac{\mathcal{R}\mathrm{e}(z_0)\mathcal{R}\mathrm{e}(h)+\mathcal{I}\mathrm{m}(z_0)\mathcal{I}\mathrm{m}(h)}{|z_0|}.$$

3. L'application $f : z \in \mathbb{C} \mapsto |z|$ n'est pas différentiable en 0. En effet si tel était le cas, il existerait une application linéaire L telle que pour tout complexe h dans un voisinage de 0 on ait $|h| = L(h)+o_0(|h|)$. On aurait alors également $|-h| = L(-h)+o_0(|-h|)$ autrement dit $|h| = -L(h)+o_0(|h|)$, ce qui en sommant les deux relations impliquerait que $2|h| = o_0(|h|)$. Cela constitue une contradiction : $2|h|$ n'est pas négligeable devant $|h|$.

La proposition suivante indique que toute fonction holomorphe est différentiable. On a déjà vu avec l'application $z \in \mathbb{C} \mapsto \mathcal{I}\mathrm{m}(z)$, qui est différentiable sur $\mathbb{C}$ mais qui n'est holomorphe nulle part, que la réciproque est fausse. Les

deux notions différentiable et $\mathbb{C}$-dérivable (*i.e.* holomorphe) ne sont donc pas équivalentes.

PROPOSITION 4.13 *Soient U un ouvert non vide de $\mathbb{C}$ et f une application de U dans $\mathbb{C}$. Si f est holomorphe en $z_0 \in U$ (resp. sur U) alors f est différentiable en z_0 (resp. sur U).*

Démonstration Pour tout $z \in U \setminus \{z_0\}$ on a $f(z) = f(z_0) + (z - z_0)\, g(z)$ où

$$g : z \in U \setminus \{z_0\} \longmapsto \frac{f(z) - f(z_0)}{z - z_0}.$$

Il en résulte que pour tout complexe h non nul dans un voisinage de 0 on a

$$f(z_0 + h) = f(z_0) + h\, g(z_0 + h) = f(z_0) + h\, f'(z_0) + \phi(h)$$

où $\phi : h \mapsto h(g(z_0 + h) - f'(z_0))$. Si f est holomorphe en z_0, alors

$$\lim_{h \to 0} g(z_0 + h) = f'(z_0) \quad \text{et par conséquent} \quad \lim_{h \to 0} \frac{h\, \phi(h)}{|h|} = 0.$$

On en conclut que si f est holomorphe en z_0 alors f est différentiable en z_0 et $df_{z_0} : h \in \mathbb{C} \mapsto h\, f'(z_0)$. □

Dans la suite de ce chapitre nous ne considérerons plus que la notion de fonction holomorphe. La notion de fonction différentiable sera approfondie et généralisée au chap. 10.

4.6.2 Propriétés des fonctions holomorphes

PROPOSITION 4.14 *Soient U un ouvert non vide de $\mathbb{C}$ et f une application de U dans $\mathbb{C}$. Si f est holomorphe en $z_0 \in U$ (resp. sur U) alors f est continue en z_0 (resp. sur U).*

Démonstration Pour tout élément z de U différent de z_0, on a (vérifier que le membre de droite se simplifie) :

$$f(z) = f(z_0) + (z - z_0)\, \frac{f(z) - f(z_0)}{z - z_0}.$$

Or, f étant holomorphe en z_0 on a $\lim_{z \to z_0} \dfrac{f(z) - f(z_0)}{z - z_0} = f'(z_0)$ et par conséquent

$$\lim_{z \to z_0} (z - z_0)\, \frac{f(z_0 + h) - f(z_0)}{z - z_0} = 0.$$

On en conclut que $\lim_{z \to z_0} f(z) = f(z_0)$, autrement dit f est continue en z_0. □

PROPOSITION 4.15 *Soient U un ouvert non vide de $\mathbb{C}$, λ un complexe et f et g deux applications de U dans $\mathbb{C}$. Si f et g sont holomorphes en $z_0 \in U$ (resp. sur U) alors,*

✗ *$f+g$ est holomorphe en z_0 (resp. sur U) et*

$$(f+g)'(z_0) = f'(z_0) + g'(z_0);$$

✗ *la fonction $\lambda \cdot f$ est holomorphe en z_0 (resp. sur U) et*

$$(\lambda \cdot f)'(z_0) = \lambda\, f'(z_0);$$

✗ *$f \times g$ est holomorphe en z_0 (resp. sur U) et*

$$(f \times g)'(z_0) = f'(z_0)\, g(z_0) + f(z_0)\, g'(z_0);$$

✗ *si de plus $g(z_0) \neq 0$ (resp. g ne s'annule pas sur U), alors $\dfrac{f}{g}$ est holomorphe en z_0 (resp. sur U) et*

$$\left(\frac{f}{g}\right)'(z_0) = \frac{f'(z_0)\, g(z_0) - f(z_0)\, g'(z_0)}{g(z_0)^2}.$$

Démonstration Ces relations se démontrent à partir de la définition 4.13 en utilisant les propriétés des limites. Démontrons la dernière relation ; les autres vérifications, plus simples, sont laissées en exercice. On a d'après la définition 4.13,

$$\left(\frac{f}{g}\right)'(z_0) = \lim_{z \to z_0} \frac{\frac{f}{g}(z) - \frac{f}{g}(z_0)}{z - z_0}.$$

Or,

$$\begin{aligned}
\frac{\frac{f}{g}(z) - \frac{f}{g}(z_0)}{z - z_0} &= \frac{1}{z - z_0} \frac{g(z_0)f(z) - g(z)f(z_0)}{g(z)g(z_0)} \\
&= \frac{1}{z - z_0} \frac{g(z_0)(f(z) - f(z_0)) + f(z_0)(g(z_0) - g(z))}{g(z)g(z_0)} \\
&= \frac{1}{g(z)} \frac{f(z) - f(z_0)}{z - z_0} - \frac{f(z_0)}{g(z)g(z_0)} \frac{g(z_0) - g(z)}{z_0 - z}.
\end{aligned}$$

Étudions la limite du premier terme du membre de droite. Comme g est holomorphe en z_0, g est continue en z_0 et puisque $g(z_0) \neq 0$, l'application $1/g$ est continue en z_0 (voir la proposition 4.10, p. 191). On a donc,

$$\lim_{z \to z_0} \frac{1}{g(z)} \frac{f(z) - f(z_0)}{z - z_0} = \lim_{z \to z_0} \frac{1}{g(z)} \times \lim_{z \to z_0} \frac{f(z) - f(z_0)}{z - z_0} = \frac{1}{g(z_0)}\, f'(z_0).$$

Étudions la limite du second terme du membre de droite. On a,

$$\lim_{z\to z_0} \frac{f(z_0)}{g(z)g(z_0)} \frac{g(z_0)-g(z)}{z_0-z} = \lim_{z\to z_0} \frac{f(z_0)}{g(z)g(z_0)} \times \lim_{z\to z_0} \frac{g(z_0)-g(z)}{z_0-z}$$
$$= \frac{f(z_0)}{g(z_0)^2}\, g'(z_0)$$

car $1/g$ est continue en z_0. On obtient finalement :

$$\left(\frac{f}{g}\right)'(z_0) = \lim_{z\to z_0} \frac{\frac{f}{g}(z)-\frac{f}{g}(z_0)}{z-z_0} = \frac{1}{g(z_0)}\, f'(z_0) - \frac{f(z_0)}{g(z_0)^2}\, g'(z_0)$$
$$= \frac{f'(z_0)\, g(z_0) - f(z_0)\, g'(z_0)}{g(z_0)^2}.$$

La relation est démontrée. □

Exemples

1. On a démontré précédemment que l'application $z \in \mathbb{C} \mapsto z$ est holomorphe sur $\mathbb{C}$. En utilisant la troisième assertion de la proposition 4.15, on peut montrer par récurrence que pour tout entier n avec $n \geqslant 2$, l'application $z \in \mathbb{C} \mapsto z^n$ est holomorphe sur $\mathbb{C}$ de dérivée l'application $z \in \mathbb{C} \mapsto nz^{n-1}$. La troisième et la quatrième assertion de la proposition 4.15 permettent alors d'en déduire que toute fonction polynomiale associée à un polynôme de $\mathbb{C}[X]$ est holomorphe sur $\mathbb{C}$.

2. La cinquième des assertions permet d'établir qu'une fonction rationnelle[23] de $\mathbb{C}(X)$ est holomorphe en tout complexe qui n'est pas un pôle de la fraction rationnelle.

La proposition suivante que nous admettrons indique que la composée de deux fonctions holomorphes est holomorphe.

PROPOSITION 4.16 *Soient U un ouvert non vide de $\mathbb{C}$, f une application de U dans $\mathbb{C}$, Z un ouvert de $\mathbb{C}$ telle que $f(U) \subset Z$ et g une application de Z dans $\mathbb{C}$. Si f est holomorphe en $z_0 \in U$ et si g est holomorphe en $f(z_0)$ alors l'application composée $g \circ f$ est holomorphe en z_0 et*

$$(g \circ f)'(z_0) = f'(z_0) \times g'(f(z_0)).$$

On a également le résultat suivant.

[23] Voir le chap. 7 du *Cours de première année*.

PROPOSITION 4.17 *Soient ϕ une application de $]a,b[\subset \mathbb{R}$ dans $\mathbb{C}$, U un ouvert de $\mathbb{C}$ tels que $\phi(]a,b[) \subset U$ et f une application de U dans $\mathbb{C}$.*

Si ϕ est dérivable[24] *en $t_0 \in]a,b[$ (resp. sur $]a,b[$) et si f est holomorphe en $z_0 = \phi(t_0)$ (resp. sur U) alors $f \circ \phi : t \in]a,b[\mapsto f(\phi(t)) \in \mathbb{C}$ est dérivable*[24] *en t_0 (resp. sur $]a,b[$) et*

$$(f \circ \phi)'(t) = \phi'(t) \times f'(\phi(t)).$$

Nous allons établir un théorème des accroissements finis dans le cas des fonctions complexes de la variable complexe. Comme cela a déjà été mentionné dans le cas des fonctions complexes de la variable réelle (théorème 4.1 p. 163), le théorème des accroissement finis ne peut être énoncé sous forme d'égalité comme c'est le cas pour les fonctions réelles de la variable réelle.

Étant donné deux complexes a et b, on appelle *segment d'extrémités a et b*, l'ensemble, noté $[a,b]$, défini par

$$[a,b] = \{ta + (1-t)b \ ; \ t \in [0,1]\} = \{(1-s)a + sb \ ; \ s \in [0,1]\}.$$

THÉORÈME 4.2 (Inégalité des accroissements finis)
Soit f une application holomorphe sur un ouvert U non vide de $\mathbb{C}$, a et b deux complexes tels que le segment $[a,b]$ soit inclus dans U. On a

$$|f(b) - f(a)| \leqslant |b-a| \sup_{z \in [a,b]} |f'(z)|.$$

Démonstration Nous admettons l'existence de la quantité $\sup_{z\in[a,b]} |f'(z)|$. Considérons l'application $\phi : t \in [0,1] \mapsto f((1-t)a + tb) \in \mathbb{C}$. Comme f est holomorphe sur U, elle est continue sur U et comme par hypothèse le segment $[a,b]$ est inclus dans U, ϕ est continue sur $[0,1]$ et dérivable sur $]0,1[$. Pour tout $t \in]0,1[$, on a $\phi'(t) = (b-a)f'((1-t)a + tb)$ et, en utilisant le théorème 4.1, on obtient :

$$|\phi(1) - \phi(0)| \leqslant \sup_{t\in[0,1]} |(b-a)f'((1-t)a + tb)|.$$

On a $\phi(1) = f(b)$, $\phi(0) = f(a)$ et

$$\sup_{t\in[0,1]} |(b-a)f'((1-t)a+tb)| = |b-a| \sup_{t\in[0,1]} |f'((1-t)a+tb)| = |b-a| \sup_{z\in[a,b]} |f'(z)|,$$

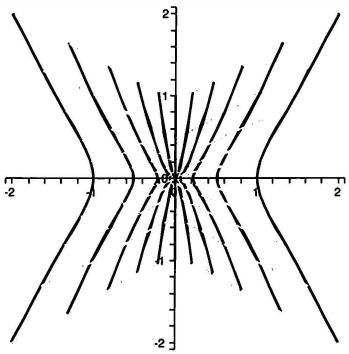
ce qui établit l'inégalité des accroissements finis. □

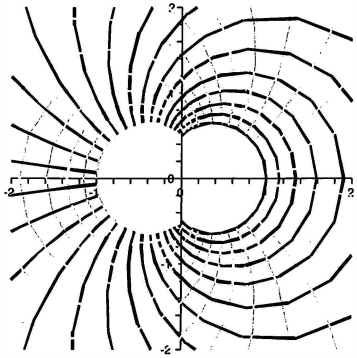
[24] Voir p. 161 pour la notion de dérivabilité pour une application de $\mathbb{R}$ dans $\mathbb{C}$.

Les fonctions holomorphes possèdent la propriété de conserver les angles (on dit qu'il s'agit de *transformations conformes*). Plus précisément, soit f une application holomorphe sur un ouvert U et z_0 un point de U tel que $f'(z_0) \neq 0$. Si γ_1 et γ_2 sont deux arcs géométriques admettant une « tangente » en z_0, alors l'angle orienté existant entre les deux droites tangentes est le même que celui existant entre les deux droites tangentes en $f(z_0)$ aux deux arcs géométriques images par f de γ_1 et γ_2.

Par exemple, l'application $z \in \mathbb{C} \mapsto z\,|z|^2$ est holomorphe nulle part car nous avons établi que le module définissait une application holomorphe en aucun point. On peut facilement illustrer avec la commande `conformal` de MAPLE le fait que les angles ne sont pas conservés par cette application, voir la représentation de gauche de la figure ci-dessous. Au contraire, l'application $z \in \mathbb{C}^* \mapsto 1/z^2$ qui est holomorphe dans $\mathbb{C}^*$ préserve les angles, voir la représentation de droite de la figure ci-dessous.

```
> with(plots):
> conformal(z*abs(z)^2,z=-1-I..1+I,scaling=constrained);
> conformal(1/z^2,z=-1-I..1+I,-2*(1+I)..2*(1+I),scaling=constrained);
```





4.6.3 Dérivées successives

Soient U un ouvert non vide de $\mathbb{C}$ et f une application de U dans $\mathbb{C}$, holomorphe sur U. Si la dérivée de f est à son tour holomorphe, on note f'' ou $f^{(2)}$ la dérivée de f' qui est appelée *dérivée seconde* de f. On peut ainsi de proche en proche définir pour tout $n \in \mathbb{N}^*$ la *dérivée n-ième* (ou d'ordre n) de f que l'on note $f^{(n)}$. Par convention $f^{(0)} = f$.

Exemple On vérifie, en utilisant un raisonnement par récurrence, que pour tout $(n,p) \in \mathbb{N}^2$ la dérivée n-ième de l'application $z \in \mathbb{C} \mapsto z^p$ est

$$z \in \mathbb{C} \longmapsto \left(z^p\right)^{(n)} = \begin{cases} 0 & \text{si } p < n \\ n! & \text{si } p = n \\ \dfrac{p!}{(p-n)!} z^{p-n} & \text{si } p > n \end{cases}$$

On dit que f est *indéfiniment dérivable* sur un ouvert U de $\mathbb{C}$ si pour tout $n \in \mathbb{N}^*$ l'application f admet une dérivée n-ième sur U.

Une propriété très particulière des fonctions complexes de la variable complexe est que si elles sont dérivables sur un ouvert U donné alors elles sont automatiquement indéfiniment dérivables sur cet ouvert. Ce résultat sera précisé dans le chapitre suivant consacré aux séries entières (voir le théorème 5.6 et ses conséquences). Cette propriété n'est évidement pas vraie pour les fonctions réelles de la variable réelle. Cette différence s'explique par le fait que pour une fonction réelle de la variable réelle les deux notions « dérivable » et « différentiable » coïncident. Comme nous l'avons vu, ce n'est pas le cas pour les fonctions complexes de la variable complexe. La condition « $\mathbb{C}$-dérivable » (*i.e.* holomorphe) est plus forte (ou plus contraignante) que la condition « différentiable ».

On a le résultat suivant concernant les propriétés algébriques de la dérivée n-ième d'une fonction complexe de la variable complexe (résultat qui est du même type que celui existant pour les fonctions réelles de la variable réelle[25]).

PROPOSITION 4.18 *Soient U un ouvert non vide de $\mathbb{C}$, λ un complexe, f et g deux applications de U dans $\mathbb{C}$. Si f et g sont holomorphes en $z_0 \in U$ (resp. sur U), alors pour tout $n \in \mathbb{N}^*$,*

✗ *$f+g$ admet une dérivée n-ième en z_0 (resp. sur U) et*

$$(f+g)^{(n)}(z_0) = f^{(n)}(z_0) + g^{(n)}(z_0);$$

✗ *$\lambda \cdot f$ admet une dérivée n-ième en z_0 (resp. sur U) et*

$$(\lambda \cdot f)^{(n)}(z_0) = \lambda\ f^{(n)}(z_0);$$

✗ *$f \times g$ admet une dérivée n-ième en z_0 (resp. sur U) et*

$$(f \times g)^{(n)}(z_0) = \sum_{k=0}^{n} C_n^k\ f^{(k)}(z_0)\ g^{(n-k)}(z_0) \qquad \textit{(formule de Leibniz)} ;$$

✗ *de plus, si $g(z_0) \neq 0$ (resp. g ne s'annule pas sur U) alors f/g admet une dérivée n-ième en z_0 (resp. sur U).*

Démonstration Ces formules se démontrent par récurrence à partir des relations pour la dérivée première, en admettant provisoirement qu'une fonction holomorphe sur un ouvert U est indéfiniment dérivable sur cet ouvert. La formule donnant l'expression de la dérivée n-ième du quotient f/g en fonction des dérivées de f et de g est trop compliquée pour être mentionnée. On s'en convaincra en calculant les premières dérivées. □

[25] Voir la proposition 16.6 p. 726 du *Cours de première année*.

4.7 Fonctions usuelles de la variable complexe

Nous allons voir dans cette partie comment un certain nombre de fonctions usuelles de la variable réelle admettent un prolongement en tant que fonctions de la variable complexe.

4.7.1 La fonction exponentielle complexe

Pour tout complexe z fixé, la série numérique complexe $\sum_n \frac{1}{n!} z^n$ est absolument convergente (on peut par exemple utiliser la règle de D'Alembert[26] pour l'établir). Cette propriété permet d'introduire la définition suivante.

DÉFINITION 4.15 (Fonction exponentielle complexe)
On appelle exponentielle complexe *et on note*[27] exp *l'application de* $\mathbb{C}$ *dans* $\mathbb{C}$ *définie par :*

$$\forall z \in \mathbb{C} \qquad \exp(z) = \sum_{n=0}^{+\infty} \frac{1}{n!} z^n.$$

Le choix de l'appellation *exponentielle complexe* est justifié par le fait que la fonction exponentielle complexe exp prolonge à $\mathbb{C}$ la fonction exponentielle réelle, au sens où si z désigne un complexe de partie imaginaire nulle ($z = x$ avec $x \in \mathbb{R}$) l'exponentielle du nombre complexe z est un nombre réel égal à l'exponentielle réelle du réel x : $\exp(z) = \mathrm{e}^x$. Ceci résulte du fait que pour tout réel x la série numérique réelle $\sum_n \frac{1}{n!} x^n$ a pour somme e^x, voir p. 5.

PROPOSITION 4.19 *Pour tout* $(z_1, z_2) \in \mathbb{C}^2$ *et tout* $n \in \mathbb{Z}$, *on a les relations suivantes :*

1. $\exp(z_1 + z_2) = \exp(z_1) \times \exp(z_2)$;
2. $\exp(-z_1) = \dfrac{1}{\exp(z_1)}$;
3. $\exp(z_1 - z_2) = \dfrac{\exp(z_1)}{\exp(z_2)}$;
4. $\exp(nz_1) = \Big(\exp(z_1)\Big)^n$.

[26] Voir la proposition 1.13 p. 34.

[27] Par souci de clarté et afin de distinguer dans un premier temps les deux fonctions exponentielles « réelle » et « complexe », nous noterons exp la fonction exponentielle complexe et nous désignerons par e^x l'exponentielle du réel x. Du point de vue mathématique, cette distinction n'a pas lieu d'être puisque nous verrons que la fonction exponentielle réelle coïncide avec la fonction exponentielle complexe lorsque celle-ci est restreinte aux complexes de partie imaginaire nulle.

Démonstration ▷ Pour tout $(z_1, z_2) \in \mathbb{C}^2$, on a d'après la définition 4.15 : $\exp(z_1 + z_2) = \sum_{n=0}^{+\infty} \frac{1}{n!}(z_1 + z_2)^n$. Par ailleurs, $\exp(z_1) = \sum_{n=0}^{+\infty} \frac{1}{n!} z_1^n$ et $\exp(z_2) = \sum_{n=0}^{+\infty} \frac{1}{n!} z_2^n$. On en déduit[28] que $\exp(z_1) \times \exp(z_2) = \sum_{n=0}^{+\infty} c_n$ où pour tout $n \in \mathbb{N}$ on a

$$c_n = \sum_{k=0}^{n} \frac{1}{k!} z_1^k \frac{1}{(n-k)!} z_2^{n-k} = \frac{1}{n!} \sum_{k=0}^{n} C_n^k z_1^k z_2^{n-k} = \frac{1}{n!} (z_1 + z_2)^n,$$

la dernière égalité découlant de la formule du binôme de Newton[29]. On en déduit que :

$$\exp(z_1) \times \exp(z_2) = \sum_{n=0}^{+\infty} \frac{1}{n!} (z_1 + z_2)^n.$$

La première relation de la proposition est vérifiée.

▷ Prenant $z_2 = -z_1$ dans la relation que l'on vient d'établir, on obtient les égalités :

$$\exp(z_1 - z_1) = \exp(0) = 1 = \exp(z_1) \times \exp(-z_1).$$

On en déduit dans un premier temps que $\exp(z_1) \neq 0$, puis dans un deuxième temps que

$$\exp(-z_1) = \frac{1}{\exp(z_1)}.$$

▷ En utilisant la première relation avec $-z_2$ à la place de z_2 on obtient :

$$\exp(z_1 - z_2) = \exp(z_1) \times \exp(-z_2).$$

D'après la seconde relation on a $\exp(-z_2) = 1/\exp(z_2)$ et on déduit que :

$$\exp(z_1 - z_2) = \frac{\exp(z_1)}{\exp(z_2)}.$$

▷ La dernière des relations s'obtient dans le cas où n est un entier naturel non nul par récurrence à partir de la première relation. On étend alors la propriété aux entiers relatifs négatifs en ayant recours à la seconde relation. □

Remarque On a établi dans la démonstration de la proposition 4.19 la relation $\exp(z) \times \exp(-z) = 1$ pour tout $z \in \mathbb{C}$. Cette relation implique en particulier que la fonction exponentielle complexe ne s'annule pas sur $\mathbb{C}$. ◇

---

[28] Voir la proposition 1.14 p. 39.

[29] Voir la proposition 2.12 p. 72 du *Cours de première année*.

Exponentielle d'un imaginaire pur

Considérons le cas d'un complexe z imaginaire pur : $z = \mathrm{i}\theta$ avec $\theta \in \mathbb{R}$. On obtient par la définition 4.15, en séparant la partie réelle et la partie imaginaire de l'expression :

$$\exp(\mathrm{i}\theta) = \sum_{n=0}^{+\infty} \frac{1}{n!}\mathrm{i}^n\theta^n = \sum_{k=0}^{+\infty} \frac{(-1)^k}{(2k)!}\theta^{2k} + \mathrm{i}\sum_{k=0}^{+\infty} \frac{(-1)^k}{(2k+1)!}\theta^{2k+1}. \tag{8}$$

Les séries réelles $\sum_k \frac{(-1)^k}{(2k)!}\theta^{2k}$ et $\sum_k \frac{(-1)^k}{(2k+1)!}\theta^{2k+1}$ convergent ; elles ont pour somme respectives la partie réelle et la partie imaginaire du complexe $\exp(\mathrm{i}\theta)$.

Intéressons-nous aux deux fonctions de la variable réelle f_c et f_s définies par :

$$f_c : x \in \mathbb{R} \longmapsto \sum_{k=0}^{+\infty} \frac{(-1)^k}{(2k)!}\, x^{2k} \quad \text{et} \quad f_s : x \in \mathbb{R} \longmapsto \sum_{k=0}^{+\infty} \frac{(-1)^k}{(2k+1)!}\, x^{2k+1}.$$

Nous allons montrer que la première est la fonction cosinus et la seconde la fonction sinus. On peut remarquer d'une part que $f_c(0) = 1 = \cos(0)$ et f_c est paire comme la fonction cosinus, et d'autre part que $f_s(0) = 0 = \sin(0)$ et f_s est impaire comme la fonction sinus. Il suffit donc de s'intéresser à ces fonctions sur $\mathbb{R}_+^*$ uniquement.

Pour tout $n \in \mathbb{N}$ et pour tout $x \in \mathbb{R}_+^*$, la formule de Taylor-Lagrange appliquée à la fonction cosinus entre 0 et x à l'ordre $2n$, nous assure de l'existence d'un réel c avec $0 < c < x$ tel que

$$\cos(x) = \sum_{k=0}^{2n} \frac{\cos^{(k)}(0)}{k!}\, x^k + \frac{\cos^{(2n+1)}(c)}{(2n+1)!}\, x^{2n+1}.$$

Rappelons[30] que la dérivée k–ième de la fonction cosinus est l'application $x \in \mathbb{R} \mapsto \cos(x + k\frac{\pi}{2})$. Les dérivées d'ordre k, où k est un entier impair, en 0 de la fonction cosinus valent donc 0 et les dérivées d'ordre k, où k est un entier pair, en 0 valent $(-1)^{k/2}$. On en déduit que

$$\cos(x) = \sum_{\ell=0}^{n} \frac{(-1)^\ell}{(2\ell)!}\, x^{2\ell} + \frac{\cos(c + (2n+1)\frac{\pi}{2})}{(2n+1)!}\, x^{2n+1}.$$

Pour tout $x \in \mathbb{R}_+^*$ et pour tout $n \in \mathbb{N}$, on a :

$$\left|\frac{\cos(c + (2n+1)\frac{\pi}{2})}{(2n+1)!}\, x^{2n+1}\right| \leqslant \frac{1}{(2n+1)!}\, x^{2n+1}.$$

D'après la formule de Stirling[31], $(2n+1)! \underset{+\infty}{\sim} \sqrt{2\pi(2n+1)}\left(\frac{2n+1}{\mathrm{e}}\right)^{2n+1}$. On en déduit que

$$\frac{1}{(2n+1)!}\, x^{2n+1} \underset{+\infty}{\sim} \frac{1}{\sqrt{2\pi(2n+1)}}\left(\frac{\mathrm{e}\,x}{2n+1}\right)^{2n+1} \underset{+\infty}{\sim} \frac{1}{2\sqrt{\pi n}}\left(\frac{\mathrm{e}\,x}{2n}\right)^{2n+1}.$$

[30] Voir le *Cours de première année* p. 726.

[31] Voir le *Cours de première année* p. 199 : $n! \underset{+\infty}{\sim} \sqrt{2\pi n}\,(n/\mathrm{e})^n$.

Pour tout $x \in \mathbb{R}_+^*$ fixé, on a pour n assez grand $\frac{\mathrm{e}\,x}{2n} < 1$, ce qui permet d'affirmer que

$$\lim_{n\to+\infty} \frac{1}{(2n+1)!}\, x^{2n+1} = \lim_{n\to+\infty} \frac{1}{2\sqrt{\pi n}} \left(\frac{\mathrm{e}\,x}{2n}\right)^{2n+1} = 0.$$

On en conclut que

$$\sum_{k=0}^{+\infty} \frac{(-1)^k}{(2k)!}\, x^{2k} = \lim_{n\to+\infty} \sum_{k=0}^{n} \frac{(-1)^k}{(2k)!}\, x^{2k} = \cos(x),$$

et par conséquent que f_c est la fonction cosinus. Un raisonnement similaire permet de montrer que f_s est la fonction sinus.

Ainsi, la relation (8) permet d'établir la relation suivante pour tout $\theta \in \mathbb{R}$:

$$\boxed{\exp(\mathrm{i}\theta) = \cos(\theta) + \mathrm{i}\sin(\theta)}$$

et on retrouve la notation introduite pour l'écriture polaire[32] d'un nombre complexe. En particulier la quatrième relation de la proposition 4.19 nous donne la formule de Moivre[32] :

$$\boxed{(\cos(\theta) + \mathrm{i}\sin(\theta))^n = \exp(\mathrm{i}\theta)^n = \exp(\mathrm{i}n\theta) = \cos(n\theta) + \mathrm{i}\sin(n\theta).}$$

Pour tout complexe z de partie réelle x et de partie imaginaire y, on a d'après la proposition 4.19 : $\exp(z) = \exp(x + \mathrm{i}y) = \exp(x)\,\exp(\mathrm{i}y)$. Compte tenu du fait que x et y sont réels, on a $\exp(\mathrm{i}y) = \cos(y) + \mathrm{i}\sin(y)$ et $\exp(x) = \mathrm{e}^x$. On a donc établi le lien suivant entre la fonction exponentielle de la variable complexe et les fonctions sinus, cosinus et exponentielle de la variable réelle.

Proposition 4.20 *Pour tout complexe z de partie réelle x et de partie imaginaire y, on a*

$$\exp(z) = \mathrm{e}^x\,(\cos(y) + \mathrm{i}\sin(y)).$$

La proposition 4.20 permet de représenter graphiquement la fonction exponentielle complexe, voir la figure 6.

Exercice 4.6 Déterminer les complexes z pour lesquels $\exp(z)$ est un réel puis ceux pour lesquels il s'agit d'un imaginaire pur.

On remarquera que la proposition 4.20 implique que la fonction exponentielle complexe est une fonction périodique de période $2\pi\mathrm{i}$. L'ensemble de ses périodes est $\{2k\pi\mathrm{i}\ ;\ k \in \mathbb{Z}\}$.

Corollaire 4.2 *Pour tout $z \in \mathbb{C}$ on a $\exp(\overline{z}) = \overline{\exp(z)}$.*

[32] Voir le *Cours de première année* p. 138.

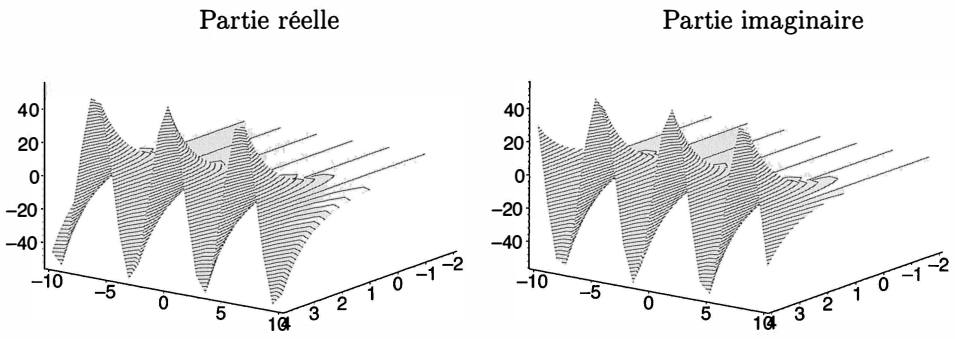


Fig. 6 Représentation graphique de la partie réelle (à gauche) et de la partie imaginaire (à droite) de la fonction exponentielle complexe.

Démonstration Soit z un complexe de partie réelle x et de partie imaginaire y ; on a $\overline{z} = x - \mathrm{i}y$. On déduit de la proposition 4.20 que

$$\exp(\overline{z}) = \mathrm{e}^x\,(\cos(-y) + \mathrm{i}\sin(-y)) = \mathrm{e}^x\,(\cos(y) - \mathrm{i}\sin(y)).$$

D'une part, puisque x est un réel, e^x est réel et par conséquent $\mathrm{e}^x = \overline{\mathrm{e}^x}$. D'autre part, y étant réel, $\cos(y)$ et $\sin(y)$ sont tous les deux réels et

$$\cos(y) - \mathrm{i}\sin(y) = \overline{\cos(y) + \mathrm{i}\sin(y)}.$$

On en déduit que :

$$\exp(\overline{z}) = \overline{\mathrm{e}^x} \times \overline{(\cos(y) + \mathrm{i}\sin(y))} = \overline{\mathrm{e}^x \times (\cos(y) + \mathrm{i}\sin(y))} = \overline{\exp(z)}.$$

La relation $\exp(\overline{z}) = \overline{\exp(z)}$ est démontrée pour tout $z \in \mathbb{C}$. □

Exercice 4.7
1 - Montrer que pour tout $z \in \mathbb{C}$ on a $|\exp(z)| = \mathrm{e}^{\mathcal{R}\mathrm{e}(z)}$.
2 - En déduire que $|\exp(z)| = 1$ si et seulement si z est imaginaire pur.

4.7.2 Fonctions trigonométriques circulaires et hyperboliques complexes

La fonction exponentielle complexe permet d'étendre de manière naturelle à $\mathbb{C}$ les fonctions circulaires et hyperboliques[33] de la variable réelle. Rappelons que pour tout réel x on a les formules d'Euler :

$$\cos(x) = \frac{\mathrm{e}^{\mathrm{i}x} + \mathrm{e}^{-\mathrm{i}x}}{2} \quad \text{et} \quad \sin(x) = \frac{\mathrm{e}^{\mathrm{i}x} - \mathrm{e}^{-\mathrm{i}x}}{2\mathrm{i}},$$

[33] Voir la section 14.6 p. 642 du *Cours de première année*.

et les fonctions hyperboliques de la variable réelle sont définies par les relations :

$$\mathrm{ch}(x) = \frac{\mathrm{e}^{x} + \mathrm{e}^{-x}}{2} \quad \text{et} \quad \mathrm{sh}(x) = \frac{\mathrm{e}^{x} - \mathrm{e}^{-x}}{2}.$$

Il est alors naturel de définir les fonctions trigonométriques circulaires complexes et hyperboliques complexes de la manière suivante.

DÉFINITION 4.16

✗ *On appelle* fonction cosinus complexe *et on note* cos *l'application*

$$z \in \mathbb{C} \longmapsto \frac{\exp(\mathrm{i}z) + \exp(-\mathrm{i}z)}{2} \in \mathbb{C}.$$

✗ *On appelle* fonction sinus complexe *et on note* sin *l'application*

$$z \in \mathbb{C} \longmapsto \frac{\exp(\mathrm{i}z) - \exp(-\mathrm{i}z)}{2\mathrm{i}} \in \mathbb{C}.$$

✗ *On appelle* fonction cosinus hyperbolique complexe *et on note* ch *l'application*

$$z \in \mathbb{C} \longmapsto \frac{\exp(z) + \exp(-z)}{2} \in \mathbb{C}.$$

✗ *On appelle* fonction sinus hyperbolique complexe *et on note* sh *l'application*

$$z \in \mathbb{C} \longmapsto \frac{\exp(z) - \exp(-z)}{2} \in \mathbb{C}.$$

Il résulte de manière directe de cette définition que les fonctions circulaires complexes et hyperboliques complexes sont liées par les relations suivantes pour tout $z \in \mathbb{C}$:

$\cos(z) = \mathrm{ch}(\mathrm{i}z)$	$\sin(z) = -\mathrm{i}\,\mathrm{sh}(\mathrm{i}z)$
$\mathrm{ch}(z) = \cos(\mathrm{i}z)$	$\mathrm{sh}(z) = -\mathrm{i}\sin(\mathrm{i}z)$

On a également les propriétés suivantes.

✗ Les applications cosinus complexe et cosinus hyperbolique complexe sont des fonctions paires. Les applications sinus complexe et sinus hyperbolique complexe sont des fonctions impaires.

✗ Les applications cosinus et sinus complexes sont des fonctions périodiques dont les périodes sont $2k\pi$, $k \in \mathbb{Z}$. Les applications cosinus et sinus hyperboliques complexes sont des fonctions périodiques dont les périodes sont $2k\pi\mathrm{i}$, $k \in \mathbb{Z}$.

PROPOSITION 4.21 *Pour tout $z \in \mathbb{C}$ on a*

$$\cos(z) = \sum_{n=0}^{+\infty} \frac{(-1)^n}{(2n)!}\, z^{2n}, \qquad \sin(z) = \sum_{n=0}^{+\infty} \frac{(-1)^n}{(2n+1)!}\, z^{2n+1},$$

et

$$\mathrm{ch}(z) = \sum_{n=0}^{+\infty} \frac{1}{(2n)!}\, z^{2n}, \qquad \mathrm{sh}(z) = \sum_{n=0}^{+\infty} \frac{1}{(2n+1)!}\, z^{2n+1}.$$

Démonstration Ces relations découlent de l'égalité $\exp(z) = \sum_{n=0}^{+\infty} \frac{1}{n!} z^n$ et de la définition 4.16. □

Exemples

1. Résolvons dans $\mathbb{C}$ l'équation $\cos(z) = 0$. D'après la définition de la fonction cosinus complexe, pour tout $z \in \mathbb{C}$ on a $\cos(z) = 0$ si et seulement si $\exp(-\mathrm{i}z) = -\exp(\mathrm{i}z)$. Désignons par x la partie réelle de z et par y sa partie imaginaire. On a

$$\exp(-\mathrm{i}z) = -\exp(\mathrm{i}z) \iff$$
$$\mathrm{e}^y \cos(x) - \mathrm{i}\mathrm{e}^y \sin(x) = -\mathrm{e}^{-y}\cos(x) - \mathrm{i}\mathrm{e}^{-y}\sin(x). \quad (9)$$

Deux complexes égaux ayant même module, on déduit de la relation (9) que $\mathrm{e}^y = \mathrm{e}^{-y}$. Cela implique que $\mathrm{e}^{2y} = 1$ donc que $y = 0$. De la relation (9) on déduit alors, en identifiant les parties réelles, que $\cos(x) = -\cos(x)$ (on a toujours égalité des parties imaginaires). On a donc $\cos(x) = 0$ autrement dit $x = \frac{\pi}{2} + k\pi$ avec $k \in \mathbb{Z}$. L'ensemble des solutions de l'équation $\cos(z) = 0$ est $S = \{\frac{1}{2}\pi + k\pi \ ; \ k \in \mathbb{Z}\}$.

Dans $\mathbb{C}$, $\cos(z) = 0$ si et seulement s'il existe $k \in \mathbb{Z}$ tel que $z = \frac{1}{2}\pi + k\pi$.

2. Effectuons un raisonnement analogue pour déterminer l'ensemble des solutions dans $\mathbb{C}$ de l'équation $\mathrm{ch}(z) = 0$. D'après la définition de la fonction cosinus hyperbolique complexe, pour tout $z \in \mathbb{C}$ on a $\mathrm{ch}(z) = 0$ si et seulement si $\exp(z) = -\exp(-z)$. Désignons par x la partie réelle de z et par y sa partie imaginaire. On a $\exp(z) = -\exp(-z)$ si et seulement si $\mathrm{e}^x\cos(y) + \mathrm{i}\mathrm{e}^x \sin(y) = -\mathrm{e}^{-x}\cos(y) + \mathrm{i}\mathrm{e}^{-x}\sin(y)$. Deux complexes égaux ayant même module, on déduit de la relation précédente que $\mathrm{e}^x = \mathrm{e}^{-x}$ ce qui implique que $x = 0$. On obtient alors, en identifiant les parties réelles, que $\cos(y) = -\cos(y)$, autrement dit que $\cos(y) = 0$. Par conséquent on a $y = \frac{\pi}{2} + k\pi$ avec $k \in \mathbb{Z}$. L'ensemble des solutions de l'équation $\mathrm{ch}(z) = 0$ est donc $S = \{\frac{1}{2}\pi\mathrm{i} + \mathrm{i}k\pi \ ; \ k \in \mathbb{Z}\}$.

Dans $\mathbb{C}$, $\mathrm{ch}(z) = 0$ si et seulement s'il existe $k \in \mathbb{Z}$ tel que $z = \frac{1}{2}\pi\mathrm{i} + k\pi\mathrm{i}$.

3. Résolvons dans $\mathbb{C}$ l'équation $\sin(z) = 2$. D'après la définition de la fonction sinus complexe, pour tout $z \in \mathbb{C}$ on a $\sin(z) = 2$ si et seulement si

$\exp(iz) - \exp(-iz) = 4i$. Désignons par x la partie réelle de z et par y sa partie imaginaire. On a

$$\begin{aligned}\exp(iz) - \exp(-iz) &= e^{-y}(\cos(x) + i\sin(x)) - e^{y}(\cos(x) - i\sin(x)) \\ &= \cos(x)(e^{-y} - e^{y}) + i\sin(x)(e^{-y} + e^{y}) \\ &= -2\cos(x)\,\mathrm{sh}(y) + 2i\sin(x)\,\mathrm{ch}(y).\end{aligned}$$

On en déduit que $\sin(z) = 2$ si et seulement si $\cos(x)\,\mathrm{sh}(y) = 0$ et $\sin(x)\,\mathrm{ch}(y) = 2$. Intéressons-nous à la première des deux conditions ; on a $\cos(x)\,\mathrm{sh}(y) = 0$ si et seulement si $y = 0$ ou $x = \frac{\pi}{2} + k\pi$, $k \in \mathbb{Z}$. Si $y = 0$ alors $\sin(x)\,\mathrm{ch}(y) = 2$ devient $\sin(x) = 2$ et aucun réel x ne satisfait cette relation. Si $x = \frac{\pi}{2} + k\pi$ avec $k \in \mathbb{Z}$ alors $\sin(x) = (-1)^{|k|}$ et $\sin(x)\,\mathrm{ch}(y) = 2$ devient $\mathrm{ch}(y) = (-2)^{|k|}$. Comme la fonction cosinus hyperbolique est minorée par 1, il ne peut y avoir de solution que si l'entier naturel $|k|$ est pair ; cette solution est $y = \mathrm{argch}(2)$. Finalement, on en conclut que les solutions dans $\mathbb{C}$ de l'équation $\sin(z) = 2$ sont les complexes $z = \frac{\pi}{2} + 2\ell\pi + i\,\mathrm{argch}(2)$ avec $\ell \in \mathbb{Z}$.

Exercice 4.8 Résoudre dans $\mathbb{C}$ les équations suivantes :

1 - $\sin(z) = 0$ 2 - $\mathrm{sh}(z) = 0$ 3 - $\cos(z) = i$

DÉFINITION 4.17

✗ *On appelle* fonction tangente complexe *et on note* tan *l'application définie pour tout* $z \in \mathbb{C} \setminus \{\frac{1}{2}\pi + k\pi \ ; \ k \in \mathbb{Z}\}$ *par*

$$\tan(z) = \frac{\sin(z)}{\cos(z)}.$$

✗ *On appelle* fonction tangente hyperbolique complexe *et on note* th *l'application définie pour tout* $z \in \mathbb{C} \setminus \{\frac{1}{2}\pi i + k\pi i \ ; \ k \in \mathbb{Z}\}$ *par*

$$\mathrm{th}(z) = \frac{\mathrm{sh}(z)}{\mathrm{ch}(z)}.$$

La fonction tangente complexe (resp. tangente hyperbolique complexe) « prolonge » à $\mathbb{C}$ la fonction réelle de la variable réelle tangente (resp. tangente hyperbolique) au sens où si z désigne un complexe de partie imaginaire nulle ces deux fonctions prennent les mêmes valeurs.

Formulaire de trigonométrie circulaire complexe

Pour tout complexe z, on obtient en utilisant la relation $a^2 - b^2 = (a-b)(a+b)$,

$$\begin{aligned}\cos^2(z) + \sin^2(z) &= \tfrac{1}{4}\big((\exp(iz) + \exp(-iz))^2 - (\exp(iz) - \exp(-iz))^2\big) \\ &= \exp(-iz) \times \exp(iz).\end{aligned}$$

On en déduit la relation suivante :

$$\boxed{\forall z \in \mathbb{C} \qquad \cos^2(z) + \sin^2(z) = 1}$$

De la même manière, en revenant à la définition des fonctions sinus et cosinus à l'aide de la fonction exponentielle complexe et en utilisant les propriétés de cette dernière, on montre pour tout $(z, z') \in \mathbb{C}^2$ les quatre relations suivantes, appelées *formules d'addition des arcs* :

$$\boxed{\begin{array}{rcl} \cos(z+z') & = & \cos(z)\cos(z') - \sin(z)\sin(z') \\ \cos(z-z') & = & \cos(z)\cos(z') + \sin(z)\sin(z') \\ \sin(z+z') & = & \sin(z)\cos(z') + \cos(z)\sin(z') \\ \sin(z-z') & = & \sin(z)\cos(z') - \cos(z)\sin(z') \end{array}}$$

En prenant z' égal à z dans les égalités précédentes, on obtient les formules dites *de doublement des arcs* :

$$\boxed{\begin{array}{rcl} \cos(2z) & = & \cos^2(z) - \sin^2(z) = 2\cos^2(z) - 1 = 1 - 2\sin^2(z) \\ \sin(2z) & = & 2\cos(z)\sin(z) \end{array}}$$

Les formules suivantes se déduisent immédiatement des quatre formules *d'addition des arcs*. Pour tout $(z, z') \in \mathbb{C}^2$,

$$\boxed{\begin{array}{rcl} 2\cos(z) \times \cos(z') & = & \cos(z+z') + \cos(z-z') \\ 2\sin(z) \times \sin(z') & = & \cos(z-z') - \cos(z+z') \\ 2\sin(z) \times \cos(z') & = & \sin(z+z') + \sin(z-z') \end{array}}$$

En remplaçant z par $\frac{1}{2}(z+z')$ et z' par $\frac{1}{2}(z-z')$ dans les relations précédentes on obtient les relations :

$$\boxed{\begin{array}{rcl} \sin(z) + \sin(z') & = & 2\sin\left(\dfrac{z+z'}{2}\right)\cos\left(\dfrac{z-z'}{2}\right) \\ \sin(z) - \sin(z') & = & 2\sin\left(\dfrac{z-z'}{2}\right)\cos\left(\dfrac{z+z'}{2}\right) \\ \cos(z) + \cos(z') & = & 2\cos\left(\dfrac{z+z'}{2}\right)\cos\left(\dfrac{z-z'}{2}\right) \\ \cos(z) - \cos(z') & = & -2\sin\left(\dfrac{z+z'}{2}\right)\sin\left(\dfrac{z-z'}{2}\right) \end{array}}$$

Formulaire de trigonométrie hyperbolique complexe

Pour tout $z \in \mathbb{C}$, on a les relations suivantes qui s'obtiennent aisément à partir de la définition 4.16 :

$$\boxed{\begin{array}{rcl} \mathrm{ch}(z) + \mathrm{sh}(z) & = & \exp(z) \\ \mathrm{ch}(z) - \mathrm{sh}(z) & = & \exp(-z) \\ \mathrm{ch}^2(z) - \mathrm{sh}^2(z) & = & 1 \end{array}}$$

Pour tout $(z, z') \in \mathbb{C}^2$, on a les quatre relations suivantes :

$$\begin{array}{rcl} \mathrm{ch}(z+z') & = & \mathrm{ch}(z)\,\mathrm{ch}(z') + \mathrm{sh}(z)\,\mathrm{sh}(z') \\ \mathrm{ch}(z-z') & = & \mathrm{ch}(z)\,\mathrm{ch}(z') - \mathrm{sh}(z)\,\mathrm{sh}(z') \\ \mathrm{sh}(z+z') & = & \mathrm{sh}(z)\,\mathrm{ch}(z') + \mathrm{ch}(z)\,\mathrm{sh}(z') \\ \mathrm{sh}(z-z') & = & \mathrm{sh}(z)\,\mathrm{ch}(z') - \mathrm{ch}(z)\,\mathrm{sh}(z') \end{array}$$

En prenant z' égal à z dans les égalités précédentes, on obtient :

$$\mathrm{ch}(2z) = \mathrm{ch}^2(z) + \mathrm{sh}^2(z) = 2\,\mathrm{ch}^2(z) - 1 = 1 + 2\,\mathrm{sh}^2(z)$$
$$\mathrm{sh}(2z) = 2\,\mathrm{ch}(z)\,\mathrm{sh}(z)$$

Exercice 4.9

1 - Déterminer les ensembles de définition des fonctions complexes de la variable complexe suivantes :

$$f_1 : z \mapsto \frac{\sin(\mathcal{R}\mathrm{e}(z))}{\mathrm{sh}(z)} \qquad f_2 : z \mapsto \frac{\mathcal{R}\mathrm{e}(\sin(z))}{\mathrm{sh}(z)} \qquad f_3 : z \longmapsto \frac{\sin(z)}{\mathcal{R}\mathrm{e}(\mathrm{sh}(z))}.$$

2 - En considérant des arcs paramétrés particuliers dans $\mathbb{C}$, montrer que ces fonctions n'ont pas de limite dans $\mathbb{C}$ en 0.

4.7.3 La fonction logarithme complexe

Dans le cadre des fonctions de la variable réelle, la fonction exponentielle définit une bijection de $\mathbb{R}$ dans $\mathbb{R}_+^*$ dont la bijection réciproque est la fonction logarithme. Nous allons voir que dans le cas des fonctions de la variable complexe, la fonction exponentielle complexe n'est pas injective sur $\mathbb{C}$.

Soient w un nombre complexe de partie réelle u et de partie imaginaire v et z un nombre complexe non nul tel que $\exp(w) = z$. On a $\exp(w) = \mathrm{e}^u \exp(\mathrm{i}v)$ et $z = |z| \exp(\mathrm{i}\mathrm{Arg}(z))$ où Arg désigne la détermination principale de l'argument (*i.e.* $\mathrm{Arg}(z) \in]-\pi, \pi]$). On déduit de la relation $\mathrm{e}^u \exp(\mathrm{i}v) = |z| \exp(\mathrm{i}\,\mathrm{Arg}(z))$ que u est entièrement déterminé par z via la relation $u = \ln(|z|)$. Par contre, v est égal à $\mathrm{Arg}(z)$ modulo 2π. Cela implique que la fonction exponentielle complexe n'est pas une bijection sur $\mathbb{C}$ (alors que la fonction exponentielle réelle est une bijection de $\mathbb{R}$ dans $\mathbb{R}_+^*$). Si l'on fait le choix de prendre $v = \mathrm{Arg}(z)$ ($v \in]-\pi, \pi]$) alors on peut définir l'application :

$$L : z \in \mathbb{C}^* \longmapsto \ln(|z|) + \mathrm{i}\mathrm{Arg}(z) \in G$$

où $G = \{w \in \mathbb{C}\ ;\ \mathcal{I}\mathrm{m}(w) \in]-\pi, \pi]\}$. Cette application vérifie :

$$\exp(L(z)) = z \qquad \forall z \in \mathbb{C}^* \qquad \text{et} \qquad L(\exp(w)) = w \qquad \forall w \in G.$$

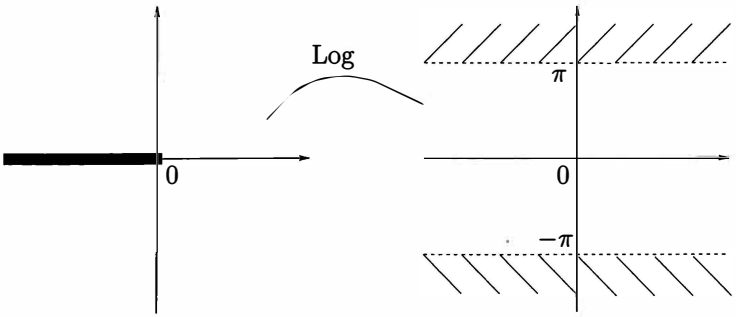


Fig. 7 Représentation graphique de l'ensemble de définition et de l'image de la fonction logarithme complexe.

Il s'agit donc de la bijection réciproque de la restriction à l'ensemble G de la fonction exponentielle.

L'application L présente la particularité de ne pas être continue sur son ensemble de définition $\mathbb{C}^*$. Considérons en effet un complexe z_0 de partie réelle négative et de partie imaginaire nulle : $z_0 = x_0$ avec $x_0 \in \mathbb{R}_-^*$. On a $L(z_0) = \ln(|z_0|) + \mathrm{i}\mathrm{Arg}(z_0) = \ln(-x_0) + \mathrm{i}\pi$. Considérons un arc géométrique de paramétrisation $\phi : t \in \mathbb{R} \mapsto x_0 + \mathrm{i}t$ dont le support est une droite parallèle à l'axe des imaginaires contenant le point d'affixe x_0. On a $\phi(0) = x_0$ et

$$L(\phi(t)) = \ln(|x_0 + \mathrm{i}t|) + \mathrm{i}\mathrm{Arg}(x_0 + \mathrm{i}t).$$

Pour $t < 0$, on a $\mathrm{Arg}(x_0 + \mathrm{i}t) \in]-\pi, -\frac{1}{2}\pi[$ et

$$\lim_{t \to 0^-} L(\phi(t)) = \ln(-x_0) - \mathrm{i}\pi \neq L(x_0).$$

Comme la fonction Arg est continue[34] sur

$$\mathbb{C}^\ddagger = \mathbb{C} \setminus \{z \in \mathbb{C} \ ; \ \mathcal{R}\mathrm{e}(z) \leqslant 0 \ \text{ et } \ \mathcal{I}\mathrm{m}(z) = 0\}$$

et que la fonction $z \mapsto \ln|z|$ est continue sur $\mathbb{C}^*$ (donc sur $\mathbb{C}^\ddagger$), on en déduit que l'ensemble de continuité de L est $\mathbb{C}^\ddagger$.

Ces considérations motivent la définition suivante.

DÉFINITION 4.18 (Détermination principale du logarithme)
On appelle détermination principale du logarithme complexe *(ou plus simplement logarithme complexe) l'application notée* Log *définie sur l'ensemble* $\mathbb{C}^\ddagger = \mathbb{C} \setminus \{z \in \mathbb{C} \ ; \ \mathcal{R}e(z) \leqslant 0 \ \text{ et } \ \mathcal{I}m(z) = 0\}$ *par*

$$\mathrm{Log} : z \in \mathbb{C}^\ddagger \longmapsto \ln(|z|) + \mathrm{i}\mathrm{Arg}(z) \in F$$

où Arg *désigne la détermination principale de l'argument :* $\mathrm{Arg}(z) \in]-\pi, \pi[$ *et* $G = \{w \in \mathbb{C} \ ; \ \mathcal{I}m(w) \in]-\pi, \pi[\}$.

[34] Voir p. 192 pour la justification de ce résultat.

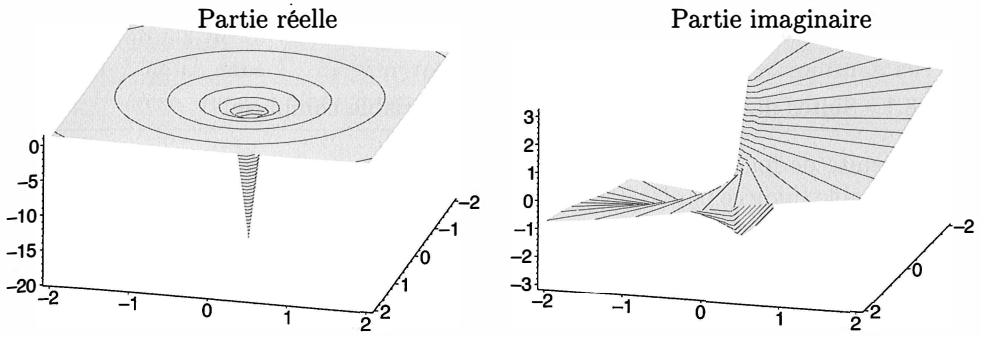


Fig. 8 Représentation graphique de la partie réelle (à gauche) et de la partie imaginaire (à droite) de la fonction logarithme complexe.

Remarques

1. Par construction, la fonction Log est une application continue sur $\mathbb{C}^{\ddagger}$.

2. Si z_1, z_2 et $z_1 z_2$ appartiennent à $\mathbb{C}^{\ddagger}$, on peut montrer qu'il existe un entier relatif k tel que

$$\mathrm{Log}(z_1 z_2) = \mathrm{Log}(z_1) + \mathrm{Log}(z_2) + 2\mathrm{i}k\pi.$$

Si de plus $\mathrm{Arg}(z_1) + \mathrm{Arg}(z_2) \in]-\pi, \pi]$ alors[35] :

$$\mathrm{Log}(z_1 z_2) = \mathrm{Log}(z_1) + \mathrm{Log}(z_2).$$

◇

Rappelons, voir p. 175, que pour tout $n \in \mathbb{N}^*$, on appelle *détermination principale de la fonction racine n-ième complexe* l'application :

$$z \in \mathbb{C}^* \longmapsto \sqrt[n]{|z|}\, \mathrm{e}^{\frac{\mathrm{i}}{n}\mathrm{Arg}(z)}.$$

Pour $z \in \mathbb{C}^{\ddagger}$, on peut écrire :

$$\sqrt[n]{z} = \sqrt[n]{|z|}\, \mathrm{e}^{\frac{\mathrm{i}}{n}\mathrm{Arg}(z)} = \exp\left(\frac{1}{\ln|z|} + \frac{\mathrm{i}}{n}\mathrm{Arg}(z)\right) = \exp(\tfrac{1}{n}\mathrm{Log}(z)).$$

[35] Voir le *Cours de première année* p. 136.

4.8 Suites et séries de fonctions complexes de la variable complexe

4.8.1 Suites de fonctions complexes de la variable complexe

Une suite de fonctions complexes de la variable complexe est une famille dénombrable (indicée par $\mathbb{N}$) d'applications définies sur un même sous-ensemble Z de $\mathbb{C}$ à valeurs dans $\mathbb{C}$.

On dit que la suite d'applications $(f_n)_n$ *converge simplement* sur Z vers l'application f de Z dans $\mathbb{C}$ si

$$\boxed{\forall z \in Z} \quad \forall \varepsilon \in \mathbb{R}_+^* \quad \exists N_{z,\varepsilon} \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad \big(n \geqslant N_{z,\varepsilon} \implies |f_n(z) - f(z)| \leqslant \varepsilon\big).$$

On remarquera que cette définition est identique à la définition 1 p. 71 de la convergence simple d'une suite de fonctions de la variable réelle, le module remplaçant la valeur absolue.

De la même manière, en accord avec la définition 2 p. 73, on dit que la suite d'applications $(f_n)_n$ *converge uniformément* sur Z vers l'application f de Z dans $\mathbb{C}$, si

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N_\varepsilon \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad \boxed{\forall z \in Z} \quad \big(n \geqslant N_\varepsilon \implies |f_n(z) - f(z)| \leqslant \varepsilon\big).$$

On a le résultat suivant : si la suite d'applications $(f_n)_n$ de Z dans $\mathbb{C}$ converge uniformément vers l'application f de Z dans $\mathbb{C}$ alors elle converge aussi simplement vers f sur Z. Ce résultat qui prolonge au cas de la variable complexe le théorème 2.1 p. 78 se démontre de la même manière : il suffit de considérer le module à chaque endroit où figurait la valeur absolue.

Les façons de montrer qu'une suite de fonctions complexes de la variable complexe converge simplement ou uniformément sur un sous-ensemble Z de $\mathbb{C}$ sont identiques à celles qui ont été présentées au chap. 2 dans le cas des de fonctions réelles de la variable réelle. Mentionnons en particulier, voir la proposition 2.2 p. 74, qu'une condition nécessaire et suffisante pour que la suite $(f_n)_n$ converge uniformément vers f sur Z est que la suite réelle de terme général

$$u_n = \sup_{z \in Z} |f_n(z) - f(z)|$$

converge vers 0. D'une manière générale, tous les résultats de la section 2.2 du chap. 2 restent valables dans le cas d'une suite de fonctions complexes de la variable complexe. Il suffit de considérer le module à la place de la valeur absolue.

Concernant les propriétés de la limite d'une suite de fonctions de la variable complexe, on a le résultat suivant qui étend au cas des fonctions de la variable complexe le théorème 2.3 p. 86 et dont la preuve s'effectue de manière très semblable.

PROPOSITION 4.22 *Soient U un ouvert non vide de $\mathbb{C}$, et $(f_n)_n$ une suite d'applications de U dans $\mathbb{C}$, convergeant uniformément sur U vers l'application f de U dans $\mathbb{C}$. Si pour tout $n \in \mathbb{N}$ l'application f_n est continue en $z_0 \in U$ (resp. sur U) alors f est continue en z_0 (resp. sur U).*

Signalons que les théorèmes 2.5 et 2.6 p. 92 donnant des conditions de dérivabilité de la limite d'une suite de fonctions de la variable réelle n'admettent pas d'énoncé sous la même forme dans le cas des suites de fonctions de la variable complexe. Cette distinction est liée aux propriétés très spéciales que possèdent les fonctions holomorphes. En particulier, la convergence uniforme d'une suite de fonctions $(f_n)_n$ de la variable complexe sur un ouvert connexe U de $\mathbb{C}$ entraîne celle de la suite $(f_n')_n$ sur tout sous-ensemble fermé borné inclus dans U.

4.8.2 Séries de fonctions complexes de la variable complexe

Étant donné une suite d'applications $(f_n)_n$ définies sur un sous-ensemble non vide Z de $\mathbb{C}$, on appelle *série d'applications de Z dans $\mathbb{C}$ de terme général f_n* et on note $\sum_n f_n$, le couple $((f_n)_{n\in\mathbb{N}}, (\phi_n)_{n\in\mathbb{N}})$ d'éléments de $\mathcal{S}(Z,\mathbb{C}) \times \mathcal{S}(Z,\mathbb{C})$ où $(\phi_n)_n$ est la suite d'applications de terme général $\phi_n = \sum_{k=0}^{n} f_k$ appelée *suite des sommes partielles* associée à la série $\sum_n f_n$.

On dit que la série d'applications $\sum_n f_n$ *converge simplement* sur Z si la suite d'applications $(\phi_n)_n$ converge simplement sur Z. On appelle alors somme de la série d'applications $\sum_n f_n$, la limite ϕ de la suite d'applications $(\phi_n)_n$ et on note

$$\forall z \in Z \qquad \phi(z) = \sum_{n=0}^{+\infty} f_n(z).$$

On dit que la série d'applications $\sum_n f_n$ *converge uniformément* (resp. absolument) sur Z si la suite d'applications $(\phi_n)_n$ converge uniformément (resp. absolument) sur Z.

On dit que la série d'applications $\sum_n f_n$ de Z dans $\mathbb{C}$ converge normalement sur Z si la série numérique réelle $\sum_n u_n$ de terme général $u_n = \sup_{z\in Z} |f_n(z)|$ converge.

Comme dans le cas des séries de fonctions de la variable réelle, l'étude de la convergence uniforme d'une série de fonctions de la variable complexe passe dans un premier temps par l'étude de la convergence normale. Si la série converge normalement sur un ensemble Z donné alors elle converge uniformément sur cet ensemble. Dans le cas où il n'y a pas convergence normale, on peut avoir recours au résultat suivant qui étend au cas des séries de fonctions de la variable complexe le théorème 3.6.

THÉORÈME 4.3 (Règle d'Abel uniforme)
Soit $(f_n)_n$ une suite d'applications d'un sous-ensemble Z de $\mathbb{C}$ dans $\mathbb{C}$ dont le terme général est de la forme

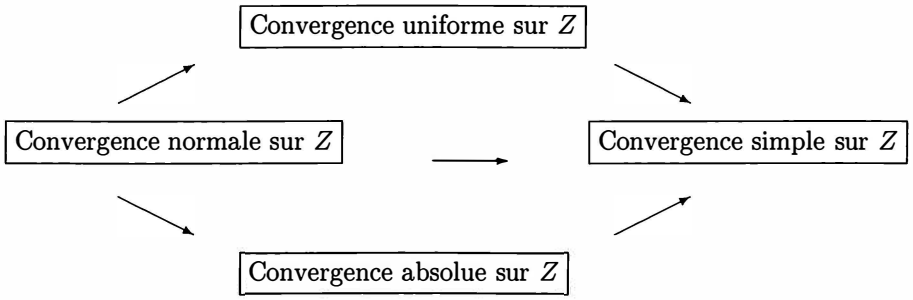
$$f_n(z) = g_n(z) \times h_n(z)$$

où $(g_n)_n$ est une suite d'applications de Z dans $\mathbb{R}_+^*$ et $(h_n)_n$ est une suite d'applications de Z dans $\mathbb{C}$. Si,

1. *pour tout $z \in Z$ fixé, la suite numérique réelle $(g_n(z))_n$ est une suite décroissante ;*
2. *la suite d'applications $(g_n)_n$ converge uniformément sur Z vers l'application nulle ;*
3. *il existe un réel positif M tel que : $\forall z \in Z \quad \forall n \in \mathbb{N} \quad |\sum_{k=0}^{n} h_k(z)| \leqslant M$;*

alors la série d'applications $\sum_n f_n$ converge uniformément sur Z.

Comme dans le cas des séries de fonctions réelles, les liens d'implication entre les différents types de convergence peuvent être représentés de la manière suivante :



Si l'on s'intéresse aux propriétés de la somme d'une série de fonctions de la variable complexe, on a le résultat suivant qui étend le théorème 3.7 p. 135 et qui se démontre d'une manière très semblable.

PROPOSITION 4.23 *Soient U un ouvert non vide de $\mathbb{C}$ et $\sum_n f_n$ une série d'applications de U dans $\mathbb{C}$ convergeant uniformément sur U. Si pour tout $n \in \mathbb{N}$, l'application f_n est continue en $z_0 \in U$ (resp. sur U) alors la somme de la série $\phi : z \in U \mapsto \sum_{n=0}^{+\infty} f_n(z)$ est continue en z_0 (resp. sur U).*

4.8.3 Continuité des fonctions usuelles de la variable complexe

La fonction exponentielle complexe a été définie comme la somme de la série $\sum_n \frac{1}{n!} z^n$, voir la définition 4.15. La série d'applications $\sum_n f_n$ de terme général

$f_n : z \in \mathbb{C} \mapsto \frac{1}{n!}z^n$ converge normalement sur la boule $D(0,R)$ pour tout $R \in \mathbb{R}_+^*$ puisque

$$\sup_{z\in D(0,R)} |f_n(z)| = \frac{1}{n!}R^n$$

et le critère de D'Alembert permet d'établir que la série $\sum_n \frac{1}{n!}R^n$ converge.

D'après la proposition 4.23, la fonction exponentielle complexe est continue sur la boule $D(0,R)$ pour tout $R \in \mathbb{R}_+^*$, autrement dit elle est continue sur $\mathbb{C}$.

Il résulte alors de la définition 4.16 p. 209 et de la proposition 4.10 p. 191 que les fonctions sinus et sinus hyperbolique complexes, cosinus et cosinus hyperbolique complexes sont continues sur $\mathbb{C}$.

La proposition 4.23 permet également d'établir certains résultats concernant les limites des fonctions usuelles. Pour tout $z \in \mathbb{C}^*$ on a $\exp(z) = \sum_{n=0}^{+\infty} \frac{1}{n!}z^n$ et par conséquent

$$\frac{\exp(z)-1}{z} = \sum_{n=0}^{+\infty} \frac{1}{(n+1)!}z^n = 1 + \sum_{n=1}^{+\infty} \frac{1}{(n+1)!}z^n.$$

La série d'applications $\sum_n g_n$ de terme général $g_n : z \in \mathbb{C} \mapsto \frac{1}{(n+1)!}z^n$ convergeant normalement sur la boule $D(0,R)$ pour tout $R \in \mathbb{R}_+^*$, on déduit de la proposition 4.23 que $\lim_{z\to 0} \sum_{n=1}^{+\infty} \frac{1}{(n+1)!}z^n = \sum_{n=1}^{+\infty} \lim_{z\to 0} \frac{1}{(n+1)!}z^n = 0$. On a donc :

$$\boxed{\lim_{\substack{z\to 0\\ z\in\mathbb{C}}} \frac{\exp(z)-1}{z} = 1.}$$

Pour $z \in \mathbb{C}^*$ on a,

$$\frac{\sin(z)}{z} = \frac{\exp(iz)-\exp(-iz)}{2iz} = \exp(-iz)\frac{\exp(2iz)-1}{2iz}.$$

Comme $\lim\limits_{\substack{z\to 0\\ z\in\mathbb{C}}} \frac{\exp(2iz)-1}{2iz} = \lim\limits_{\substack{\zeta\to 0\\ \zeta\in\mathbb{C}}} \frac{\exp(\zeta)-1}{\zeta} = 1$ et comme $\lim\limits_{\substack{z\to 0\\ z\in\mathbb{C}}} \exp(-iz) = 1$, on en déduit que :

$$\boxed{\lim_{\substack{z\to 0\\ z\in\mathbb{C}}} \frac{\sin(z)}{z} = 1.}$$

Pour $z \in \mathbb{C}^*$ on a $\mathrm{sh}(z) = -i\sin(iz)$ et

$$\frac{\mathrm{sh}(z)}{z} = -\frac{i\sin(iz)}{z} = \frac{\sin(iz)}{iz}.$$

Comme $\lim\limits_{z\to 0} \frac{\sin(iz)}{iz} = \lim\limits_{\zeta\to 0} \frac{\sin(\zeta)}{\zeta}$, on en déduit que

$$\boxed{\lim_{\substack{z\to 0\\ z\in\mathbb{C}}} \frac{\mathrm{sh}(z)}{z} = 1.}$$

Compte tenu du fait que la fonction cosinus (resp. cosinus hyperbolique) est continue en 0 et prend la valeur 1 en 0, on déduit des relations précédentes que :

$$\boxed{\lim_{\substack{z\to 0\\ z\in\mathbb{C}}} \frac{\tan(z)}{z} = 1, \qquad \lim_{\substack{z\to 0\\ z\in\mathbb{C}}} \frac{\mathrm{th}(z)}{z} = 1.}$$

Enfin, pour tout $z \in \mathbb{C}$ on a $\cos(z) = \cos(2\frac{z}{2}) = 1 - 2\sin^2(\frac{z}{2})$ et par conséquent

$$\frac{\cos(z)-1}{z^2} = -2\frac{\sin^2(\frac{z}{2})}{z^2} = -\frac{1}{2}\left(\frac{\sin(\frac{z}{2})}{\frac{z}{2}}\right)^2.$$

Comme $\lim\limits_{\substack{z\to 0\\ z\in\mathbb{C}}} \frac{\sin(\frac{z}{2})}{\frac{z}{2}} = \lim\limits_{\substack{\zeta\to 0\\ \zeta\in\mathbb{C}}} \frac{\sin(\zeta)}{\zeta} = 1$, on en déduit que

$$\boxed{\lim_{\substack{z\to 0\\ z\in\mathbb{C}}} \frac{\cos(z)-1}{z^2} = -\frac{1}{2}.}$$

De plus, comme $\mathrm{ch}(z) = \cos(\mathrm{i}z)$ on a

$$\frac{\mathrm{ch}(z)-1}{z^2} = -\frac{\cos(\mathrm{i}z)-1}{(\mathrm{i}z)^2}$$

et puisque $\lim\limits_{\substack{z\to 0\\ z\in\mathbb{C}}} \frac{\cos(\mathrm{i}z)-1}{(\mathrm{i}z)^2} = \lim\limits_{\substack{\zeta\to 0\\ \zeta\in\mathbb{C}}} \frac{\cos(\zeta)-1}{\zeta^2} = -\frac{1}{2}$, on en déduit que

$$\boxed{\lim_{\substack{z\to 0\\ z\in\mathbb{C}}} \frac{\mathrm{ch}(z)-1}{z^2} = \frac{1}{2}.}$$

4.9 Exercices de synthèse

Exercice 4.10

1 - Soit z un nombre complexe de partie réelle x et de partie imaginaire y.

a) Exprimer $|\sin(z)|$ à l'aide des fonctions réelles de la variable réelle sinus, cosinus, sinus hyperbolique et cosinus hyperbolique en x et y.

b) En déduire que $|\mathrm{sh}(y)| \leqslant |\sin(z)| \leqslant \mathrm{ch}(y)$.

c) Montrer que les fonction sinus et cosinus ne sont pas bornées sur $\mathbb{C}$.

2 - Montrer que pour tout $z \in \mathbb{C}$ on a $|\cos(z)|^2 + |\sin(z)|^2 \geqslant 1$ et que l'égalité a lieu uniquement si z est réel.

3 - a) Montrer que la série de fonctions complexes $\sum_n f_n$ de terme général $f_n : z \mapsto \frac{1}{n^2}\sin\left(\frac{z}{n^2}\right)$ converge absolument dans $\mathbb{C}$.

On désigne par f l'application de $\mathbb{C}$ dans $\mathbb{C}$ qui est la somme de cette série.

b) Montrer que cette série ne converge pas normalement sur $\mathbb{C}$ mais qu'elle converge normalement sur le disque $D(0, R)$ pour tout $R \in \mathbb{R}_+$.

c) Justifier que f est une application continue sur $\mathbb{C}$.

4.10 Solutions des exercices

Solution de l'exercice 4.1

Considérons la boule ouverte $D(z_0, r_0)$ où $z_0 \in \mathbb{C}$ et $r_0 \in \mathbb{R}_+^*$ et montrons que pour tout $z_1 \in D(z_0, r_0)$, il existe $r_1 \in \mathbb{R}_+^*$ tel que la boule ouverte $D(z_1, r_1)$ soit incluse dans $D(z_0, r_0)$.

Pour $z_1 \in D(z_0, r_0)$ posons $r_1 = r_0 - |z_0 - z_1|$. Quel que soit $z \in D(z_1, r_1)$ on a $|z - z_1| < r_1$ et

$$|z - z_0| \leqslant |z - z_1| + |z_1 - z_0| < r_1 + |z_1 - z_0| = r_0,$$

ce qui signifie que $z \in D(z_0, r_0)$. On en déduit que pour tout $z_1 \in D(z_0, r_0)$, la boule ouverte $D(z_1, r_1)$ avec $r_1 = r_0 - |z_0 - z_1|$ est incluse dans $D(z_0, r_0)$ et on peut en conclure que toute boule ouverte de $\mathbb{C}$ est un ensemble ouvert.

Solution de l'exercice 4.2

1 - Pour tous complexes z et z_0 on a

$$|f_1(z) - f_1(z_0)| = |\mathcal{R}\mathrm{e}(z) - \mathcal{R}\mathrm{e}(z_0)| \leqslant |z - z_0|.$$

On en déduit que pour tout réel ε strictement positif fixé, il existe un réel strictement positif η (prendre par exemple $\eta = \varepsilon$) tel que si $|z - z_0| \leqslant \eta$ alors $|f_1(z) - f_1(z_0)| \leqslant \varepsilon$.

2 - Pour tous complexes z et z_0 on a

$$|f_2(z) - f_2(z_0)| = |\mathcal{I}\mathrm{m}(z) - \mathcal{I}\mathrm{m}(z_0)| \leqslant |z - z_0|.$$

On en déduit que pour tout réel ε strictement positif fixé, il existe un réel strictement positif η (prendre par exemple $\eta = \varepsilon$) tel que si $|z - z_0| \leqslant \eta$ alors $|f_2(z) - f_2(z_0)| \leqslant \varepsilon$.

3 - Pour tous complexes z et z_0 on a

$$|f_3(z) - f_3(z_0)| = |\overline{z} - \overline{z_0}| = |\overline{z - z_0}| = |z - z_0|.$$

On en déduit que pour tout réel ε strictement positif fixé, il existe un réel strictement positif η (prendre par exemple $\eta = \varepsilon$) tel que si $|z - z_0| \leqslant \eta$ alors $|f_3(z) - f_3(z_0)| \leqslant \varepsilon$.

4 - La seconde inégalité triangulaire implique que pour tous complexes z et z_0 on a

$$|f_4(z) - f_4(z_0)| = ||z| - |z_0|| \leqslant |z - z_0|.$$

On en déduit que pour tout réel ε strictement positif fixé, il existe un réel strictement positif η (prendre par exemple $\eta = \varepsilon$) tel que si $|z - z_0| \leqslant \eta$ alors $|f_4(z) - f_4(z_0)| \leqslant \varepsilon$.

Solution de l'exercice 4.3

Supposons que $z_0 \neq 0$ et considérons un premier arc géométrique qui est la droite passant par z_0 et qui est parallèle à l'axe des réels. Cet arc admet pour paramétrisation $\phi_1 : t \in \mathbb{R} \mapsto z_0 + t$. En désignant par x_0 la partie réelle de z_0 et par y_0 sa partie imaginaire, on obtient pour tout réel t non nul,

$$\begin{aligned} f(\phi_1(t)) &= \frac{|z_0+t|-|z_0|}{t} = \frac{1}{t}\left(\sqrt{(x_0+t)^2+y_0^2}-\sqrt{x_0^2+y_0^2}\right) \\ &= \frac{2x_0+t}{\sqrt{(x_0+t)^2+y_0^2}+\sqrt{x_0^2+y_0^2}}. \end{aligned}$$

On a donc

$$\lim_{t\to 0} f(\phi_1(t)) = \frac{x_0}{\sqrt{x_0^2+y_0^2}}.$$

Considérons à présent un arc géométrique qui est la droite passant par z_0 et qui est parallèle à l'axe des imaginaires. Cet arc admet pour paramétrisation $\phi_2 : t \in \mathbb{R} \mapsto z_0 + \mathrm{i}t$. On obtient pour tout réel t non nul,

$$\begin{aligned} f(\phi_2(t)) &= \frac{|z_0+t|-|z_0|}{\mathrm{i}t} = \frac{1}{\mathrm{i}t}\left(\sqrt{x_0^2+(y_0+t)^2}-\sqrt{x_0^2+y_0^2}\right) \\ &= \frac{2y_0+t}{\mathrm{i}\sqrt{x_0^2+(y_0+t)^2}+\sqrt{x_0^2+y_0^2}}. \end{aligned}$$

On a donc

$$\lim_{t\to 0} f(\phi_2(t)) = \frac{y_0}{\mathrm{i}\sqrt{x_0^2+y_0^2}}.$$

Puisque la limite est différente selon les deux arcs paramétrés considérés, f n'a pas de limite en z_0.

Si $z_0 = 0$, on a $f : z \in \mathbb{C}^* \mapsto |z|/z$. Considérons un arc géométrique qui est l'axe des réels de paramétrisation $\phi_3 : t \in \mathbb{R} \mapsto t$. Pour tout réel t non nul, on a

$$f(\phi_3(t)) = \frac{|t|}{t} = \begin{cases} 1 & \text{si } t > 0 \\ -1 & \text{si } t < 0 \end{cases}.$$

La fonction $f \circ \phi_3$ n'ayant pas de limite en 0, f n'a pas non plus de limite en 0.

Solution de l'exercice 4.4

Supposons que ϕ soit continue en $t_0 \in]a,b[$ *i.e.* supposons que

$$\forall \varepsilon_1 \in \mathbb{R}_+^* \quad \exists \eta_1 \in \mathbb{R}_+^* \quad \forall t \in]a,b[\quad \Big(|t-t_0| \leqslant \eta_1 \Longrightarrow |\phi(t)-\phi(t_0)| \leqslant \varepsilon_1\Big). \quad (10)$$

et supposons que f soit continue en $z_0 = \phi(t_0)$, *i.e.* supposons que

$$\forall \varepsilon_2 \in \mathbb{R}_+^* \quad \exists \eta_2 \in \mathbb{R}_+^* \quad \forall z \in U \quad \Big(|z-z_0| \leqslant \eta_2 \Longrightarrow |f(z)-f(z_0)| \leqslant \varepsilon_2\Big). \quad (11)$$

Pour montrer que $f \circ \phi$ est continue en t_0, il faut montrer que

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists \eta \in \mathbb{R}_+^* \quad \forall t \in]a,b[\quad \Big(|t - t_0| \leqslant \eta \Longrightarrow |f(\phi(t)) - f(\phi(t_0)| \leqslant \varepsilon \Big). \quad (12)$$

Soit ε un réel strictement positif fixé, t un élément de $]a,b[$ et $z = \phi(t)$ un élément de U. D'après (11) considérée avec $\varepsilon_2 = \varepsilon$, il existe $\eta_2 \in \mathbb{R}_+^*$ tel que si $|z - z_0| = |\phi(t) - \phi(t_0)| \leqslant \eta_2$ alors $|f(z) - f(z_0)| = |f(\phi(t)) - f(\phi(t_0))| \leqslant \varepsilon$. D'après l'assertion (10) considérée avec $\varepsilon_1 = \eta_2$, il existe $\eta_1 \in \mathbb{R}_+^*$ tel que si $|t - t_0| \leqslant \eta_1$ alors $|\phi(t) - \phi(t_0)| \leqslant \eta_2$. On en déduit que si $t \in]a,b[$ est tel que $|t - t_0| \leqslant \eta_1$ alors $|f(\phi(t)) - f(\phi(t_0))| \leqslant \varepsilon$. L'assertion (12) est démontrée.

Solution de l'exercice 4.5

1 - Si z et z_0 désignent deux complexes distincts et n un entier supérieur ou égal à 2 on a :

$$f(z) - f(z_0) = z^n - z_0^n = (z - z_0)(z^{n-1} + z_0 z^{n-2} + \ldots + z_0^{n-2} z + z_0^{n-1}).$$

On en déduit que :

$$\Delta_{z_0}(z) = \frac{z^n - z_0^n}{z - z_0} = z^{n-1} + z_0 z^{n-2} + \ldots + z_0^{n-2} z + z_0^n$$

et $\lim_{z \to z_0} \Delta_{z_0}(z) = n z_0^{n-1}$.

L'application f est donc holomorphe sur $\mathbb{C}$ et admet pour dérivée l'application $f' : z \in \mathbb{C} \mapsto n z^{n-1}$. Les cas $n = 0$ et $n = 1$ sont triviaux.

2 - Pour $z_0 \in \mathbb{C}$, la quantité $\Delta_{z_0}(z) = \frac{|z| - |z_0|}{z - z_0}$ n'a pas de limite lorsque z tend vers z_0, voir l'exercice 4.3. On en déduit que l'application $z \in \mathbb{C} \mapsto |z|$ n'est holomorphe nulle part $z_0 \in \mathbb{C}$.

Solution de l'exercice 4.6

Pour tout $z \in \mathbb{C}$, d'après la proposition 4.20, on a

$$\exp(z) = \mathrm{e}^x \cos(y) + \mathrm{i}\mathrm{e}^x \sin(y).$$

On en déduit que le complexe $\exp(z)$ de partie réelle x et de partie imaginaire y est un réel si et seulement si $\mathrm{e}^x \sin(y) = 0$. Comme la fonction exponentielle ne s'annule pas sur $\mathbb{R}$, cette relation est vraie si et seulement s'il existe $k \in \mathbb{Z}$ tel que $y = k\pi$. Les complexes pour lesquels $\exp(z)$ est un réel sont donc tous les complexes dont la partie imaginaire est multiple de π.

Par ailleurs, $\exp(z)$ est un imaginaire pur si et seulement si $\mathrm{e}^x \cos(y) = 0$, autrement dit, si $\cos(y) = 0$. On en déduit que $\exp(z)$ est un imaginaire pur si et seulement s'il existe $k \in \mathbb{Z}$ tel que $\mathcal{I}\mathrm{m}(z) = (2k+1)\pi/2$.

Solution de l'exercice 4.7

1 - Pour tout $z \in \mathbb{C}$, on a

$$\begin{aligned} |\exp(z)|^2 &= \exp(z) \times \overline{\exp(z)} = \exp(z) \times \exp(\overline{z}) \\ &= \exp(z + \overline{z}) = \exp(2\mathcal{R}\mathrm{e}(z)) = \exp(\mathcal{R}\mathrm{e}(z))^2 = \left(\mathrm{e}^{\mathcal{R}\mathrm{e}(z)}\right)^2. \end{aligned}$$

De plus, comme $|\exp(z)|$ et $e^{\mathcal{R}e(z)}$ sont deux réels positifs, la relation $|\exp(z)|^2 = \left(e^{\mathcal{R}e(z)}\right)^2$ est équivalente à $|\exp(z)| = e^{\mathcal{R}e(z)}$.

2 - La relation précédente indique que $|\exp(z)| = 1$ si et seulement si $e^{\mathcal{R}e(z)} = 1$. D'après les propriétés de la fonction exponentielle réelle, ceci équivaut à $\mathcal{R}e(z) = 0$, autrement dit à ce que z soit un imaginaire pur.

Solution de l'exercice 4.8

1 - D'après la définition de la fonction sinus complexe, pour tout $z \in \mathbb{C}$ on a $\sin(z) = 0$ si et seulement si $\exp(-iz) = \exp(iz)$. Désignons par x la partie réelle de z et par y sa partie imaginaire. On a $\exp(-iz) = \exp(iz)$ si et seulement si

$$e^y \cos(x) - ie^y \sin(x) = e^{-y} \cos(x) + ie^{-y} \sin(x). \tag{13}$$

Deux complexes égaux ayant même module, on déduit de la relation (13) que $e^y = e^{-y}$. Ceci implique que $e^{2y} = 1$ donc que $y = 0$. De la relation (13) on déduit alors en identifiant les parties imaginaires que $\sin(x) = \sin(x)$ (on a toujours égalité des parties réelles). On a donc $2\sin(x) = 0$ et par conséquent les réels x solutions sont tous les réels de la forme $x = k\pi$ avec $k \in \mathbb{Z}$. Finalement l'ensemble des solutions de l'équation $\sin(z) = 0$ est $S = \{k\pi \ ; \ k \in \mathbb{Z}\}$.

2 - Effectuons un raisonnement analogue pour déterminer l'ensemble des solutions dans $\mathbb{C}$ de l'équation $\mathrm{sh}(z) = 0$. D'après la définition de la fonction sinus hyperbolique complexe, pour tout $z \in \mathbb{C}$ on a $\mathrm{sh}(z) = 0$ si et seulement si $\exp(z) = \exp(-z)$. Désignons par x la partie réelle de z et par y sa partie imaginaire. On a $\exp(z) = \exp(-z)$ si et seulement si

$$e^x \cos(y) + ie^x \sin(y) = e^{-x} \cos(y) - ie^{-x} \sin(y). \tag{14}$$

Deux complexes égaux ayant même module, on déduit de la relation (14) que $e^x = e^{-x}$ ce qui implique que $x = 0$. De la relation (14) on déduit alors en identifiant les parties réelles que $\sin(y) = -\sin(y)$ autrement dit que $\sin(y) = 0$. Par conséquent on a $y = k\pi$ avec $k \in \mathbb{Z}$. L'ensemble des solutions de l'équation $\mathrm{sh}(z) = 0$ est donc $S = \{ik\pi \ ; \ k \in \mathbb{Z}\}$.

3 - D'après la définition de la fonction cosinus complexe, pour tout $z \in \mathbb{C}$ on a $\cos(z) = i$ si et seulement si $\exp(iz) + \exp(-iz) = 2i$. Désignons par x la partie réelle de z et par y sa partie imaginaire. On a

$$\begin{aligned} \exp(iz) + \exp(-iz) &= e^{-y}(\cos(x) + i\sin(x)) + e^y(\cos(x) - i\sin(x)) \\ &= \cos(x)(e^{-y} + e^y) + i\sin(x)(e^{-y} - e^y) \\ &= 2\cos(x)\,\mathrm{ch}(y) - 2i\sin(x)\,\mathrm{sh}(y). \end{aligned}$$

On en déduit que l'on a $\cos(z) = \mathrm{i}$ si et seulement si $\cos(x)\operatorname{ch}(y) = 0$ et $\sin(x)\operatorname{sh}(y) = -1$. Intéressons-nous à la première des deux conditions ; on a $\cos(x)\operatorname{ch}(y) = 0$ si et seulement si $x = \frac{\pi}{2} + k\pi$, $k \in \mathbb{Z}$ car la fonction cosinus hyperbolique réelle est minorée par 1. Pour $x = \frac{\pi}{2} + k\pi$ avec $k \in \mathbb{Z}$, la seconde relation devient $(-1)^{|k|}\operatorname{sh}(y) = -1$. Si l'entier naturel $|k|$ est impair, on obtient $y = \operatorname{argsh}(1)$ et si $|k|$ est pair, on obtient $y = \operatorname{argsh}(-1)$. Comme la fonction argsh est impaire, on a $\operatorname{argsh}(-1) = -\operatorname{argsh}(1)$ et on peut regrouper les deux cas en exprimant la solution y sous la forme $y = (-1)^{|k|+1}\operatorname{argsh}(1)$. Finalement, on en conclut que les solutions dans $\mathbb{C}$ de l'équation $\cos(z) = \mathrm{i}$ sont les complexes $z = \frac{\pi}{2} + k\pi - \mathrm{i}(-1)^{|k|}\operatorname{argsh}(1)$ avec $k \in \mathbb{Z}$.

Solution de l'exercice 4.9

1 - a) f_1 et f_2 sont définies en tout $z \in \mathbb{C}$ tel que $\operatorname{sh}(z) \neq 0$. Or on a vu dans l'exercice 4.8 que $\operatorname{sh}(z) = 0$ si et seulement si $z = \mathrm{i}k\pi$ avec $k \in \mathbb{Z}$. Les fonctions f_1 et f_2 sont donc définies sur $\mathbb{C} \setminus \mathrm{i}\pi\mathbb{Z}$.

b) f_3 est définie en tout $z \in \mathbb{C}$ telle que $\mathcal{R}\mathrm{e}(sh(z)) \neq 0$. Soit z un complexe de partie réelle x et de partie imaginaire y. On a d'après le formulaire de trigonométrie hyperbolique complexe :

$$\operatorname{sh}(z) = \operatorname{sh}(x + \mathrm{i}y) = \operatorname{sh}(x)\operatorname{ch}(\mathrm{i}y) + \operatorname{ch}(x)\operatorname{sh}(\mathrm{i}y).$$

Comme $\operatorname{ch}(\mathrm{i}y) = \cos(\mathrm{i}^2 y) = \cos(y)$ et $\operatorname{sh}(\mathrm{i}y) = -\mathrm{i}\sin(\mathrm{i}^2 y) = \mathrm{i}\sin(y)$, on a $\operatorname{sh}(z) = \operatorname{sh}(x)\cos(y) + \mathrm{i}\sin(y)\operatorname{ch}(x)$. On a par conséquent $\mathcal{R}\mathrm{e}(sh(z)) = 0$ si et seulement si $\operatorname{sh}(x)\cos(y) = 0$, c'est à dire si $x = 0$ ou si $y = \frac{\pi}{2} + k\pi$, $k \in \mathbb{Z}$. L'ensemble de définition de f_3 est donc $\mathbb{C} \setminus E$ où

$$E = \{z \in \mathbb{C} \ ; \ \mathcal{R}\mathrm{e}(z) = 0\} \cup \{z \in \mathbb{C} \ ; \ \exists k \in \mathbb{Z} \ \mathcal{I}\mathrm{m}(z) = \tfrac{\pi}{2} + k\pi\}.$$

2 - a) Considérons l'arc paramétré $\phi : t \in \mathbb{R} \mapsto t \in \mathbb{C}$ dont l'image est l'axe des réels. On a $\phi(0) = 0$ et pour $t \in \mathbb{R}^*$, on obtient

$$f_1(\phi(t)) = \frac{\sin(t)}{\operatorname{sh}(t)} = \frac{\sin(t)}{t}\left(\frac{\operatorname{sh}(t)}{t}\right)^{-1}$$

Comme $\lim\limits_{\substack{t\to 0\\ t\in\mathbb{R}}} \frac{\sin(t)}{t} = \lim\limits_{\substack{t\to 0\\ t\in\mathbb{R}}} \frac{\operatorname{sh}(t)}{t} = 1$, on en déduit que $\lim\limits_{\substack{t\to 0\\ t\in\mathbb{R}}} f_1(\phi(t)) = 1$ et par conséquent que si f_1 a une limite en 0 ce ne peut être que 1. Considérons à présent l'arc paramétré $\phi : t \in]-1, 1[\mapsto \mathrm{i}\pi t$. On a $\phi(0) = 0$ et pour $t \in]-1, 0[\cup]0, 1[$ on obtient

$$f_1(\phi(t)) = \frac{\sin(\mathcal{R}\mathrm{e}(\mathrm{i}\pi t))}{\operatorname{sh}(\mathrm{i}\pi t)} = 0.$$

On en déduit que $\lim\limits_{\substack{t\to 0\\ t\in\mathbb{R}}} f_1(\phi(t)) = 0$ ce qui permet de conclure que f_1 n'a pas de limite en 0.

b) Considérons l'arc paramétré $\phi : t \in \mathbb{R} \mapsto t \in \mathbb{C}$ dont l'image est l'axe des réels. On a $\phi(0) = 0$ et pour $t \in \mathbb{R}^*$ on obtient

$$f_2(\phi(t)) = \frac{\sin(t)}{\operatorname{sh}(t)} = \frac{\sin(t)}{t}\left(\frac{\operatorname{sh}(t)}{t}\right)^{-1}$$

Comme $\lim_{\substack{t\to 0\\ t\in\mathbb{R}}} \frac{\sin(t)}{t} = \lim_{\substack{t\to 0\\ t\in\mathbb{R}}} \frac{\mathrm{sh}(t)}{t} = 1$, on en déduit que $\lim_{\substack{t\to 0\\ t\in\mathbb{R}}} f_2(\phi(t)) = 1$ et par conséquent que si f_2 a une limite en 0 ce ne peut être que 1. Considérons à présent l'arc paramétré $\phi : t \in]-1,1[\mapsto \mathrm{i}\pi t$. On a $\phi(0) = 0$ et pour $t \in]-1,0[\cup]0,1[$ on obtient

$$f_2(\phi(t)) = \frac{\mathcal{R}\mathrm{e}(\sin(\mathrm{i}\pi t))}{\mathrm{sh}(\mathrm{i}\pi t)}.$$

Or $\sin(\mathrm{i}\pi t) = -\mathrm{i}\,\mathrm{sh}(\mathrm{i}^2\pi t) = \mathrm{i}\,\mathrm{sh}(\pi t)$ est un imaginaire pour donc $f_2(\phi(t)) = 0$. On en déduit que $\lim_{\substack{t\to 0\\ t\in\mathbb{R}}} f_2(\phi(t)) = 0$ ce qui permet de conclure que f_2 n'a pas de limite en 0.

c) Considérons l'arc paramétré $\phi : t \in \mathbb{R} \mapsto t \in \mathbb{C}$ dont l'image est l'axe des réels. On a $\phi(0) = 0$ et pour $t \in \mathbb{R}^*$ on obtient

$$f_3(\phi(t)) = \frac{\sin(t)}{\mathrm{sh}(t)} = \frac{\sin(t)}{t}\left(\frac{\mathrm{sh}(t)}{t}\right)^{-1}$$

Comme $\lim_{\substack{t\to 0\\ t\in\mathbb{R}}} \frac{\sin(t)}{t} = \lim_{\substack{t\to 0\\ t\in\mathbb{R}}} \frac{\mathrm{sh}(t)}{t} = 1$, on en déduit que $\lim_{\substack{t\to 0\\ t\in\mathbb{R}}} f_3(\phi(t)) = 1$ et par conséquent que si f_3 a une limite en 0 ce ne peut être que 1.

Considérons à présent l'arc paramétré $\phi : t \in]-1,1[\mapsto t + \mathrm{i}t$. On a $\phi(0) = 0$ et pour $t \in]-1,,1[$ on obtient

$$f_3(\phi(t)) = \frac{\sin(t+\mathrm{i}t)}{\mathcal{R}\mathrm{e}(\mathrm{sh}(t+\mathrm{i}t))}.$$

Or $\mathrm{sh}(t+\mathrm{i}t) = \mathrm{sh}(t)\,\mathrm{ch}(\mathrm{i}t) + \mathrm{ch}(t)\,\mathrm{sh}(\mathrm{i}t) = \mathrm{sh}(t)\cos(t) + \mathrm{i}\sin(t)\,\mathrm{ch}(t)$
et $\sin(t+\mathrm{i}t) = \sin(t)\cos(\mathrm{i}t) + \cos(t)\sin(\mathrm{i}t) = \sin(t)\,\mathrm{ch}(t) + \mathrm{i}\cos(t)\,\mathrm{sh}(t)$ donc

$$f_3(\phi(t)) = \frac{\sin(t)\,\mathrm{ch}(t)}{\mathrm{sh}(t)\cos(t)} + \mathrm{i}.$$

On en déduit que $\lim_{\substack{t\to 0\\ t\in\mathbb{R}}} f_3(\phi(t)) = 1 + \mathrm{i}$ ce qui permet de conclure que f_3 n'a pas de limite en 0.

Solution de l'exercice 4.10

1 - a) Soit z un complexe de partie réelle x et de partie imaginaire y. On a

$$\sin(z) = \sin(x+\mathrm{i}y) = \sin(x)\cos(\mathrm{i}y) + \sin(\mathrm{i}y)\cos(x).$$

Or $\cos(\mathrm{i}y) = \frac{1}{2}\left(\exp(\mathrm{i}^2 y) + \exp(-\mathrm{i}^2 y)\right) = \frac{1}{2}\left(\exp(-y) + \exp(y)\right) = \mathrm{ch}(y)$ et $\sin(\mathrm{i}y) = \mathrm{i}\,\mathrm{sh}(y)$ donc $\sin(z) = \sin(x)\,\mathrm{ch}(y) + \mathrm{i}\,\mathrm{sh}(y)\cos(x)$. On en déduit que :

$$|\sin(z)|^2 = \sin(x)^2\,\mathrm{ch}(y)^2 + \mathrm{sh}(y)^2\cos(x)^2 = \mathrm{ch}(y)^2 - \cos(x)^2 = \sin(x)^2 + \mathrm{sh}(y)^2,$$

les deux dernières égalités étant obtenues grâce aux relations $\mathrm{ch}(y)^2 - \mathrm{sh}(y)^2 = 1$ et $\sin(x)^2 + \cos(x)^2 = 1$.

b) Comme $\cos(x)^2 \geqslant 0$, on en déduit que $|\sin(z)|^2 \leqslant \operatorname{ch}(y)^2$ et comme $\sin(x)^2 \geqslant 0$, on en déduit que $|\sin(z)|^2 \geqslant \operatorname{sh}(y)^2$. On en conclut que pour tout $z \in \mathbb{C}$ on a $|\operatorname{sh}(y)| \leqslant |\sin(z)| \leqslant |\operatorname{ch}(y)|$ et comme la fonction ch est positive, on a $|\operatorname{ch}(y)| = \operatorname{ch}(y)$.

c) Pour montrer que la fonction sinus complexe n'est pas bornée sur $\mathbb{C}$, il faut montrer que[36]

$$\forall M \in \mathbb{R}_+^* \quad \exists z \in \mathbb{C} \quad |\sin(z)| > M. \tag{15}$$

Soit M un réel strictement positif fixé et z le complexe $2\mathrm{i}\operatorname{argsh}(M)$. On a $|\sin(z)| \geqslant |\operatorname{sh}(\mathcal{I}\mathrm{m}(z))| = |\operatorname{sh}(2\operatorname{argsh}(M))|$. Comme M est strictement positif, $\operatorname{argsh}(M)$ est lui aussi strictement positif et $2\operatorname{argsh}(M) > \operatorname{argsh}(M)$. Comme la fonction sinus hyperbolique est strictement croissante et à valeurs positives sur $\mathbb{R}_+^*$, on en déduit que :

$$|\operatorname{sh}(2\operatorname{argsh}(M))| = \operatorname{sh}(2\operatorname{argsh}(M)) > \operatorname{sh}(\operatorname{argsh}(M)) = M.$$

L'assertion est établie pour tout $M \in \mathbb{R}_+^*$. Si $M = 0$, le complexe $z = \frac{\pi}{2}$ vérifie $|\sin(z)| = 1 > 0 = M$. On peut donc en conclure que la fonction sinus complexe n'est pas bornée sur $\mathbb{C}$. Comme pour tout $z \in \mathbb{C}$ on a $\cos(z)^2 = 1 - \sin(z)^2$, on en déduit en utilisant la seconde inégalité triangulaire que

$$|\cos(z)|^2 = |\cos(z)^2| = |1 - \sin(z)^2| \geqslant |1 - |\sin(z)|^2|.$$

Soit $M' \in \mathbb{R}_+^*$; d'après l'assertion (15) considérée avec le réel $M = \sqrt{M'} + 1$, on peut trouver un complexe z_M tel que $|\sin(z_M)| > \sqrt{M'+1}$. On a alors $|\cos(z_M)|^2 \geqslant |1 - |\sin(z_M)|^2| > M'$. On en conclut que la fonction cosinus complexe n'est pas bornée elle non plus sur $\mathbb{C}$.

> Même si une partie des propriétés connues des fonctions circulaires réelles restent valables pour les fonctions circulaires complexes (par exemple les identités trigonométriques), on sera vigilant au fait que les fonctions circulaires complexes ne possèdent pas toutes les propriétés des fonctions circulaires réelles. Ainsi comme on vient de le voir, alors que les fonctions circulaires réelles sont bornées, les fonctions circulaires complexes ne le sont pas.

d) Pour tout $z \in \mathbb{C}$, on a

$$|\cos(z)|^2 + |\sin(z)|^2 = |\cos(z)^2| + |\sin(z)|^2 = |1 - \sin^2(z)| + |\sin(z)|^2.$$

Or $\sin(z) = \sin(x)\operatorname{ch}(y) + \mathrm{i}\operatorname{sh}(y)\cos(x)$ donc

$$\begin{aligned} 1 - \sin^2(z) &= \left(1 - \sin(x)^2\operatorname{ch}(y)^2 + \operatorname{sh}(y)^2\cos(x)^2\right) + 2\mathrm{i}\sin(x)\operatorname{ch}(y)\operatorname{sh}(y)\cos(x) \\ &= \left(\cos(x)^2\operatorname{ch}(y)^2 - \sin(x)^2\operatorname{sh}(y)^2\right) + 2\mathrm{i}\sin(x)\operatorname{ch}(y)\operatorname{sh}(y)\cos(x) \end{aligned}$$

[36] Cette assertion quantifiée est la négation de l'assertion quantifiée exprimant que la fonction sinus complexe est bornée sur $\mathbb{C}$: $\exists M \in \mathbb{R}_+ \quad \forall z \in \mathbb{C} \quad |\sin(z)| \leqslant M$.

(on a utilisé la relation $1 = \cos(x)^2 + \sin(x)^2$ et factorisé). On en déduit que

$$\begin{aligned}|1 - \sin^2(z)|^2 &= \left(\cos(x)^2 \operatorname{ch}(y)^2 - \sin(x)^2 \operatorname{sh}(y)^2\right)^2 + \left(2\sin(x)\operatorname{ch}(y)\operatorname{sh}(y)\cos(x)\right)^2 \\ &= \left(\cos(x)^2 \operatorname{ch}(y)^2 + \operatorname{sh}(y)^2 \sin(x)^2\right)^2.\end{aligned}$$

On a également établi que $|\sin(z)|^2 = \sin(x)^2 \operatorname{ch}(y)^2 + \operatorname{sh}(y)^2 \cos(x)^2$ ce qui permet d'obtenir

$$\begin{aligned}|\cos(z)|^2 + |\sin(z)|^2 &= \cos(x)^2 \operatorname{ch}(y)^2 + \operatorname{sh}(y)^2 \sin(x)^2 \\ &\quad + \sin(x)^2 \operatorname{ch}(y)^2 + \operatorname{sh}(y)^2 \cos(x)^2 \\ &= \operatorname{ch}(y)^2 + \operatorname{sh}(y)^2 = \operatorname{ch}(2y).\end{aligned}$$

Comme la fonction cosinus hyperbolique réelle est minorée par 1, on en déduit que pour tout $z \in \mathbb{C}$ on a $|\cos(z)|^2 + |\sin(z)|^2 \geqslant 1$.

La relation précédente nous indique que $|\cos(z)|^2 + |\sin(z)|^2 = 1$ si et seulement si $\operatorname{ch}(2y) = 1$ c'est-à-dire si $y = 0$. L'égalité a donc lieu uniquement si z est réel.

2 - a) Montrons que la série de fonctions $\sum_n \frac{1}{n^2} \sin\left(z/n^2\right)$ converge absolument dans $\mathbb{C}$. Pour tout complexe z de partie réelle x et de partie imaginaire y, on a compte tenu de ce qui a été établi à la question précédente :

$$\left|\frac{1}{n^2} \sin\left(\frac{z}{n^2}\right)\right| \leqslant \frac{1}{n^2} \operatorname{ch}\left(\frac{y}{n^2}\right).$$

Comme $\lim_{n\to+\infty} \operatorname{ch}\left(y/n^2\right) = 1$ on en déduit que $\frac{1}{n^2} \operatorname{ch}\left(\frac{y}{n^2}\right) \underset{n\to+\infty}{\sim} \frac{1}{n^2}$. Comme la série $\sum_n \frac{1}{n^2}$ converge, on peut affirmer que la série $\sum_n \frac{1}{n^2} \sin\left(\frac{z}{n^2}\right)$ converge absolument.

b) La série $\sum_n f_n$ ne convergera pas normalement sur $\mathbb{C}$ si la série numérique de terme général $\sup_{z\in\mathbb{C}} |f_n(z)|$ converge. On a

$$\sup_{z\in\mathbb{C}} |f_n(z)| = \sup_{z\in\mathbb{C}} \left|\frac{1}{n^2} \sin\left(\frac{z}{n^2}\right)\right| = \frac{1}{n^2} \sup_{\zeta\in\mathbb{C}} |\sin(\zeta)|.$$

Comme la fonction sinus n'est pas bornée dans $\mathbb{C}$, la série $\sum_n f_n$ ne converge pas normalement sur $\mathbb{C}$.

Considérons à présent la convergence normale sur la boule $D(0,R)$, $R \in \mathbb{R}_+^*$. On a

$$\sup_{z\in D(0,R)} \left|\frac{1}{n^2} \sin\left(\frac{z}{n^2}\right)\right| = \frac{1}{n^2} \sup_{z\in D(0,\frac{R}{n^2})} |\sin(\zeta)| \leqslant \frac{1}{n^2} \sup_{y\in]-\frac{R}{n^2},\frac{R}{n^2}[} \operatorname{ch}(y).$$

Comme la fonction cosinus hyperbolique réelle est paire et strictement croissante sur $\mathbb{R}^+$, on a

$$\sup_{z\in D(0,R)} |f_n(z)| = \frac{1}{n^2} \operatorname{ch}\left(\frac{R}{n^2}\right).$$

Or $\frac{1}{n^2}\operatorname{ch}\left(R/n^2\right) \underset{n\to+\infty}{\sim} \frac{1}{n^2}$ et la série $\sum_n 1/n^2$ converge. On en déduit donc que la série $\sum_n f_n$ converge normalement sur le disque $D(0,R)$ pour tout $R \in \mathbb{R}_+$.

c) Soit z un complexe fixé et $R = |z| + 1$. On a $z \in D(0,R)$ et comme la fonction sinus complexe est continue sur $\mathbb{C}$, pour tout $n \in \mathbb{N}^*$ les fonctions f_n sont continues sur $D(0,R)$. La convergence normale donc uniforme de la série $\sum_n f_n$ sur $D(0,R)$ permet d'affirmer que la fonction f somme de la série est continue en z. On en conclut que f est continue sur $\mathbb{C}$ puisqu'elle est continue en tout $z \in \mathbb{C}$.

CHAPITRE 5

Séries entières

Une série entière de la variable complexe (resp. de la variable réelle) est une série d'applications de $\mathbb{C}$ dans $\mathbb{C}$ (resp. de $\mathbb{R}$ dans $\mathbb{R}$) de terme général $f_n : z \mapsto a_n z^n$ où pour tout $n \in \mathbb{N}$, a_n désigne un nombre complexe (resp. un nombre réel) appelé *coefficient de rang n* de la série entière. Les séries entières sont donc des cas particuliers de séries de fonctions, et les résultats généraux concernant les séries de fonctions s'appliquent aux séries entières. La présence d'un chapitre spécifique consacré aux séries entières est motivée par les propriétés de convergence spécifiques que possèdent les séries entières et les liens étroits existant entre les séries entières de la variable complexe et les fonctions holomorphes.

Dans un premier temps, nous nous intéresserons dans ce chapitre exclusivement aux séries entières de la variable complexe, nous contentant dans un deuxième temps d'indiquer les différences notables existant pour les séries entières de la variable réelle. Nous terminerons ce chapitre avec une partie consacrée à l'utilisation des séries entières pour la résolution d'équations différentielles.

5.1 Série entière de la variable complexe

5.1.1 Généralités sur les séries entières

DÉFINITION 5.1 *On appelle* série entière *de la variable complexe (ou série entière complexe) toute série d'applications de terme général*

$$f_n : z \in \mathbb{C} \longmapsto a_n z^n \in \mathbb{C} \qquad \text{où } a_n \in \mathbb{C}.$$

On note $\sum_n a_n z^n$ la série entière dont le terme général est $f_n : z \in \mathbb{C} \mapsto a_n z^n$. Pour tout $n \in \mathbb{N}$, le complexe a_n est appelé *coefficient de rang n* de la série entière. Pour $r \in \mathbb{R}^+$, on note $D(0,r)$ l'ensemble $\{z \in \mathbb{C} \ ; \ |z| < r\}$.

Exemples

1. La série $\sum_n z^n$, qui est une série géométrique de raison z, est une série entière complexe dont tous les coefficients sont égaux à 1. Elle converge sur $D(0,1)$ et sa somme est l'application $z \in D(0,1) \mapsto (1-z)^{-1}$ (voir le chap. 1 p. 3).

2. $\sum_n z^{2n}$ est une série entière complexe dont les coefficients d'ordre pair valent 1 et les coefficients d'ordre impairs sont nuls. On parle de série entière *lacunaire* ; cette dénomination s'applique à toute série entière possédant un nombre infini de termes nuls.

3. La fonction polynomiale $p : z \in \mathbb{C} \mapsto \sum_{k=0}^{d} a_k z^k$ $(d \in \mathbb{N})$ définit une série entière complexe dont les coefficients d'ordre supérieur à d sont nuls et dont les coefficients d'ordre inférieur où égal à d sont égaux aux coefficients de même ordre du polynôme.

4. $\sum_n \frac{1}{n!} z^n$ est une série entière complexe convergente sur $\mathbb{C}$ dont la somme est la fonction exponentielle complexe, voir la section 4.7.1 du chap. 4.

5. $\sum_n \frac{(-1)^n}{(n!)^2} \left(\frac{z}{2}\right)^{2n}$ est une série entière complexe lacunaire qui converge sur $\mathbb{C}$ (voir l'exercice 5.4 p. 244) dont la somme est appelée *fonction de Bessel d'ordre* 0.

Lois de composition

On peut munir l'ensemble des séries entières $\Sigma_e(\mathbb{C},\mathbb{C})$ de deux lois de composition internes : somme $(+_{\Sigma_e})$ et produit $(\times_{\Sigma_e})$ définies de la manière suivante : pour toutes séries entières $\sum_n a_n z^n$ et $\sum_n b_n z^n$ on a

i) $(\sum_n a_n z^n) +_{\Sigma_e} (\sum_n b_n z^n) = \sum_n c_n z^n$ où $c_n = a_n + b_n$ pour tout $n \in \mathbb{N}$;

ii) $(\sum_n a_n z^n) \times_{\Sigma_e} (\sum_n b_n z^n) = \sum_n c_n z^n$ où $c_n = \sum_{k=0}^{n} a_k \times b_{n-k}$ pour tout $n \in \mathbb{N}$.

$(\Sigma_e(\mathbb{C},\mathbb{C}), +_{\Sigma_e}, \times_{\Sigma_e})$ est un anneau[1] commutatif. Le produit $\times_{\Sigma_e}$ est appelé *produit de Cauchy*, voir la définition 1.7 p. 38.

On peut également munir l'ensemble $\Sigma_e(\mathbb{C},\mathbb{C})$ d'une loi de composition externe notée $\cdot$ définie de la manière suivante : pour toute série entière $\sum_n a_n z^n$ et pour tout complexe λ on a :

iii) $\lambda \cdot (\sum_n a_n z^n) = \sum_n c_n z^n$ où $c_n = \lambda \times a_n$ pour tout $n \in \mathbb{N}$.

$(\Sigma_e(\mathbb{C},\mathbb{C}), +_{\Sigma_e}, \cdot)$ est un espace vectoriel[2] sur $\mathbb{C}$. Par la suite, on notera plus simplement $+$ et $\times$ les deux lois $+_{\Sigma_e}$ et $\times_{\Sigma_e}$, le contexte précisant la signification du symbole en tant que loi.

5.1.2 Rayon de convergence d'une série entière

En accord avec les définitions 1.2 p. 2 et 1.6 p. 28, on définit la convergence simple et la convergence absolue d'une série entière de la manière suivante.

---

[1] Voir la définition 2.36 p. 64 du *Cours de première année*.

[2] Voir le chap. 8 du *Cours de première année*.

DÉFINITION 5.2 (Convergence simple et convergence absolue)

✗ *On dit que la série entière* $\sum_n a_n z^n$ converge simplement *sur un sous-ensemble* Z *de* $\mathbb{C}$ *si pour tout* $z \in Z$ *la série numérique de terme général* $u_n = a_n z^n$ *converge. L'application* $\phi : z \in Z \mapsto \sum_{n=0}^{+\infty} a_n z^n$ *est alors appelée* somme de la série entière.

✗ *On dit que la série entière* $\sum_n a_n z^n$ converge absolument *sur* Z *si pour tout* $z \in Z$ *la série numérique de terme général* $u_n = |a_n z^n|$ *converge.*

Remarques

1. Il résulte des propriétés des séries numériques que si une série entière converge absolument sur un sous-ensemble Z de $\mathbb{C}$ alors elle converge simplement sur Z.

2. Une série entière $\sum_n a_n z^n$ converge simplement et converge absolument en $z = 0$ et a pour somme a_0. ◇

Le résultat suivant, connu sous le nom de lemme d'Abel, est l'outil avec lequel nous allons mener l'étude des séries entières.

LEMME 5.1 (Lemme d'Abel)

Soient $(a_n)_n$ *une suite de complexes et* z_0 *un complexe non nul. Si la suite numérique* $(u_n)_n$ *de terme général* $u_n = a_n z_0^n$ *est bornée*[3] *alors la série entière* $\sum_n a_n z^n$ *converge absolument sur le disque ouvert* $D(0, |z_0|)$.

Démonstration Supposons que la suite $(u_n)_n$ de terme général $u_n = a_n z_0^n$ soit bornée, *i.e.* supposons que :

$$\exists M \in \mathbb{R}^+ \quad \forall n \in \mathbb{N} \quad |a_n z_0^n| \leqslant M.$$

On a alors pour tout $z \in \mathbb{C}$ et pour tout $n \in \mathbb{N}$:

$$|a_n z^n| = |a_n\ z_0^n\ (z/z_0)^n| = |a_n\ z_0^n|\ |z/z_0|^n \leqslant M\ |z/z_0|^n .$$

Pour tout complexe z tel que $|z| < |z_0|$ (*i.e.* pour $z \in D(0, |z_0|)$), la série numérique de terme général $v_n = M\ |z/z_0|^n$ est une série géométrique de raison $|z/z_0| < 1$; elle est donc convergente. Les critères de comparaison pour les séries numériques à termes positifs (voir le théorème 1.2 p. 14) indiquent que la série numérique $\sum_n |a_n z^n|$ converge. On en conclut que la série entière $\sum_n a_n z^n$ converge absolument sur $D(0, |z_0|)$. □

COROLLAIRE 5.1 *Soient* $(a_n)_n$ *une suite de complexes et* z_0 *un complexe non nul. Si la série numérique* $\sum_n a_n z_0^n$ *converge alors la série entière* $\sum_n a_n z^n$ *converge absolument sur* $D(0, |z_0|)$.

[3] Voir la définition 5.3 p. 172 du *Cours de première année*.

Démonstration La convergence de la série numérique $\sum_n a_n z_0^n$ implique que la suite de terme général $a_n z_0^n$ admet pour limite 0 en $+\infty$, voir la proposition 1.2 p. 11. Or, toute suite numérique convergente est bornée[4]. La suite de terme général $a_n z_0^n$ est donc bornée et on conclut en utilisant le lemme d'Abel que la série entière $\sum_n a_n z^n$ converge absolument sur $D(0, |z_0|)$. □

Étant donné une suite $(a_n)_n$ de complexes, on désigne par E l'ensemble des réels positifs r tels que la suite numérique de terme général $a_n r^n$ est bornée. Cet ensemble est non vide car 0 est élément de E. De plus,

- ou bien l'ensemble E est majoré et dans ce cas il admet[5] une borne supérieure dans $\mathbb{R}$;
- ou bien l'ensemble E n'est pas majoré et dans ce cas il n'admet pas de borne supérieure dans $\mathbb{R}$.

DÉFINITION 5.3 (Rayon de convergence)
Soit $(a_n)_n$ une suite de complexes et E l'ensemble des réels positifs r tels que la suite numérique $(a_n r^n)_{n\in\mathbb{N}}$ soit bornée.

- *Si E est majoré alors on appelle* rayon de convergence *de la série entière $\sum_n a_n z^n$ la borne supérieure de l'ensemble E.*
- *Si l'ensemble E n'est pas majoré alors on convient que le rayon de convergence de la série entière $\sum_n a_n z^n$ est égal à $+\infty$.*

Remarques

1. De la définition 5.3, il résulte que le rayon de convergence d'une série entière est soit un réel positif, soit $+\infty$. On dit parfois que R est la borne supérieure dans[6] $\overline{\mathbb{R}}$ de l'ensemble E, englobant ainsi dans un même énoncé les deux possibilités.

$$\begin{aligned} R &= \sup\{r \in \mathbb{R} \ ; \ (a_n r^n)_{n\in\mathbb{N}} \text{ est bornée } \} \\ &= \sup\{r \in \mathbb{R} \ ; \ \exists M \in \mathbb{R}^+ \ \forall n \in \mathbb{N} \ \ |a_n r^n| \leqslant M\}. \end{aligned}$$

2. Si E est majoré, dire que R est la borne supérieure de l'ensemble E signifie[5] que :

$$(\forall r \in E \quad r \leqslant R) \quad \text{et} \quad (\forall \varepsilon \in \mathbb{R}_+^* \quad \exists r_0 \in E \quad r_0 > R - \varepsilon). \qquad (1)$$

◇

La nature d'une série entière est déterminée par la valeur du rayon de convergence comme l'indique le théorème suivant illustré figure 1.

---

[4] Voir la proposition 5.5 p. 173 du *Cours de première année*.

[5] Voir la proposition 3.2 p. 95 du *Cours de première année*.

[6] On rappelle, voir la section 3.3.4 p. 112 du *Cours de première année*, que $\overline{\mathbb{R}}$ est l'ensemble $\mathbb{R}$ auquel on adjoint les deux éléments $+\infty$ et $-\infty$ et que cet ensemble est appelé « droite numérique achevée ».

THÉORÈME 5.1 *Soit $\sum_n a_n z^n$ une série entière complexe de rayon de convergence R.*

- *Si $R = 0$ alors la série $\sum_n a_n z^n$ diverge pour tout $z \in \mathbb{C}^*$.*
- *Si $R \in]0, +\infty[$ alors la série $\sum_n a_n z^n$ converge absolument pour tout complexe z tel que $|z| < R$ et diverge pour tout complexe z tel que $|z| > R$.*
- *Si $R = +\infty$ alors la série $\sum_n a_n z^n$ converge absolument pour tout $z \in \mathbb{C}$.*

Démonstration $\trianglerighteq$ Supposons $R > 0$ et montrons que si z_1 est un complexe tel que $|z_1| < R$ alors la série $\sum_n a_n z_1^n$ converge absolument.

– Établissons tout d'abord l'existence d'un élément r_0 *dans* E tel que $|z_1| < r_0 \leqslant R$. Si $z_1 = 0$, le résultat est évident : $r_0 = R$ convient. Pour $z_1 \neq 0$, on peut appliquer la relation (1) avec $\varepsilon = R - |z_1|$: il existe un élément r_0 dans E tel que $r_0 > R - \varepsilon = |z_1|$ et $r_0 \leqslant R$.

– La suite numérique de terme général $a_n r_0^n$ est bornée (car $r_0 \in E$) donc d'après le lemme d'Abel, la série entière $\sum_n a_n z^n$ est absolument convergente sur $D(0, r_0)$. Elle converge en particulier en z_1. On en conclut que si $R \in \mathbb{R}_+^*$ alors pour tout $z_1 \in \mathbb{C}$ tel que $|z_1| < R$ la série $\sum_n a_n z_1^n$ est absolument convergente.

$\trianglerighteq$ Supposons $R \in \mathbb{R}^+$ et montrons que si z_1 est un complexe tel que $|z_1| > R$ alors la série $\sum_n a_n z_1^n$ diverge. Si la série $\sum_n a_n z_1^n$ convergeait absolument, la suite de terme général $a_n z_1^n$ tendrait vers 0 et donc serait bornée[7] ; cela impliquerait que $|z_1|$ appartient à E. On a une contradiction car $|z_1| > R$ où R désigne la borne supérieure de E. On en conclut que si $R \in \mathbb{R}^+$ alors pour tout $z_1 \in \mathbb{C}$ tel que $|z_1| > R$, la série $\sum_n a_n z_1^n$ diverge.

$\trianglerighteq$ Supposons que $R = +\infty$ et montrons que pour tout complexe z_1 la série numérique $\sum_n a_n z_1^n$ converge absolument. Puisque $R = +\infty$, l'ensemble E n'est pas majoré, donc[8] pour tout réel M, il existe un élément r_0 dans E (autrement dit un réel positif r_0 tel que la suite $(a_n r_0^n)_n$ soit bornée) vérifiant $r_0 > M$. En prenant $M = |z_1|$, on en déduit qu'il existe $r_0 \in \mathbb{R}^+$ tel que $r_0 > |z_1|$ et tel que la suite $(a_n r_0^n)_n$ soit bornée. Ainsi z_1 appartient au disque ouvert $D(0, r_0)$ et d'après le lemme d'Abel, la série $\sum_n a_n z_1^n$ converge absolument sur $D(0, r_0)$. On en conclut que si $R = +\infty$ alors pour tout $z_1 \in \mathbb{C}$, la série $\sum_n a_n z_1^n$ est absolument convergente.

$\trianglerighteq$ Pour conclure la démonstration du théorème, il suffit de remarquer qu'en $z = 0$ une série entière converge et a pour somme a_0. □

[7] Voir la proposition 5.5 p. 173 du *Cours de première année*.

[8] Par définition, l'ensemble E est majoré si : $\exists M \in \mathbb{R}\ \forall e \in E\ e \leqslant M$. On en déduit que l'ensemble E n'est pas majoré si : $\forall M \in \mathbb{R}\quad \exists e \in E\quad e > M$.

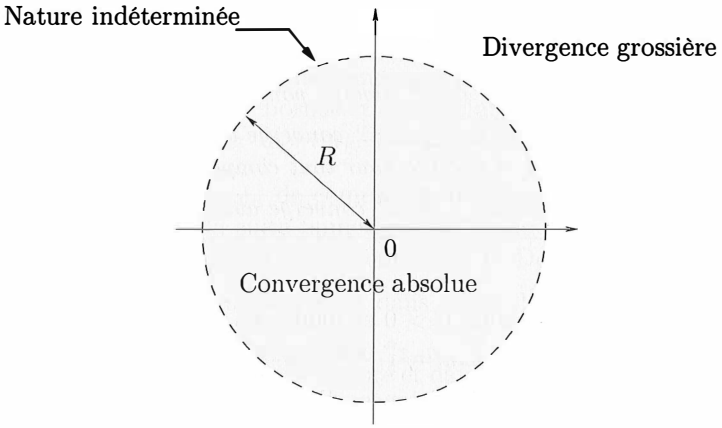


Fig. 1 Disque de convergence et nature d'une série entière complexe.

Remarques

1. On peut préciser le comportement d'une série entière de rayon de convergence R à l'extérieur du disque $\overline{D}(0,R)$. Considérons un complexe z_0 dans $\complement\overline{D}(0,R)$; la série numérique $\sum_n a_n z_0^n$ diverge. D'après le lemme d'Abel, la suite $(a_n z_0^n)_n$ n'est pas bornée. Cette suite ne converge donc pas vers 0 ; cela signifie que la série numérique $\sum_n a_n z_0^n$ diverge grossièrement. On en conclut qu'une série entière diverge grossièrement dans $\complement\overline{D}(0,R)$.

2. Le rayon de convergence d'une série entière $\sum_n a_n z^n$ vérifie

$$R \;=\; \sup\left\{r \in \mathbb{R}\,;\;\; \textstyle\sum_n a_n r^n \text{ converge}\right\}.$$

3. Soient $(a_n)_n$ et $(\alpha_n)_n$ deux suites de complexes dont tous les termes sont non nuls à partir d'un certain rang et telles que $|\alpha_n| \underset{+\infty}{\sim} |a_n|$. Les deux séries entières $\sum_n a_n z^n$ et $\sum_n \alpha_n z^n$ ont même rayon de convergence car pour tous $z \in \mathbb{C}^*$ on a $|\alpha_n z^n| \underset{+\infty}{\sim} |a_n z^n|$, ce qui d'après les critères de comparaison pour les séries numériques, voir le corollaire 1.2 p. 17, indique que les séries $\sum_n |a_n z^n|$ et $\sum_n |\alpha_n z^n|$ sont de même nature. ◇

Exemple Soit $\sum_n a_n z^n$ une série entière complexe de rayon de convergence R. Montrons que pour tout $p \in \mathbb{N}$ fixé, les séries entières $\sum_n a_n z^n$ et $\sum_n a_{p+n} z^n$ ont même rayon de convergence. Désignons par $(S_n)_n$et $(T_n)_n$ la suite des sommes partielles associée aux séries entières $\sum_n a_n z^n$ et $\sum_n a_{p+n} z^n$. Pour tous $z \in \mathbb{C}^*$ et $n \in \mathbb{N}$, on a

$$\begin{aligned} T_n(z) &= \sum_{k=0}^{n} a_{p+k} z^k = \frac{1}{z^p} \sum_{k=0}^{n} a_{p+k} z^{p+k} = \frac{1}{z^p} \sum_{k=p}^{n+p} a_k z^k \\ &= \frac{1}{z^p} \left(S_{n+p}(z) - S_{p-1}(z)\right). \end{aligned}$$

De cette relation, on peut déduire que si la suite $(S_n(z))_{n\in\mathbb{N}}$ converge vers $S(z)$ alors la suite $(T_n(z))_{n\in\mathbb{N}}$ converge vers $\frac{1}{z^p}\left(S(z) - S_{p-1}(z)\right)$. On peut en déduire aussi que si la suite $(T_n(z))_{n\in\mathbb{N}}$ converge vers $T(z)$ alors la suite $(S_n(z))_{n\in\mathbb{N}}$ converge vers $z^p T(z) + S_{p-1}(z)$. On en conclut que la suite $(T_n(z))_n$ converge si et seulement si la suite $(S_n(z))_n$ converge. Les séries entières $\sum_n a_n z^n$ et $\sum_n a_{p+n} z^n$ ont donc même ensemble de convergence, ce qui implique en particulier qu'elles ont même rayon de convergence.

DÉFINITION 5.4 (Disque de convergence)
Soit $\sum_n a_n z^n$ une série entière complexe de rayon de convergence R.

- *L'ensemble $D(0,R) = \{z \in \mathbb{C} \ ; \ |z| < R\}$ est appelé* disque de convergence *de la série entière.*
- *L'ensemble $C(0,R) = \{z \in \mathbb{C} \ ; \ |z| = R\}$ est appelé* cercle de convergence *de la série entière.*

D'après le théorème 5.1, une série entière de rayon de convergence R converge sur son disque de convergence $D(0,R)$ et diverge grossièrement sur $\complement\overline{D}(0,R)$ le complémentaire de l'adhérence du disque de convergence. On ne peut rien dire de général sur la nature d'une série entière sur le cercle de convergence $C(0,R)$. Elle peut selon les cas converger, diverger (grossièrement ou non), ou encore converger en certains points du cercle de convergence et diverger en d'autres.

Considérons à titre d'exemple la série entière $\sum_{n\geqslant 1} n^{-\alpha} z^n$ (où α désigne un réel) dont le rayon de convergence est égal à 1 (cette affirmation sera établie au cours de l'exercice 5.2). Selon les valeurs de α on a les comportements suivants sur le cercle de convergence $C(0,1)$:

- si $\alpha \in]-\infty, 0]$, la série diverge en tout point du cercle de convergence ;
- si $\alpha \in]0, 1]$, elle converge en tout point du cercle de convergence sauf en $z = 1$;
- si $\alpha \in]1, +\infty[$, elle converge en tout point du cercle de convergence.

On pourra également considérer l'exercice 5.1 pour un autre exemple illustrant cette mise en garde.

5.1.3 Détermination du rayon de convergence

Le théorème 5.1 permet d'obtenir plusieurs moyens de déterminer le rayon de convergence d'une série entière. Commençons par un résultat très utile dans la pratique.

PROPOSITION 5.1 *Soient $\sum_n a_n z^n$ et $\sum_n b_n z^n$ deux séries entières de rayon de convergence respectifs R_a et R_b. Si $|a_n| \underset{+\infty}{\sim} |b_n|$ alors $R_a = R_b$.*

Démonstration Puisque $|a_n| \underset{+\infty}{\sim} |b_n|$, pour tout $z \in \mathbb{C}^*$ on a $|a_n z^n| \underset{+\infty}{\sim} |b_n z^n|$. On en déduit[9] que les séries numériques réelles $\sum_n |a_n z^n|$ et $\sum_n |b_n z^n|$ sont de même nature : elles convergent (resp. divergent) toutes les deux pour les mêmes valeurs de z. Les séries entières $\sum_n a_n z^n$ et $\sum_n b_n z^n$ convergent donc absolument sur le même sous-ensemble de $\mathbb{C}$. Le théorème 5.1 indique que l'ensemble sur lequel une série entière converge absolument est le disque de convergence de la série, auquel s'ajoutent éventuellement certains points du cercle de convergence. On a donc $R_a = R_b$. □

Exemples

1. Le rayon de convergence de la série entière $\sum_{n\geqslant 1} \sin(\frac{1}{n})\, z^n$ est le même que celui de la série entière $\sum_{n\geqslant 1} \frac{1}{n} z^n$ car $\sin(\frac{1}{n}) \underset{+\infty}{\sim} \frac{1}{n}$.

2. Le rayon de convergence de la série entière $\sum_{n\geqslant 1} \frac{n^{n-1}-1}{n^n} z^n$ est le même que celui de la série entière $\sum_{n\geqslant 1} \frac{1}{n} z^n$ car $n^{n-1} - 1 \underset{+\infty}{\sim} n^{n-1}$ d'où $\frac{n^{n-1}-1}{n^n} \underset{+\infty}{\sim} \frac{1}{n}$.

Remarque Plus généralement, on peut, sur le principe de la démonstration de la proposition 5.1, déduire des résultats de comparaison pour les séries numériques à termes positifs (voir la section 1.2.2 du chap. 1) des résultats pour les rayons de convergence de séries entières. On montre par exemple que si $\sum_n a_n z^n$ et $\sum_n b_n z^n$ sont deux séries entières de rayon de convergence respectifs R_a et R_b telles que à partir d'un certain rang on ait $|a_n| \leqslant |b_n|$ alors $R_a \leqslant R_b$. ◇

Nous allons donner une autre utilisation du théorème 5.1 pour déterminer le rayon de convergence d'une série entière à travers deux exemples.

Exemples

1. Considérons la série entière $\sum_{n\geqslant 1} \frac{1}{n^2} z^n$.
– Si z_0 est tel que $|z_0| < 1$ alors la série numérique $\sum_{n\geqslant 1} \frac{1}{n^2} z_0^n$ converge absolument puisque $|\frac{1}{n^2} z_0^n| \leqslant \frac{1}{n^2}$ et que la série de Riemann $\sum_{n\geqslant 1} \frac{1}{n^2}$ converge. On a donc nécessairement $R \geqslant 1$.
– Si z_0 est tel que $|z_0| > 1$ alors $\lim_{n\to+\infty} \frac{1}{n^2} |z_0|^n = +\infty$ et la série numérique $\sum_{n\geqslant 1} \frac{1}{n^2} z_0^n$ diverge grossièrement. On a donc $R \leqslant 1$. On en conclut que $R = 1$.
Cette méthode est souvent utilisée pour déterminer le rayon de convergence d'une série entière en l'absence d'hypothèse permettant d'avoir recours à d'autres méthodes calculatoires comme celles qui vont être présentées dans la partie suivante.

2. Utilisons cette méthode pour montrer qu'étant donnée une suite complexes $(a_n)_n$, les séries entières $\sum_n a_n z^n$ et $\sum_n (n+1) a_{n+1} z^n$ ont même rayon de convergence. Désignons par R le rayon de convergence de $\sum_n a_n z^n$ et par R' le rayon de convergence de $\sum_n (n+1) a_{n+1} z^n$; il s'agit de montrer que $R = R'$.

[9] Voir le corollaire 1.2 p. 17.

– Soit z_0 un complexe n'appartenant pas au disque de convergence de la série entière $\sum_n a_n z^n$ (*i.e.* tel que $|z_0| > R$). La série numérique $\sum_n a_n z_0^n$ divergeant, la série numérique $\sum_n |a_n z_0^n|$ diverge également. Pour tout $n \in \mathbb{N}$ tel que $(n+1) \geqslant |z_0|$, on a

$$|(n+1)\, a_{n+1}\, z_0^n| = (n+1)\, |a_{n+1}\, z_0^n| \geqslant \left|a_{n+1}\, z_0^{n+1}\right|,$$

et les critères de comparaison[10] pour les séries numériques indiquent que la série $\sum_n |(n+1)a_{n+1} z_0^n|$ diverge. On en déduit que z_0 n'appartient pas au disque de convergence de la série entière $\sum_n (n+1)a_{n+1}\, z^n$ et on peut en conclure que $R' \leqslant R$.
– Soit maintenant z_0 un complexe appartenant au disque de convergence de $\sum_n a_n\, z^n$ (*i.e.* tel que $|z_0| < R$). Il existe un réel η strictement positif tel que $|z_0| + \eta < R$ (on peut par exemple prendre $\eta = \frac{1}{2}(R - |z_0|)$). On a[11]

$$(|z_0| + \eta)^{n+1} = \sum_{k=0}^{n+1} C_{n+1}^k |z_0|^k \eta^{n+1-k} \geqslant \eta(n+1)|z_0|^n.$$

On en déduit l'inégalité suivante :

$$(n+1)\, |a_{n+1}|\, |z_0|^n \leqslant \frac{1}{\eta}\, |a_{n+1}|\, (|z_0| + \eta)^{n+1}.$$

Comme $|z_0| + \eta < R$, la série numérique $\sum_n |a_{n+1}|(|z_0| + \eta)^{n+1}$ converge. Les critères de comparaison pour les séries numériques[11] impliquent que la série numérique $\sum_n (n+1)|a_{n+1}||z_0|^n$ converge. On en déduit que z_0 appartient au disque de convergence de la série entière $\sum_n (n+1)a_{n+1}\, z^n$ et on peut en conclure que $R' \geqslant R$.
Finalement des relations $R' \leqslant R$ et $R' \geqslant R$, on déduit que $R = R'$.

Nous avons par ailleurs les résultats suivants.

PROPOSITION 5.2 *Soit $\sum_n a_n\, z^n$ une série entière complexe de rayon de convergence R. Les séries entières*

$$\sum_n n a_n\, z^n, \qquad \sum_n (n+1) a_n\, z^n, \qquad \sum_{n \geqslant 1} \frac{1}{n}\, a_n\, z^n, \qquad \sum_n \frac{1}{n+1}\, a_n\, z^n$$

et $\displaystyle\sum_{n \geqslant 1} \frac{1}{n}\, a_{n+p}\, z^n$ (où $p \in \mathbb{N}$ est fixé) ont pour rayon de convergence R.

[10] Voir le théorème 1.2 p. 14.
[11] Tous les termes de la somme étant positifs, cette somme est supérieure au terme d'indice $k = n$.

Démonstration Nous avons établi en préambule à la proposition que si $(\alpha_n)_n$ désigne une suite de complexes alors les séries entières $\sum_n \alpha_n z^n$ et $\sum_n (n+1)\alpha_{n+1} z^n$ ont même rayon de convergence. Nous avons par ailleurs montré p. 236 que pour toute suite complexe $(\beta_n)_n$, les séries entières $\sum_n \beta_n z^n$ et $\sum_n \beta_{p+n} z^n$ ont même rayon de convergence.

$\trianglerighteq$ En prenant $\alpha_n = a_n$, on établit que les séries entières $\sum_n a_n z^n$ et $\sum_n (n+1)a_{n+1} z^n$ ont même rayon de convergence R. De plus, en prenant $\beta_n = na_n$ et $p = 1$, on établit que les séries entières $\sum_n (n+1)a_{n+1} z^n$ et $\sum_n na_n z^n$ ont même rayon de convergence R. Comme pour tout $z \neq 0$ on a $|(n+1)a_n z^n| \underset{+\infty}{\sim} |na_n z^n|$, la série entière $\sum_n (n+1)a_n z^n$ admet aussi pour rayon de convergence R.

$\trianglerighteq$ En posant cette fois-ci $\alpha_n = a_n$, on établit que les séries entières $\sum_{n\geqslant 1} \frac{1}{n} a_n z^n$ et $\sum_n a_{n+1} z^n$ ont même rayon de convergence. Comme le rayon de convergence de la série $\sum_n a_{n+1} z^n$ est le même que celui de la série $\sum_n a_n z^n$ (prendre $\beta_n = a_n$ et $p = 1$), on en déduit que le rayon de convergence de la série $\sum_{n\geqslant 1} \frac{1}{n} a_n z^n$ est R. De plus, puisque pour tout $z \neq 0$ on a $\left|\frac{a_n}{n+1} z^n\right| \underset{+\infty}{\sim} \left|\frac{a_n}{n} z^n\right|$, la série entière $\sum_n \frac{1}{n+1} a_n z^n$ admet aussi pour rayon de convergence R.

$\trianglerighteq$ Pour $p \in \mathbb{N}$ fixé et $z \in \mathbb{C}^*$ on a $\left|\frac{a_{n+p}}{n} z^n\right| \underset{+\infty}{\sim} \left|\frac{a_{n+p}}{n+p} z^n\right|$. Comme les séries entières $\sum_n \frac{1}{n+p} a_{n+p} z^n$ et $\sum_n \frac{1}{n} a_n z^n$ ont même rayon de convergence, on en déduit que la série entière $\sum_{n\geqslant 1} \frac{1}{n} a_{n+p} z^n$ a pour rayon de convergence R. □

On déduit également du théorème 5.1 la proposition suivante qui fournit un autre moyen de déterminer le rayon de convergence d'une série entière.

PROPOSITION 5.3 *Soit $(a_n)_n$ une suite de complexes. Si z_0 est un complexe tel que la série numérique $\sum_n a_n z_0^n$ est semi-convergente alors le rayon de convergence de la série entière $\sum_n a_n z^n$ est $R = |z_0|$.*

Démonstration La série numérique $\sum_n a_n z_0^n$ est semi-convergente donc elle converge mais ne converge pas absolument. Si elle converge, elle ne diverge pas donc d'après le théorème 5.1 le complexe z_0 n'appartient pas à $\complement\overline{D}(0, R)$. Si elle ne converge pas absolument, d'après ce même théorème, on peut en déduire que z_0 n'appartient pas à $D(0, R)$. On en conclut que nécessairement z_0 appartient au cercle de convergence $C(0, R)$. On a donc bien $R = |z_0|$. □

Exemple La série numérique $\sum_{n\geqslant 1} \frac{(-1)^n}{\sqrt{n}}$ est une série alternée qui n'est pas absolument convergente (car la série $\sum_{n\geqslant 1} \frac{1}{\sqrt{n}}$ est une série de Riemann divergente) mais qui satisfait au critère spécial pour les séries alternées[12]. C'est donc une série semi-convergente. On en déduit que le rayon de convergence de la série entière complexe $\sum_{n\geqslant 1} \frac{1}{\sqrt{n}} z^n$ est $R = 1$ (on a pris ici $z_0 = -1$).

[12] Voir le théorème 1.3 p. 42.

Exercice 5.1

1 - En utilisant les propositions 5.2 et 5.3, montrer que les trois séries entières suivantes ont pour rayon de convergence $R = 1$:

$$\sum_n z^n, \qquad \sum_{n\geqslant 1} \frac{z^n}{n}, \qquad \sum_{n\geqslant 1} \frac{z^n}{n^2}.$$

2 - En utilisant les résultats sur les séries numériques, montrer que la première ne converge en aucun point du cercle de convergence, que la seconde converge en tout point du cercle de convergence excepté $z = 1$ et que la troisième converge en tout point du cercle de convergence.

PROPOSITION 5.4 (Règle de Cauchy pour les séries entières)
Soit $\sum_n a_n z^n$ une série entière de rayon de convergence R telle que la suite de terme général $\sqrt[n]{|a_n|}$ admette une limite $\ell \in [0, +\infty[$ ou tende vers $+\infty$. On a les propriétés suivantes :

- *si $\ell = 0$ alors $R = +\infty$;*
- *si $\ell \in]0, +\infty[$ alors $R = 1/\ell$;*
- *si $\ell = +\infty$ alors $R = 0$.*

Démonstration $\trianglerighteq$ Pour tous $z \in \mathbb{C}$ et $n \in \mathbb{N}$, on a $\sqrt[n]{|a_n z^n|} = \sqrt[n]{a_n}\,|z|$. Par conséquent, si la suite de terme général $\sqrt[n]{|a_n|}$ admet une limite $\ell \in]0, +\infty[$ alors la suite $(u_n)_n$ de terme général $u_n = \sqrt[n]{|a_n z^n|}$ converge vers $\ell\,|z|$. D'après la règle de Cauchy pour les séries numériques[13], si la suite $(u_n)_n$ converge vers une limite k alors la série numérique $\sum_n |a_n z^n|$ converge absolument si $k < 1$ et diverge si $k > 1$. On en déduit que :

- si $\ell\,|z| < 1$ (*i.e.* si $|z| < 1/\ell$) alors la série $\sum_n a_n z^n$ converge absolument ;
- si $\ell\,|z| > 1$ (*i.e.* si $|z| > 1/\ell$) alors elle diverge.

D'après le théorème 5.1, le rayon de convergence de la série entière $\sum_n a_n z^n$ est $R = 1/\ell$.

$\trianglerighteq$ Si la suite de terme général $\sqrt[n]{|a_n|}$ admet pour limite 0 alors la suite $(u_n)_n$ de terme général $u_n = \sqrt[n]{|a_n z^n|}$ converge vers 0. D'après la règle de Cauchy pour les séries numériques, la série numérique $\sum_n |a_n z^n|$ converge quelle que soit la valeur de z ; on a donc $R = +\infty$.

$\trianglerighteq$ Si la suite de terme général $\sqrt[n]{|a_n|}$ tend vers $+\infty$ alors la suite $(u_n)_n$ de terme général $u_n = \sqrt[n]{|a_n z^n|}$ tend également vers $+\infty$ pour tout $z \neq 0$. D'après la règle de Cauchy pour les séries numériques, la série numérique $\sum_n |a_n z^n|$ diverge quel que soit $z \in \mathbb{C}^*$; on a donc $R = 0$. □

[13] Voir la proposition 1.12 p. 30.

Exemples

1. Le rayon de convergence de la série entière $\sum_{n\geqslant 1} \frac{1}{n^n} z^n$ est $+\infty$ car

$$\lim_{n\to+\infty} \sqrt[n]{n^{-n}} = \lim_{n\to+\infty} n^{-1} = 0.$$

2. Le rayon de convergence de la série entière $\sum_n \frac{1}{(1+\mathrm{i})^n} z^n$ est $\sqrt{2}$ car

$$\lim_{n\to+\infty} \sqrt[n]{|(1+\mathrm{i})^{-n}|} = \lim_{n\to+\infty} \sqrt[n]{2^{-\frac{1}{2}n}} = \frac{1}{\sqrt{2}}.$$

Exercice 5.2 En utilisant la règle de Cauchy, calculer le rayon de convergence des séries entières suivantes :

$$1 - \sum_{n\geqslant 1} n^{\sqrt{n}} z^n \qquad 2 - \sum_{n\geqslant 1} n^{\alpha} z^n \ (\alpha \in \mathbb{R}) \qquad 3 - \sum_{n\geqslant 1} \frac{n^{\ln(n)}}{\ln(n)^n} z^n$$

Proposition 5.5 (Règle de D'Alembert pour les séries entières) *Soit $\sum_n a_n z^n$ une série entière de rayon de convergence R dont les coefficients sont tous non nuls à partir d'un certain rang et telle que la suite de terme général $|a_{n+1}|/|a_n|$ admette une limite $\ell \in [0, +\infty[$ ou tende vers $+\infty$. On a les propriétés suivantes :*

- *si $\ell = 0$ alors $R = +\infty$;*
- *si $\ell \in]0, +\infty[$ alors $R = 1/\ell$;*
- *si $\ell = +\infty$ alors $R = 0$.*

Démonstration $\trianglerighteq$ Pour $z \in \mathbb{C}^*$, considérons la suite $(u_n)_n$ de terme général

$$u_n = \frac{|a_{n+1} z^{n+1}|}{|a_n z^n|} = \frac{|a_{n+1}|}{|a_n|}|z|.$$

Si la suite de terme général $|a_{n+1}|/|a_n|$ admet pour limite le réel $\ell \in]0, +\infty[$ alors la suite $(u_n)_n$ converge vers $\ell\,|z|$. D'après la règle de D'Alembert pour les séries numériques[(14)], si la suite $(u_n)_n$ converge vers une limite k alors la série numérique $\sum_n |a_n z^n|$ converge absolument si $k < 1$ et diverge si $k > 1$. On en déduit que :

- si $\ell\,|z| < 1$ (*i.e.* si $|z| < 1/\ell$), la série $\sum_n a_n z^n$ converge absolument ;
- si $\ell\,|z| > 1$ (*i.e.* si $|z| > 1/\ell$), elle diverge.

Le rayon de convergence de la série entière $\sum_n a_n z^n$ est donc $R = 1/\ell$.

$\trianglerighteq$ Si la suite de terme général $|a_{n+1}|/|a_n|$ admet pour limite 0, alors la suite $(u_n)_n$ converge vers 0. D'après la règle de D'Alembert pour les séries numériques, la série numérique $\sum_n |a_n z^n|$ converge quelle que soit la valeur de z ; on a donc $R = +\infty$.

(14) Voir la proposition 1.13 p. 34.

$\trianglerighteq$ Si la suite de terme général $|a_{n+1}|/|a_n|$ tend vers $+\infty$ alors la suite $(u_n)_n$ tend également vers $+\infty$. D'après la règle de D'Alembert pour les séries numériques, la série numérique $\sum_n |a_n z^n|$ diverge quel que soit $z \in \mathbb{C}^*$; on en déduit que $R = 0$. □

Exemples

1. Le rayon de convergence de la série entière $\sum_n \frac{1}{n!} z^n$ est $+\infty$ car :

$$\lim_{n\to+\infty} \frac{\frac{1}{(n+1)!}}{\frac{1}{n!}} = \lim_{n\to+\infty} \frac{1}{n+1} = 0.$$

2. Le rayon de convergence de la série entière $\sum_n \mathrm{i}^n z^n$ est 1 car :

$$\lim_{n\to+\infty} \left| \frac{\mathrm{i}^{n+1}}{\mathrm{i}^n} \right| = 1.$$

3. Le rayon de convergence de la série entière $\sum_n n!\, z^n$ est 0 car :

$$\lim_{n\to+\infty} \frac{(n+1)!}{n!} = \lim_{n\to+\infty} (n+1) = +\infty.$$

Remarque Comme dans le cas des séries numériques (voir la remarque p. 35), la règle de Cauchy s'applique dans tous les cas où la règle de D'Alembert s'applique. Toutefois selon l'expression du terme général de la série, la règle de D'Alembert peut être plus simple à appliquer que la règle de Cauchy (ce qui justifie l'énoncé des deux règles). Il existe des cas où la règle de Cauchy permet d'obtenir le rayon de convergence alors que la règle de D'Alembert ne le permet pas. ◇

Exercice 5.3 En utilisant la règle de D'Alembert, calculer le rayon de convergence des séries entières suivantes :

1 - $\sum_{n\geqslant 1} \frac{\mathrm{ch}(n)}{n} z^n$ 2 - $\sum_{n\geqslant 2} \ln(n)\, z^n$ 3 - $\sum_{n\geqslant 1} (\sqrt{n})^n z^n$

La proposition 5.5 ne s'applique pas aux séries entières lacunaires, c'est-à-dire aux séries ayant un nombre infini de termes nuls (la suite de terme général $|a_{n+1}|/|a_n|$ n'étant alors pas définie). Dans ce cas, on peut toutefois revenir au critère de D'Alembert pour les séries numériques et l'utiliser comme cela a été fait dans la démonstration de la proposition 5.5.

Considérons pour illustrer cette mise en garde la série entière $\sum_{n\geqslant 1} \frac{1}{n} z^{2n}$. Il s'agit d'une série entière lacunaire puisque ses coefficients sont donnés pour tout $m \in \mathbb{N}^*$ par :

$$a_m = \begin{cases} 2/m & \text{si } m \text{ est pair} \\ 0 & \text{si } m \text{ est impair} \end{cases}$$

Appliquons le critère de D'Alembert pour les séries numériques à la série $\sum_n \frac{1}{n} z^{2n}$, z étant fixé dans $\mathbb{C}^*$. Notant $u_n = \frac{1}{n} z^{2n}$, on obtient pour tout $n \in \mathbb{N}$:

$$\frac{|u_{n+1}|}{|u_n|} = \frac{n+1}{n} |z|^2.$$

Comme cette quantité tend vers $|z|^2$ quand n tend vers $+\infty$, on en déduit que
- si $|z|^2 < 1$ (*i.e.* si $|z| < 1$), la série numérique $\sum_n u_n$ converge absolument ;
- si $|z|^2 > 1$ (*i.e.* si $|z| > 1$), elle diverge.

Le rayon de convergence de la série entière $\sum_{n \geqslant 1} \frac{1}{n} z^{2n}$ est donc $R = 1$.

Dans le cas de cet exemple, on peut aussi procéder par changement de variable. Si l'on pose $Z = z^2$, on est ramené à la série entière non lacunaire $\sum_{n \geqslant 1} \frac{1}{n} Z^n$ pour laquelle on peut appliquer le critère de D'Alembert pour les séries entières. On a

$$R = \lim_{n \to +\infty} \left| \frac{\frac{1}{n+1}}{\frac{1}{n}} \right| = 1.$$

Le rayon de convergence de la série entière $\sum_{n \geqslant 1} \frac{1}{n} Z^n$ est donc $R = 1$. Compte tenu du changement de variable qui a été effectué, on en conclut que le rayon de convergence de la série entière $\sum_{n \geqslant 1} \frac{1}{n} z^{2n}$ est $\sqrt{R}$ et vaut donc 1.

Exercice 5.4 Déterminer le rayon de convergence de la série $\sum_n \frac{(-1)^n}{(n!)^2} \left(\frac{z}{2}\right)^{2n}$; en déduire l'ensemble de définition de la *fonction de Bessel d'ordre* 0 qui est la somme de cette série.

Nous avons évoqué p. 232 que $(\Sigma_e(\mathbb{C}, \mathbb{C}), +, \times)$ est un anneau commutatif. La proposition suivante précise les propriétés du rayon de convergence de la somme et du produit de deux séries entières.

PROPOSITION 5.6 *Soient $\sum_n a_n z^n$ et $\sum_n b_n z^n$ deux séries entières de rayon de convergence respectifs R_a et R_b.*
- *La série entière qui est la somme de $\sum_n a_n z^n$ et de $\sum_n b_n z^n$ admet un rayon de convergence R_s qui vérifie :*
  - *$R_s = \min(R_a, R_b)$ si $R_a \neq R_b$;*
  - *$R_s \geqslant \min(R_a, R_b)$ si $R_a = R_b$.*
- *La série entière qui est le produit de $\sum_n a_n z^n$ et de $\sum_n b_n z^n$ admet un rayon de convergence R_p qui vérifie :*

$$R_p \geqslant \min(R_a, R_b).$$

Démonstration $\trianglerighteq$ Supposons sans restreindre la généralité[15] que $R_a \geqslant R_b$. Les deux séries entières $\sum_n a_n z^n$ et $\sum_n b_n z^n$ convergeant absolument sur

[15] Les deux séries entières étant données, on peut toujours décider d'appeler R_a le plus grand des deux rayons, R_b le plus petit et $\sum_n a_n z^n$, $\sum_n b_n z^n$ les deux séries entières correspondantes.

$D(0, R_b)$, on déduit de l'inégalité $|a_n + b_n| \leqslant |a_n| + |b_n|$ et des critères de comparaison pour les séries à termes positifs[16] que la série $\sum_n c_n z^n$ de terme général $c_n = a_n + b_n$ converge absolument sur $D(0, R_b)$. On a donc $R_s \geqslant R_b$.

$\trianglerighteq$ Dans le cas où $R_a \neq R_b$, pour montrer que $R_s = R_b$, raisonnons par l'absurde et supposons que $R_s > R_b$. Il existe alors un complexe z_0 tel que $R_b < |z_0| < \min(R_s, R_a)$. Les séries $\sum_n a_n z_0^n$ et $\sum_n c_n z_0^n$ convergent et puisque $b_n z_0^n = c_n z_0^n - a_n z_0^n$, la série $\sum_n b_n z_0^n$ converge aussi. Cela est absurde car z_0 n'appartient pas au disque de convergence de la série entière $\sum_n b_n z^n$. On en conclut que si $R_a > R_b$ alors $R_s = R_b = \min(R_a, R_b)$.

$\trianglerighteq$ Dans le cas où $R_a = R_b$, on n'a pas nécessairement $R_s = \min(R_a, R_b)$ comme le prouve l'exemple suivant : les séries $\sum_n \frac{1}{n} z^n$ et $\sum_n \frac{-1}{n} z^n$ ont pour rayon de convergence 1 et leur somme qui est la série nulle a pour rayon de convergence $+\infty$.

$\trianglerighteq$ Le résultat concernant le rayon de convergence du produit de deux séries entières est admis. □

Remarque Les résultats donnés par la proposition 5.6 s'étendent au cas où les séries ont pour rayon de convergence $+\infty$. ◇

Exemples

1. Les séries $\sum_n \frac{1}{n} z^n$ et $\sum_n \frac{-1}{n^n} z^n$ ont pour rayon de convergence respectifs $R_a = 1$ et $R_b = +\infty$. On a $R_a \neq R_b$ et la série entière $\sum_n \frac{n^{n-1}-1}{n^n} z^n$ qui est la somme de ces deux séries entières a pour rayon de convergence $R_s = \min(R_a, R_b) = 1$.

2. Les séries $\sum_n (-1)^n z^n$ et $\sum_n z^n$ ont pour rayon de convergence $R_a = R_b = 1$. Leur somme qui est la série $\sum_n 2z^{2n}$ a un rayon de convergence R_s vérifiant $R_s \geqslant \min(R_a, R_b) = 1$. Comme par ailleurs la série entière lacunaire $\sum_n 2z^{2n}$ diverge en $z = 1$, on peut affirmer que son rayon de convergence vaut $R_s = 1$.

3. Les séries entières $\sum_{n \geqslant 1} \frac{1}{n} z^n$ et $\sum_{n \geqslant 1} \frac{1-n^{n-1}}{n^n} z^n$ ont pour rayon de convergence $R_a = R_b = 1$ et leur somme qui est la série $\sum_n \frac{1}{n^n} z^n$ a un rayon de convergence $R_s = +\infty$. On a donc $R_s > \min(R_a, R_b) = 1$.

4. Les séries entières $\sum_n \pi^n z^n$ et $\sum_n (1/\pi)^n z^n$ ont pour rayon de convergence respectifs $R_a = 1/\pi$ et $R_b = \pi$. La série produit est $\sum_n c_n z^n$ où

$$c_n = \sum_{k=0}^{n} \pi^{n-2k} = \pi^n \frac{\pi^2 - \pi^{-2n}}{\pi^2 - 1} = \frac{\pi^{n+2} - \pi^{-n}}{\pi^2 - 1}.$$

En utilisant la règle de D'Alembert, on établit que le rayon de convergence de cette série entière est $1/\pi = \min(R_a, R_b)$.

5. Les séries entières $\sum_n, z^n$ et $\sum_n \frac{1-n}{n!} z^n$ ont pour rayon de convergence respectifs $R_a = 1$ et $R_b = +\infty$. La série produit qui est $\sum_n \frac{1}{n!} z^n$ a pour rayon de convergence $R = +\infty > \min(R_a, R_b)$.

---

[16] Voir le théorème 1.2 p. 14.

5.1.4 Convergence uniforme d'une série entière

La notion de convergence uniforme pour une série entière complexe correspond à un cas particulier de convergence uniforme pour une série de fonctions de la variable complexe et la définition suivante n'est qu'une reformulation de la définition 3.4 p. 120 dans le cas particulier des séries entières.

DÉFINITION 5.5 (Convergence uniforme)
Soit $(a_n)_n$ une suite de complexes et Z un sous-ensemble de $\mathbb{C}$. On dit que la série entière $\sum_n a_n z^n$ converge uniformément sur Z s'il existe une application ϕ de Z dans $\mathbb{C}$ telle que :

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad \forall z \in Z$$
$$\left(n \geqslant N \implies \left|\sum_{k=0}^{n} a_k z^k - \phi(z)\right| \leqslant \varepsilon\right). \tag{2}$$

L'assertion (2) peut encore s'exprimer sous la forme :

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad \left(n \geqslant N \implies \sup_{z \in Z} \left|\sum_{k=0}^{n} a_k z^k - \phi(z)\right| \leqslant \varepsilon\right) \tag{3}$$

ou encore sous la forme :

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad \left(n \geqslant N \implies \sup_{z \in Z} \left|\sum_{k=n+1}^{+\infty} a_k z^k\right| \leqslant \varepsilon\right). \tag{4}$$

L'assertion (4) indique que la série entière $\sum_n a_n z^n$ converge uniformément sur Z si

$$\lim_{n \to +\infty} \sup_{z \in Z} \left|\sum_{k=n+1}^{+\infty} a_k z^k\right| = 0.$$

Exemple La série entière $\sum_n z^n$ ne converge pas uniformément sur $D(0,1)$ car $\sum_{k=n+1}^{+\infty} z^k = \frac{z^{n+1}}{1-z}$ et cette quantité n'est pas bornée sur $D(0,1)$ (considérer la limite en $z=1$). Par contre, pour tout $r \in]0,1[$, la série entière $\sum_n z^n$ converge uniformément sur $D(0,r)$ car

$$\sup_{z \in D(0,r)} \left|\sum_{k=n+1}^{+\infty} z^k\right| = \frac{r^{n+1}}{1-r} \quad \text{et} \quad \lim_{n \to +\infty} \frac{r^{n+1}}{1-r} = 0.$$

De même, la notion de convergence normale pour une série entière complexe correspond à un cas particulier de convergence normale pour une série de fonctions de la variable complexe.

DÉFINITION 5.6 (Convergence normale)
Soit $(a_n)_n$ une suite de complexes et Z un sous-ensemble de $\mathbb{C}$. On dit que la série entière $\sum_n a_n z^n$ converge normalement sur Z si la série numérique réelle $\sum_n u_n$ de terme général $u_n = \sup_{z\in Z} |a_n z^n|$ converge.

Les séries entières étant des cas particuliers de séries de fonctions, si une série entière converge uniformément sur un sous-ensemble de $\mathbb{C}$ alors elle converge également simplement sur cet ensemble. Par ailleurs, si elle converge normalement sur un sous-ensemble de $\mathbb{C}$ alors elle converge aussi uniformément, absolument et simplement sur cet ensemble.

Par ailleurs, les résultats permettant d'établir la convergence uniforme ou la convergence normale des séries de fonctions énoncés dans les chapitres 3 et 4, s'appliquent au cas des séries entières.

Le résultat suivant est propre aux séries entières ; il indique que les ensembles pour lesquels il y a convergence normale ou convergence uniforme d'une série entière sont inclus dans le disque de convergence. La figure 2 illustre le théorème.

THÉORÈME 5.2 *Soit $\sum_n a_n z^n$ une série entière de rayon de convergence R. Pour tout réel $r \in]0, R[$, la série entière $\sum_n a_n z^n$ converge normalement et converge uniformément sur le disque fermé $\overline{D}(0,r)$.*

Démonstration Soit z_0 un complexe de module $r \in]0, R[$. Puisque $r < R$, z_0 appartient au disque de convergence de la série entière et d'après le théorème 5.1, la série numérique $\sum_n a_n z_0^n$ converge absolument. Or,

$$\sup_{z\in\overline{D}(0,r)} |a_n z^n| = |a_n| \sup_{z\in\overline{D}(0,r)} |z^n| = |a_n| r^n = |a_n| |z_0|^n = |a_n z_0^n|.$$

La convergence absolue de la série numérique $\sum_n a_n z_0^n$ implique donc la convergence normale sur $\overline{D}(0,r)$ de la série entière $\sum_n a_n z^n$ et par conséquent la convergence uniforme sur $\overline{D}(0,r)$. □

Une série entière converge simplement et absolument sur son disque de convergence, mais la convergence normale ou uniforme sur le disque de convergence n'est pas automatique comme le montrent les exemples suivants.

Exemples

1. La série entière $\sum_n z^n$ a pour rayon de convergence 1 mais ne converge pas normalement sur le disque $D(0,1)$ car $\sup_{z\in D(0,1)} |z^n| = 1$ (et la série $\sum_n 1$ diverge). Elle ne converge pas non plus uniformément sur $D(0,1)$ comme cela a été établi p. 246.

2. La série entière $\sum_{n\geqslant 1} \frac{1}{n^2} z^n$ a pour rayon de convergence 1 et elle converge normalement sur le disque fermé $\overline{D}(0,1)$ car $\sup_{z\in\overline{D}(0,1)} |\frac{1}{n^2} z^n| = \frac{1}{n^2}$ et la série de Riemann $\sum_{n\geqslant 1} \frac{1}{n^2}$ converge.

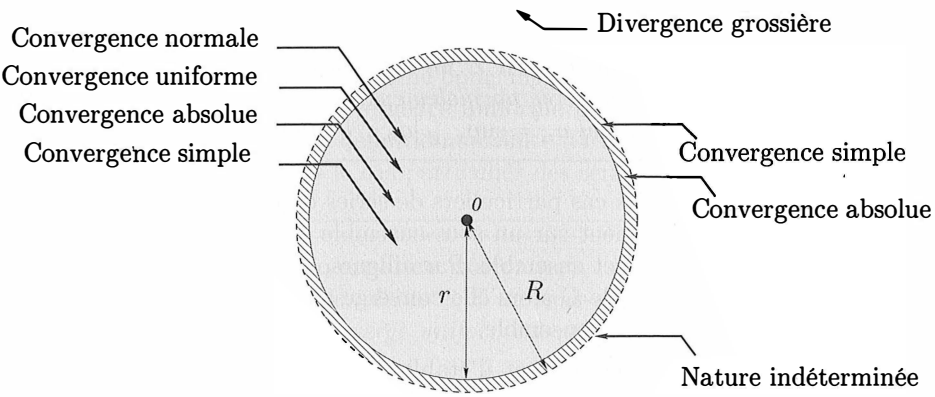


Fig. 2 Visualisation des différents type de convergence dans le disque de convergence d'une série entière.

5.2 Fonctions définies par une série entière

5.2.1 Continuité de la fonction somme

Le terme général d'une série entière étant polynomial, il définit une application continue sur $\mathbb{C}$. On déduit des résultats généraux concernant les séries de fonctions que la somme d'une série entière est continue sur tout ensemble où la série converge uniformément ; en particulier d'après le théorème 5.2 sur tout disque fermé $\overline{D}(0,r)$ où $r \in]0,R[$ inclus dans le disque de convergence. Le théorème suivant indique qu'une série entière est en fait continue sur tout son disque de convergence.

THÉORÈME 5.3 (Continuité de la somme d'une série entière)
Soit $\sum_n a_n z^n$ une série entière complexe dont le rayon de convergence R est non nul. Sa somme est une application continue sur le disque de convergence $D(0,R)$.

Démonstration Désignons par ϕ l'application, définie sur le disque de convergence $D(0,R)$, qui est la somme de la série entière $\sum_n a_n z^n$. Montrer que l'application ϕ est continue sur $D(0,R)$ revient à montrer qu'elle est continue en z_0 pour tout $z_0 \in D(0,R)$. Pour $z \in D(0,R)$ fixé, désignons par r le réel $\frac{1}{2}(R + |z_0|)$; il vérifie $|z_0| < r < R$. Pour montrer que ϕ est continue en z_0, il faut montrer[17] que pour tout réel ε strictement positif fixé, on peut trouver un réel η strictement positif tel que pour tout complexe $z \in D(0,r)$ avec $|z - z_0| \leqslant \eta$ on ait $|\phi(z) - \phi(z_0)| \leqslant \varepsilon$.

[17] Voir la définition 4.11 p. 189.

Pour tout $z \in D(0,r)$ et pour tout $n \in \mathbb{N}$ on a

$$\begin{aligned} |\phi(z) - \phi(z_0)| &= \Big|\phi(z) - \sum_{k=0}^{n} a_k z^k + \sum_{k=0}^{n} a_k (z^k - z_0^k) + \sum_{k=0}^{n} a_k z_0^k - \phi(z_0)\Big| \\ &\leqslant \Big|\phi(z) - \sum_{k=0}^{n} a_k z^k\Big| + \Big|\sum_{k=0}^{n} a_k (z^k - z_0^k)\Big| + \Big|\phi(z_0) - \sum_{k=0}^{n} a_k z_0^k\Big|. \end{aligned}$$

Comme la série entière $\sum_n a_n z^n$ converge uniformément sur $D(0,r)$ et a pour somme ϕ, d'après l'assertion (2) p. 246, il existe $N \in \mathbb{N}$ tel que pour tout $z \in D(0,r)$ on ait $|\phi(z) - \sum_{k=0}^{N} a_k z^k| \leqslant \frac{1}{3}\varepsilon$. De plus, puisque $z_0 \in D(0,r)$ on a en particulier $|\phi(z_0) - \sum_{k=0}^{N} a_k z_0^k| \leqslant \frac{1}{3}\varepsilon$. On en déduit que pour tout $z \in D(0,r)$,

$$|\phi(z) - \phi(z_0)| \leqslant \tfrac{2}{3}\varepsilon + \Big|\sum_{k=0}^{N} a_k (z^k - z_0^k)\Big|. \tag{5}$$

Comme la fonction $\phi_N : z \mapsto \sum_{k=0}^{N} a_k z^k$ est continue sur $D(0,r)$ (il s'agit d'une fonction polynomiale), il existe un réel η strictement positif tel que[17]

$$\forall z \in D(0,r) \quad \big(|z - z_0| \leqslant \eta \implies |\phi_N(x) - \phi_N(x_0)| \leqslant \tfrac{1}{3}\varepsilon\big).$$

On déduit alors de la relation (5) que pour tout $z \in D(0,r)$ tel que $|z - z_0| \leqslant \eta$ on a $|\phi(z) - \phi(z_0)| \leqslant \frac{1}{3}\varepsilon + \frac{1}{3}\varepsilon + \frac{1}{3}\varepsilon = \varepsilon$. On a ainsi montré que l'application ϕ est continue en z_0 pour tout $z_0 \in D(0,R)$. Elle est donc continue sur $D(0,R)$. □

5.2.2 Dérivabilité de la fonction somme

Les résultats concernant la dérivabilité de la fonction qui est la somme d'une série entière correspondent à des cas particuliers des résultats généraux concernant la dérivabilité de la somme d'une série de fonctions. Dans le chap. 3, nous avons établi ces résultats dans le cas des séries de fonctions réelles en ayant recours à l'intégrale de Riemann. Dans le cas des séries de fonctions de la variable complexe, la justification des résultats concernant la dérivabilité de la fonction somme ne peut utiliser la même approche. Dans le chap. 4 nous avons admis les résultats concernant la dérivabilité de la somme d'une série de fonctions de la variable complexe. Nous allons en donner ici une justification dans le cas particulier des séries entières.

THÉORÈME 5.4 *Soit $\sum_n a_n z^n$ une série entière complexe de rayon de convergence R non nul. Sa somme ϕ est une application holomorphe sur le disque de convergence $D(0,R)$ et pour tout $z \in D(0,R)$ on a*

$$\phi'(z) = \sum_{n=0}^{+\infty} (n+1)\, a_{n+1}\, z^n = \sum_{n=1}^{+\infty} n\, a_n\, z^{n-1}.$$

De plus, la série entière $\sum_n a_n z^n$ et sa dérivée $\sum_n (n+1)\, a_{n+1}\, z^n$ ont même rayon de convergence.

Démonstration Commençons par remarquer que l'égalité des deux sommes $\sum_{n=0}^{+\infty}(n+1)\,a_{n+1}\,z^n$ et $\sum_{n=1}^{+\infty} n\,a_n\,z^{n-1}$ résulte du changement d'indice n en $n-1$. Par ailleurs, d'après la proposition 5.2, les séries entières $\sum_n a_n\,z^n$ et $\sum_n (n+1)\,a_{n+1}\,z^n$ ont même rayon de convergence R. Cela implique que pour tout $r \in [0, R[$, la série numérique $\sum_n (n+1)\,a_{n+1}\,r^n$ converge absolument.

Pour montrer que l'application ϕ est holomorphe sur $D(0,R)$, il faut montrer qu'elle est holomorphe en z_0 pour tout $z_0 \in D(0,R)$. Soit z_0 un complexe fixé tel que $|z_0| < R$ et r le réel $\frac{1}{2}(R + |z_0|)$. On a $|z_0| < r < R$ et pour tout $z \in D(0,r)$ avec $z \neq z_0$ et pour tout $n \in \mathbb{N}^*$,

$$z^n - z_0^n = (z - z_0)(z^{n-1} + z_0 z^{n-2} + \ldots + z_0^{n-2} z + z_0^{n-1}) = (z - z_0)\,p_n(z)$$

où $p_n : z \in \mathbb{C} \mapsto \sum_{k=1}^{n} z^{n-k} z_0^{k-1}$ est une fonction polynomiale de degré $n-1$. On en déduit que :

$$\frac{\phi(z) - \phi(z_0)}{z - z_0} = \sum_{n=0}^{+\infty} a_n\,\frac{z^n - z_0^n}{z - z_0} = \sum_{n=0}^{+\infty} a_n\,p_n(z).$$

Pour montrer que ϕ est holomorphe en z_0, il faut montrer que cette quantité admet une limite quand z tend vers z_0. Pour tout $z \in D(0,r)$, on a

$$|a_n\,p_n(z)| \leqslant |a_n| \sum_{k=1}^{n} |z|^{n-k}\,|z_0|^{k-1} \leqslant |a_n| \sum_{k=1}^{n} r^{n-k}\,r^{k-1} = n\,|a_n|\,r^{n-1}.$$

On en déduit que

$$\sup_{z \in D(0,r)} |a_n\,p_n(z)| \leqslant n\,|a_n|\,r^{n-1}.$$

Puisque la série numérique $\sum_n (n+1)\,|a_{n+1}|\,r^n$ converge[18], on en déduit que la série $\sum_n a_n\,p_n(z)$ converge normalement sur $D(0,r)$. Sa somme $\psi : z \in D(0,r) \mapsto \sum_{n=0}^{+\infty} a_n\,p_n(z)$ est d'après le théorème 5.3 une fonction continue sur $D(0,r)$. En particulier, ψ est continue en z_0 ; on a donc

$$\lim_{z \to z_0} \frac{\phi(z) - \phi(z_0)}{z - z_0} = \lim_{z \to z_0} \psi(z) = \psi(z_0).$$

On en conclut que la dérivée de ϕ en z_0 existe et vaut $\phi'(z_0) = \psi(z_0)$. Or,

$$\psi(z_0) = \sum_{n=0}^{+\infty} a_n\,p_n(z_0) \quad \text{et} \quad p_n(z_0) = \sum_{k=1}^{n} z_0^{n-k} z_0^{k-1} = \sum_{k=1}^{n} z_0^{n-1} = n\,z_0^{n-1}.$$

On a donc $\phi'(z_0) = \sum_{n=1}^{+\infty} n\,a_n\,z_0^{n-1}$ et le théorème est démontré. □

[18] Les deux séries entières $\sum_n n\,|a_n|\,r^{n-1}$ et $\sum_n (n+1)\,|a_{n+1}|\,r^n$ sont identiques car les deux expressions correspondent aux changement d'indice n en $n+1$.

Exemples

1. La série entière complexe $\sum_n z^n$ admet pour rayon de convergence 1 et pour somme l'application $\phi : z \in D(0,1) \mapsto (1-z)^{-1}$. Elle est donc holomorphe sur $D(0,1)$ et pour tout $z \in D(0,1)$ on a

$$\phi'(z) = \frac{1}{(1-z)^2} = \sum_{n=0}^{+\infty} (n+1)\, z^n.$$

2. La série $\sum_n (n+1)\, z^n$ a pour rayon de convergence 1. Sa somme qui est l'application $\phi' : z \in D(0,1) \mapsto (1-z)^{-2}$ est donc holomorphe sur $D(0,1)$ et pour tout $z \in D(0,1)$,

$$\phi''(z) = \frac{2}{(1-z)^3} = \sum_{n=0}^{+\infty} (n+1)(n+2)\, z^n.$$

Remarque Il résulte du théorème 5.4 que les séries entières peuvent être dérivées « terme à terme » sur leur disque de convergence sans qu'il soit au préalable nécessaire d'étudier la série des dérivées comme c'est le cas pour les séries de fonctions, voir le théorème 3.9 p. 139. Attention toutefois au fait que cette propriété n'est pas valable sur le cercle de convergence. ◇

Propriétés de dérivabilité des fonctions usuelles de la variable complexe

Intéressons-nous aux propriétés de dérivabilité des fonctions usuelles de la variable complexe introduites dans la section 4.7 du chap. 4. La série entière $\sum_n \frac{1}{n!}\, z^n$ a pour rayon de convergence $R = +\infty$. Sa somme qui est la fonction exponentielle complexe est donc holomorphe sur $\mathbb{C}$ et pour tout $z \in \mathbb{C}$, on a

$$\exp'(z) = \sum_{n=0}^{+\infty} (n+1)\, \frac{1}{(n+1)!}\, z^n = \sum_{n=0}^{+\infty} \frac{1}{n!}\, z^n = \exp(z).$$

La dérivée de la fonction exponentielle est donc la fonction exponentielle elle même :

$$\boxed{\forall z \in \mathbb{C} \qquad \exp'(z) = \exp(z).}$$

Exercice 5.5 En appliquant le théorème des accroissements finis[19] à la fonction exponentielle complexe entre z et 0, montrer que pour tout $z \in D(0,1)$ on a $|\exp(z) - 1| \leqslant \mathrm{e}\,|z|$.

[19] Voir le théorème 4.2 p. 201.

Il résulte par ailleurs des relations

$$\cos(z) = \tfrac{1}{2}\left(\exp(\mathrm{i}z) + \exp(-\mathrm{i}z)\right), \qquad \sin(z) = \tfrac{1}{2\mathrm{i}}\left(\exp(\mathrm{i}z) - \exp(-\mathrm{i}z)\right)$$

et

$$\mathrm{ch}(z) = \tfrac{1}{2}\left(\exp(z) + \exp(-z)\right), \qquad \mathrm{sh}(z) = \tfrac{1}{2}\left(\exp(z) - \exp(-z)\right),$$

et de la proposition 4.15 p. 199 que les fonctions sinus et sinus hyperbolique complexes, cosinus et cosinus hyperbolique complexes sont holomorphe sur $\mathbb{C}$. Pour tout $z \in \mathbb{C}$, on a

$$\begin{array}{|rcl|rcl|}\hline \cos'(z) & = & -\sin(z) & \mathrm{ch}'(z) & = & \mathrm{sh}(z) \\ \sin'(z) & = & \cos(z) & \mathrm{sh}'(z) & = & \mathrm{ch}(z) \\ \hline \end{array}$$

La *détermination principale du logarithme complexe* est définie sur l'ensemble $\mathbb{C}^\ddagger = \mathbb{C} \setminus \{z \in \mathbb{C} \ ; \ \mathcal{R}\mathrm{e}(z) \leqslant 0 \ \text{ et } \ \mathcal{I}\mathrm{m}(z) = 0\}$ par

$$\mathrm{Log} : z \in \mathbb{C}^\ddagger \longmapsto \ln(|z|) + \mathrm{i}\mathrm{Arg}(z)$$

où Arg désigne la détermination principale de l'argument : $\mathrm{Arg}(z) \in]-\pi, \pi[$. Pour tout $z \in \mathbb{C}^\ddagger$ on a $\exp(\mathrm{Log}(z)) = z$ et par conséquent $\big(\exp(\mathrm{Log}(z))\big)' = 1$. Comme par ailleurs, $\big(\exp(\mathrm{Log}(z))\big)' = \mathrm{Log}'(z)\exp(\mathrm{Log}(z)) = z\,\mathrm{Log}'(z)$, on en déduit que

$$\boxed{\mathrm{Log}'(z) = \frac{1}{z}.}$$

Exercice 5.6 Montrer que la fonction de Bessel d'ordre 0 définie par

$$J_0(z) = \sum_{n=0}^{+\infty} \frac{(-1)^n}{(n!)^2}\left(\frac{z}{2}\right)^{2n}$$

est holomorphe sur $\mathbb{C}$ et que pour tout $z \in \mathbb{C}$ on a

$$J_0'(z) = -\sum_{n=0}^{+\infty} \frac{(-1)^n}{n! \times (n+1)!}\left(\frac{z}{2}\right)^{2n+1}.$$

Plus généralement, on a le résultat suivant concernant les dérivées successives de la somme d'une série entière. Il se démontre par récurrence sur l'entier p en utilisant le théorème 5.4.

Proposition 5.7 *Soit $\sum_n a_n z^n$ une série entière complexe de rayon de convergence R non nul. La somme ϕ de cette série entière est indéfiniment $\mathbb{C}$-dérivable sur le disque de convergence $D(0,R)$ et pour tout $p \in \mathbb{N}^*$ on a*

$$\phi^{(p)} : z \in D(0,R) \longmapsto \sum_{n=p}^{+\infty} \frac{n!}{(n-p)!}\, a_n\, z^{n-p}.$$

Exemples

1. La série entière complexe $\sum_n z^n$ admet pour rayon de convergence 1. La somme ϕ de cette série entière est donc indéfiniment $\mathbb{C}$-dérivable sur $D(0,1)$ et pour tous $p \in \mathbb{N}^*$ et $z \in D(0,1)$ on a

$$\phi^{(p)}(z) = \sum_{n=p}^{+\infty} \frac{n!}{(n-p)!} z^{n-p}.$$

Comme la somme de la série entière est l'application $\phi : z \in D(0,1) \mapsto (1-z)^{-1}$ et que $\phi^{(p)}(z) = p!/(1-z)^{p+1}$, on en déduit que

$$\boxed{\forall z \in D(0,1) \qquad \frac{1}{(1-z)^{p+1}} = \sum_{n=p}^{+\infty} C_n^p z^{n-p}.}$$

2. Considérons la fonction de Bessel d'ordre 0 qui est la somme sur $\mathbb{C}$ de la série entière $\sum_n \frac{(-1)^n}{(n!)^2} \left(\frac{z}{2}\right)^{2n}$. Nous avons établi à l'exercice 5.4 que cette série entière lacunaire a pour rayon de convergence $R = +\infty$ et pour tout $z \in \mathbb{C}$, on peut écrire $J_0(z) = \sum_{k=0}^{+\infty} a_k z^k$ où $a_k = 0$ si k est impair et $a_k = \frac{(-1)^n}{4^n (n!)^2}$ si k est pair avec $k = 2n, n \in \mathbb{N}$. Il résulte de la proposition 5.7 que J_0 est indéfiniment $\mathbb{C}$-dérivable sur $\mathbb{C}$. La dérivée seconde de J_0 est donnée par

$$\begin{aligned}
J_0''(z) &= \sum_{k=2}^{+\infty} k(k-1) a_k \, z^{k-2} \\
&= \overbrace{\sum_{n=1}^{+\infty} (2n+1)(2n) \underbrace{a_{2n+1}}_{=0} z^{2n-1}}^{\text{indices } k \text{ impairs } k=2n+1} + \overbrace{\sum_{n=0}^{+\infty} (2n+2)(2n+1)\, a_{2n+2}\, z^{2n}}^{\text{indices } k \text{ pairs } k=2n+2} \\
&= \sum_{n=0}^{+\infty} (2n+2)(2n+1) \, \frac{(-1)^{n+1}}{4^{n+1}\,((n+1)!)^2} \, z^{2n} \\
&= \sum_{n=0}^{+\infty} \frac{(-1)^{n+1}}{n! \times (n+1)!} \, \frac{2n+1}{2} \, \left(\frac{z}{2}\right)^{2n}.
\end{aligned}$$

5.2.3 Fonction développable en série entière

Jusqu'à présent notre propos a consisté, étant donné une série entière complexe de rayon de convergence R non nul, à étudier sur le disque de convergence les propriétés de la fonction qui est la somme de la série entière. Dans cette partie, nous nous intéressons au problème inverse : étant donné une fonction de la variable complexe définie dans un voisinage de l'origine, existe-t-il une série entière dont la somme soit cette fonction ?

DÉFINITION 5.7 *Soit f une fonction de $\mathbb{C}$ dans $\mathbb{C}$ définie dans un voisinage de l'origine. On dit que f est* développable en série entière dans un voisinage de l'origine *s'il existe une série entière complexe $\sum_n a_n z^n$ de rayon de convergence non nul et un voisinage V de l'origine tels que :*

$$\forall z \in V \qquad f(z) = \sum_{n=0}^{+\infty} a_n z^n.$$

Exemples

1. Une fonction polynomiale $f_1 : z \mapsto \sum_{n=0}^{d} c_n z^n$ de degré d est développable en série entière dans un voisinage de l'origine car pour tout $z \in \mathbb{C}$ on a $f_1(z) = \sum_{n=0}^{+\infty} a_n z^n$ où les coefficients $(a_n)_n$ sont définis par $a_n = c_n$ pour tout $n \in \{0, \ldots, d\}$ et $a_n = 0$ sinon.

2. La fonction $f_2 : z \mapsto (1-z)^{-1}$ est développable en série entière dans un voisinage de l'origine car pour tout $z \in D(0,1)$ on a $f_2(z) = \sum_{n=0}^{+\infty} z^n$.

3. La fonction $f_3 : z \mapsto (1-z^4)^{-1}$ est développable en série entière dans un voisinage de l'origine car par la substitution z en z^4 dans l'expression obtenue dans l'exemple précédent, on obtient $f_3(z) = \sum_{k=0}^{+\infty} z^{4k}$ pour tout $z \in D(0,1)$.

4. La fonction exponentielle complexe est développable en série entière dans un voisinage de l'origine car pour tout $z \in \mathbb{C}$ on a, voir la définition 4.15 p. 204, $\exp(z) = \sum_{n=0}^{+\infty} \frac{1}{n!} z^n$.

5. Les fonctions sinus, cosinus, sinus hyperbolique et cosinus hyperbolique sont développables en série entière dans un voisinage de l'origine puisque d'après la proposition 4.21 p. 210, on a pour tout $z \in \mathbb{C}$

$$\cos(z) = \sum_{n=0}^{+\infty} \frac{(-1)^n}{(2n)!} z^{2n}, \qquad \sin(z) = \sum_{n=0}^{+\infty} \frac{(-1)^n}{(2n+1)!} z^{2n+1},$$

$$\mathrm{ch}(z) = \sum_{n=0}^{+\infty} \frac{1}{(2n)!} z^{2n}, \qquad \mathrm{sh}(z) = \sum_{n=0}^{+\infty} \frac{1}{(2n+1)!} z^{2n+1}.$$

Remarques

1. Pour $z_0 \in \mathbb{C}^*$, on dit qu'une fonction f de $\mathbb{C}$ dans $\mathbb{C}$ est *développable en série entière dans un voisinage de z_0* s'il existe une série entière complexe $\sum_n a_n z^n$ de rayon de convergence non nul et un voisinage V de z_0 tels que

$$\forall z \in V \qquad f(z) = \sum_{n=0}^{+\infty} a_n (z - z_0)^n.$$

2. Une application f d'un ouvert U de $\mathbb{C}$ dans $\mathbb{C}$ est dite *analytique* sur U si pour tout $z_0 \in U$ l'application f est développable en série entière dans un

voisinage de z_0. On démontre[20] que la somme d'une série entière de rayon de convergence R non nul est analytique dans le disque ouvert $D(0,R)$. ◇

THÉORÈME 5.5 (Série de Mac-Laurin)
Soit f une fonction de $\mathbb{C}$ dans $\mathbb{C}$ développable en série entière dans un voisinage de l'origine. Il existe un voisinage V de l'origine tel que f soit indéfiniment $\mathbb{C}$-dérivable sur V et

$$\forall z \in V \qquad f(z) = \sum_{n=0}^{+\infty} \frac{1}{n!} f^{(n)}(0)\, z^n.$$

La série entière $\sum_n \frac{1}{n!} f^{(n)}(0)\, z^n$ est appelée série de Mac-Laurin *de f.*

Démonstration Soit f une fonction complexe de la variable complexe développable en série entière dans un voisinage de l'origine. D'après la définition 5.7, il existe un voisinage V de 0 et une série entière $\sum_n a_n z^n$ de rayon de convergence R non nul tels que pour tout $z \in V$ on ait : $f(z) = \sum_{n=0}^{+\infty} a_n z^n$.

D'après la proposition 5.7, l'application f, qui sur V est la somme de la série entière $\sum_n a_n z^n$, est indéfiniment dérivable et pour tous $p \in \mathbb{N}^*$ et $z \in V$ on a

$$f^{(p)}(z) = \sum_{k=p}^{+\infty} \frac{k!}{(k-p)!} a_k\, z^{k-p} = p!\, a_p + \sum_{k=p+1}^{+\infty} \frac{k!}{(k-p)!} a_k\, z^{k-p}.$$

En évaluant cette expression en $z = 0$, on obtient $f^{(p)}(0) = p!\, a_p$ pour tout $p \in \mathbb{N}^*$. On a aussi $f(0) = a_0 = 0!\, a_0$. On en déduit que pour tout $n \in \mathbb{N}$ on a $a_n = \frac{1}{n!} f^{(n)}(0)$ et par conséquent que pour tout $z \in V$, on a $f(z) = \sum_{n=0}^{+\infty} \frac{1}{n!} f^{(n)}(0)\, z^n$. □

COROLLAIRE 5.2 (Unicité du développement en série entière)
Soit f une fonction de $\mathbb{C}$ dans $\mathbb{C}$ définie dans un voisinage de l'origine. Si f est développable en série entière dans un voisinage de l'origine alors ce développement est unique et est donné par la série de Mac-Laurin.

Démonstration Il a été établi dans la preuve du théorème 5.5 que si f est une fonction complexe admettant un développement en série entière dans un voisinage de l'origine de la forme $f(z) = \sum_{n=0}^{+\infty} a_n z^n$ alors $a_n = \frac{1}{n!} f^{(n)}(0)$. Le développement est donc entièrement déterminé par f et est unique : il s'agit du développement de Mac-Laurin. □

[20] Voir par exemple : H. Cartan, *Théorie élémentaire des fonctions analytiques*, Hermann, 1961.

Remarque Le théorème 5.5 a une conséquence intéressante. Si deux fonctions f et g, développables en série entière dans un voisinage V de l'origine, sont telles que

$$\exists r \in \mathbb{R}_+^* \quad \forall z \in D(0,r) \quad f(z) = g(z)$$

alors ces deux fonctions sont égales sur tout l'ensemble V. ◇

Si f est une fonction de $\mathbb{C}$ dans $\mathbb{C}$ admettant un développement en série entière dans un voisinage V de l'origine de la forme $f(z) = \sum_{k=0}^{+\infty} a_k z^k$ alors, en restriction aux valeurs x réelles de la variable, on a pour tout $n \in \mathbb{N}$

$$f(x) = \sum_{k=0}^{n} a_k x^k + x^{n+1} \sum_{k=0}^{+\infty} a_{n+1+k} x^k = \sum_{k=0}^{n} a_k x^k + \mathcal{O}_0(x^n).$$

Cette expression constitue le développement limité[21] à l'ordre n de f à l'origine.

Le résultat suivant, qui est un corollaire du théorème 5.5, étend au cas des fonctions de la variable complexe la notion de développement limité.

Proposition 5.8 (Développement limité d'une fonction complexe) *Soit f une fonction de $\mathbb{C}$ dans $\mathbb{C}$ admettant un développement en série entière dans un voisinage V de l'origine de la forme $f(z) = \sum_{k=0}^{+\infty} a_k z^k$.*

Pour tout $n \in \mathbb{N}$, f admet un développement limité d'ordre n en 0 *donné par :*

$$\forall z \in V \qquad f(z) = \sum_{k=0}^{n} a_k z^k + \mathcal{O}_0(z^n).$$

Démonstration Désignons par R le rayon de convergence de la série entière $\sum_n a_n z^n$. L'ordre n du développement limité étant fixé, il s'agit de montrer que pour tout $z \in V$,

$$f(z) - \sum_{k=0}^{n} a_k z^k = z^n \varepsilon(z) \qquad \text{avec } \lim_{z \to 0} \varepsilon(z) = 0.$$

Pour $z \in V$, on a

$$f(z) - \sum_{k=0}^{n} a_k z^k = \sum_{k=n+1}^{+\infty} a_k z^k = z^{n+1} \sum_{k=0}^{+\infty} a_{n+k+1} z^k = z^n \varepsilon(z)$$

où on a posé $\varepsilon(z) = z \sum_{k=0}^{+\infty} a_{n+k+1} z^k$.

Intéressons-nous à la série entière $\sum_k a_{n+k+1} z^k$; elle a même rayon de convergence que la série entière $\sum_k a_k z^k$, voir p. 236. Elle converge donc dans $D(0,R)$ et sa somme ψ est une fonction continue sur $D(0,R)$. On en déduit que pour tout $z \in D(0,R)$ on a $\varepsilon(z) = z\,\phi(z)$ ce qui implique que $\lim_{z \to 0} \varepsilon(z) = 0$ car ϕ est continue en 0. Le résultat est démontré. □

[21] Voir le chap. 17 du *Cours de première année*.

Le théorème 5.5 indique que toute fonction complexe développable en série entière dans un voisinage de l'origine est indéfiniment $\mathbb{C}$-dérivable sur un voisinage de l'origine ; elle est donc en particulier holomorphe dans ce voisinage. Ce résultat admet une réciproque : toute fonction holomorphe dans un voisinage de l'origine est développable en série entière dans un voisinage de l'origine. Plus précisément, on a le résultat suivant que nous admettrons[(22)].

THÉORÈME 5.6 *Soit f une fonction holomorphe dans un disque ouvert $D(0,r)$ (où $r \in \mathbb{R}_+^*$). Il existe une série entière $\sum_n a_n z^n$ de rayon de convergence R avec $R \geqslant r$ telle que :*

$$\forall z \in D(0,r) \qquad f(z) = \sum_{n=0}^{+\infty} a_n z^n.$$

Remarques

1. Les théorèmes 5.4, 5.5 et 5.6 nous indiquent qu'il y a coïncidence entre l'ensemble des fonctions holomorphes sur un ouvert U de $\mathbb{C}$ et l'ensemble des fonctions analytiques sur U (*i.e.* développables en séries entières au voisinage de z_0 pour tout $z_0 \in U$).

2. Il résulte du théorème 5.6 que si une fonction f est holomorphe dans le disque ouvert $D(0,r)$ alors elle est indéfiniment $\mathbb{C}$-dérivable dans $D(0,r)$ puisque la somme d'une série entière est indéfiniment $\mathbb{C}$-dérivable sur son disque de convergence, voir la proposition 5.7. En particulier, la dérivée d'une fonction holomorphe est elle-même une fonction holomorphe. Cette propriété est très forte et est spécifique aux fonctions de la variable complexe ; elle n'est bien entendu pas vraie dans le cas des fonctions réelles de la variable réelle. ◇

Développement en série entière des fractions rationnelles

Rappelons qu'une fonction polynomiale $p : z \in \mathbb{C} \mapsto \sum_{k=0}^{d} a_k z^k$ ($d \in \mathbb{N}$) définit une série entière complexe dont les coefficients d'ordre supérieur à d sont nuls et dont les coefficients d'ordre inférieur où égal à d sont égaux aux coefficients de même ordre du polynôme.

Considérons deux polynômes P et Q de $\mathbb{C}[X]$ premiers entre eux (*i.e.* sans racine commune) et intéressons-nous au développement en série entière de la fonction rationnelle $F : z \mapsto P(z)/Q(z)$ dans un voisinage de l'origine. On peut déjà remarquer que si 0 est un pôle de F (*i.e.* si 0 est une racine de Q) alors F ne peut avoir de développement en série entière dans un voisinage de l'origine en vertu du théorème 5.5.

(22) La démonstration de ce résultat dépasse le cadre de ce cours. On pourra consulter le chap. 2 de : H. Cartan, *Théorie élémentaire des fonctions analytiques*, Hermann, 1961.

Désignons par $\alpha_1, \ldots, \alpha_q$ les racines distinctes de Q, supposées non nulles, et par $\eta_1, \ldots, \eta_q$ leur multiplicité respective. La fraction rationnelle F admet une décomposition en éléments simples dans $\mathbb{C}(X)$ de la forme[23] :

$$F(z) = E(z) + \sum_{i=1}^{q} \sum_{j=1}^{\eta_i} \frac{\beta_{ij}}{(z-\alpha_i)^j},$$

où E désigne la partie entière[24] de F et les β_{ij} désignent des complexes. La fonction polynomiale E constitue son propre développement en série entière dans un voisinage de l'origine. Pour déterminer le développement en série entière dans un voisinage de l'origine de la fonction rationnelle F, il suffit de savoir calculer le développement en série entière dans un voisinage de l'origine des fonctions de la forme $\phi : z \mapsto (z-a)^{-p}$ où a désigne un complexe non nul et p un entier naturel non nul.

Il a été établi p. 253 que pour tout $z \in D(0,1)$ on a

$$\frac{1}{(1-z)^{p+1}} = \sum_{n=p}^{+\infty} C_n^p \, z^{n-p}.$$

On en déduit que pour tout $z \in D(0,|a|)$,

$$\begin{aligned} \frac{1}{(z-a)^{p+1}} &= \frac{(-1)^{p+1}}{a^{p+1}} \frac{1}{(1-\frac{z}{a})^{p+1}} = \frac{(-1)^{p+1}}{a^{p+1}} \sum_{n=p}^{+\infty} C_n^p \left(\frac{z}{a}\right)^{n-p} \\ &= (-1)^{p+1} \sum_{n=p}^{+\infty} C_n^p \frac{1}{a^{n+1}} z^{n-p}. \end{aligned}$$

On en conclut que toute fonction rationnelle qui n'admet pas 0 pour pôle est développable en série entière dans un voisinage de l'origine et que le rayon de convergence R de la série entière est égal au minimum des distances entre 0 et les pôles de F, *i.e.* $R = \min_{1 \leqslant i \leqslant q} |\alpha_i|$.

On obtient l'expression du développement en série entière dans un voisinage de l'origine d'une fonction rationnelle en considérant successivement les différents termes de la décomposition en éléments simples de la fraction rationnelle.

Exemples

1. Le développement en série entière dans un voisinage de l'origine de la fonction rationnelle $z \mapsto \frac{1}{1-z^2}$ s'obtient de la manière suivante. Pour tout $z \in D(0,1)$ on a

$$\frac{1}{1-z^2} = \frac{\frac{1}{2}}{1-z} + \frac{\frac{1}{2}}{1+z} = \frac{1}{2} \sum_{n=0}^{+\infty} z^n + \frac{1}{2} \sum_{n=0}^{+\infty} (-1)^n z^n = \sum_{n=0}^{+\infty} z^{2n}.$$

[23] Voir la section 7.2.3 p. 279 du *Cours de première année*.

[24] Voir le *Cours de première année* p. 274.

2. Pour obtenir le développement en série entière dans un voisinage de l'origine de la fonction rationnelle $F : z \mapsto \frac{z^2+\mathrm{i}}{z^4-1}$, on part de la décomposition en éléments simples de la fraction rationnelle qui est donnée par[(24)] :

$$\frac{X^2+\mathrm{i}}{X^4-1} = \frac{\frac{1}{4}(1+\mathrm{i})}{X-1} + \frac{-\frac{1}{4}(1+\mathrm{i})}{X-\mathrm{i}} + \frac{-\frac{1}{4}(1+\mathrm{i})}{X+1} + \frac{\frac{1}{4}(1+\mathrm{i})}{X+\mathrm{i}}.$$

Pour tout $z \in D(0,1)$, on a

$$\frac{1}{z-1} = -\sum_{n=0}^{+\infty} z^n \quad \text{et} \quad \frac{1}{z-\mathrm{i}} = \frac{\mathrm{i}}{1+\mathrm{i}z} = -\sum_{n=0}^{+\infty} (-\mathrm{i})^{n+1} z^n$$

$$\frac{1}{z+1} = \sum_{n=0}^{+\infty} (-1)^n z^n \quad \text{et} \quad \frac{1}{z+\mathrm{i}} = -\frac{\mathrm{i}}{1-\mathrm{i}z} = -\sum_{n=0}^{+\infty} \mathrm{i}^{n+1} z^n.$$

On en déduit que la fonction rationnelle F admet pour développement en série entière dans un voisinage de l'origine $\sum_{n=0}^{+\infty} a_n z^n$ où pour tout $n \in \mathbb{N}$ on a

$$a_n = \tfrac{1}{4}(1+\mathrm{i})(-1+(-\mathrm{i})^{n+1} - (-1)^n - \mathrm{i}^{n+1}) = \begin{cases} 0 & \text{si } n \text{ est impair} \\ -\mathrm{i} & \text{si } n = 4k, k \in \mathbb{N} \\ -1 & \text{si } n = 4k+2, k \in \mathbb{N} \end{cases}.$$

On en conclut que pour tout $z \in D(0,1)$ on a $F(z) = -\sum_{k=0}^{+\infty} \alpha_k z^{2k}$ avec $\alpha_k = \mathrm{i}$ si k est impair et $\alpha_k = 1$ si k est pair.

On peut obtenir sur cet exemple l'expression du développement en série entière plus rapidement en remarquant que

$$\frac{z^2+\mathrm{i}}{z^4-1} = -z^2 \frac{1}{1-z^4} - \mathrm{i}\frac{1}{1-z^4} \quad \text{et} \quad \frac{1}{1-z^4} = \sum_{k=0}^{+\infty} z^{4k}.$$

3. Déterminons le développement en série entière dans un voisinage de l'origine de la fonction rationnelle $F : z \mapsto \frac{1}{z^2+z+1}$. En notant $\mathrm{j} = \mathrm{e}^{\mathrm{i}\frac{2\pi}{3}}$, la décomposition en éléments simples de la fraction rationnelle F est

$$\frac{1}{X^2+X+1} = -\frac{\mathrm{i}}{\sqrt{3}} \frac{1}{X-\mathrm{j}} + \frac{\mathrm{i}}{\sqrt{3}} \frac{1}{X-\bar{\mathrm{j}}}.$$

On en déduit que pour tout $z \in D(0,1)$, on a

$$\frac{1}{z^2+z+1} = \frac{\mathrm{i}}{\mathrm{j}\sqrt{3}} \frac{1}{1-\frac{X}{\mathrm{j}}} - \frac{\mathrm{i}}{\bar{\mathrm{j}}\sqrt{3}} \frac{1}{1-\frac{X}{\bar{\mathrm{j}}}}$$

$$= \frac{\mathrm{i}}{\mathrm{j}\sqrt{3}} \sum_{n=0}^{+\infty} \frac{1}{\mathrm{j}^n} z^n - \frac{\mathrm{i}}{\bar{\mathrm{j}}\sqrt{3}} \sum_{n=0}^{+\infty} \frac{1}{\bar{\mathrm{j}}^n} z^n = \frac{\mathrm{i}}{\sqrt{3}} \sum_{n=0}^{+\infty} \left(\frac{1}{\mathrm{j}^{n+1}} - \frac{1}{\bar{\mathrm{j}}^{n+1}} \right) z^n.$$

En remarquant que d'une part $\mathrm{j}^{n+1}\bar{\mathrm{j}}^{n+1} = (\mathrm{j}\bar{\mathrm{j}})^{n+1} = |\mathrm{j}|^{2(n+1)} = 1$ et que d'autre part $\bar{\mathrm{j}}^{n+1} - \mathrm{j}^{n+1} = \mathrm{e}^{-\mathrm{i}\frac{2}{3}\pi(n+1)} - \mathrm{e}^{\mathrm{i}\frac{2}{3}\pi(n+1)} = -2\mathrm{i}\sin(\frac{2}{3}\pi(n+1))$, on obtient

$$\frac{1}{z^2+z+1} = \frac{2}{\sqrt{3}} \sum_{n=0}^{+\infty} \sin\left(\frac{2}{3}\pi(n+1)\right) z^n.$$

La série entière obtenue est une série entière lacunaire puisque si $n+1$ est un multiple de 3 on a $\sin\left(\frac{2}{3}\pi(n+1)\right) = 0$. On peut avoir recours à MAPLE pour expliciter ses coefficients et comparer au résultat obtenu avec la commande `series` de MAPLE.

```
> a:=n->(2/sqrt(3))*sin((2*Pi*(n+1)/3)):
> seq(a(i),i=0..12);
```

$$1, -1, 0, 1, -1, 0, 1, -1, 0, 1, -1, 0, 1$$

```
> convert(series(1/(1+x+x^2), x=0,13),polynom);
```

$$1 - x + x^3 - x^4 + x^6 - x^7 + x^9 - x^{10} + x^{12}$$

Exercice 5.7 Déterminer le développement en série entière dans un voisinage de l'origine de la fonction rationnelle $z \mapsto \dfrac{z}{(z-1)^2(z-2)}$.

5.3 Série entière de la variable réelle

5.3.1 Propriétés des séries entières réelles

On appelle *série entière* de la variable réelle (ou série entière réelle) toute série d'applications de $\mathbb{R}$ dans $\mathbb{R}$ de terme général $f_n : x \in \mathbb{R} \mapsto a_n x^n$ où $a_n \in \mathbb{R}$. Pour tout $n \in \mathbb{N}$, le réel a_n est appelé *coefficient d'ordre* n de la série entière.

Une série entière de la variable réelle étant définie d'une manière analogue à une série entière de la variable complexe, on peut introduire un cadre d'étude similaire. En particulier le rayon de convergence d'une série entière réelle est défini de la même manière que celui d'une série entière complexe, voir la définition 5.3. D'un point de vue pratique, le rayon de convergence d'une série entière réelle peut lui aussi être calculé en ayant recours aux règles de Cauchy ou de D'Alembert, voir les propositions 5.4 et 5.5. Par ailleurs, le rayon de convergence d'une série entière réelle $\sum_n a_n x^n$ est le même que celui de la série entière complexe $\sum_n a_n z^n$ à coefficients a_n réels.

Les différentes propriétés de convergence d'une série entière réelle de rayon de convergence R non nul sont les suivantes :

- la série entière converge simplement et absolument sur $]-R, R[$;
- la série entière diverge grossièrement sur $]-\infty, -R[$ et sur $]R, +\infty[$;
- en R et en $-R$, la série entière peut selon les cas converger ou diverger ; une étude spécifique est requise ;
- quel que soit le réel r avec $0 < r < R$, la série entière converge normalement et uniformément sur l'intervalle $[-r, r]$.

Les résultats qui suivent donnent les propriétés de la somme d'une série entière sur l'intervalle de convergence de la série. Ils découlent des résultats généraux concernant la somme des séries de fonctions réelles de la variable réelle, voir le chap. 3, et du fait que le terme général de la série, qui est un monôme, est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$.

THÉORÈME 5.7 (Continuité d'une série entière réelle)
Soit $\sum_n a_n x^n$ une série entière réelle de rayon de convergence R non nul. Sa somme ϕ est une application continue sur l'intervalle $]-R, R[$.

Démonstration Il n'est pas possible d'appliquer le théorème 3.7 p. 135 sur l'intervalle $]-R, R[$ car on ne peut pas être assuré de la convergence uniforme de la série entière sur cet intervalle. Toutefois, pour tout $r \in]0, R[$, on a convergence uniforme de la série entière sur $]-r, r[$ et on peut donc appliquer le théorème 3.7 sur cet intervalle.

Nous allons montrer que quel que soit le réel x_0 de l'intervalle $]-R, R[$, la somme de la série est continue en x_0. Pour $x_0 \in]-R, R[$ fixé, considérons le réel $r = |x_0| + \frac{1}{2}(R - |x_0|)$; on a $x_0 \in [-r, r]$ et $r < R$. Puisque la série entière converge uniformément sur $]-r, r[$, d'après le théorème 3.7, sa somme ϕ est continue sur $]-r, r[$ et par conséquent en x_0. Le théorème est démontré. □

Remarque En complément au théorème 5.7, on a le résultat suivant : si la série numérique $\sum_n a_n R^n$ converge alors l'application ϕ est continue à gauche en R et si la série numérique $\sum_n a_n (-R)^n$ converge alors l'application ϕ est continue à droite en $-R$. ◇

THÉORÈME 5.8 (Dérivabilité d'une série entière réelle)
Soit $\sum_n a_n x^n$ une série entière réelle de rayon de convergence R non nul et de somme ϕ. L'application ϕ est une application dérivable sur $]-R, R[$ et pour tout $x \in]-R, R[$ on a

$$\phi'(x) = \sum_{n=0}^{+\infty} (n+1)\, a_{n+1}\, x^n = \sum_{n=1}^{+\infty} n\, a_n\, x^{n-1}.$$

Démonstration On procède comme dans la preuve du théorème 5.7 : quel que soit le réel x_0 de l'intervalle $]-R, R[$, il existe un réel $r \in]0, R[$ tel que $x_0 \in]-r, r[$. Puisque la série entière converge uniformément sur $]-r, r[$, on peut appliquer le théorème 3.9 sur l'intervalle $]-r, r[$. On obtient la dérivabilité en x_0 de la somme ϕ de la série entière et la relation $\phi'(x_0) = \sum_{n=1}^{+\infty} n\, a_n\, x_0^{n-1}$. Par ailleurs, en effectuant le changement d'indice n en $n-1$, on obtient l'égalité entre les deux sommes indiquée dans le théorème.. □

Plus généralement, on a le résultat suivant qui découle du théorème 5.8 par récurrence sur l'ordre de dérivation.

Corollaire 5.3 *Soit $\sum_n a_n x^n$ une série entière réelle de rayon de convergence R non nul et de somme ϕ. L'application ϕ est indéfiniment dérivable sur $]-R, R[$ et pour tout $p \in \mathbb{N}^*$ on a*

$$\phi^{(p)} : x \in]-R, R[\longmapsto \sum_{n=p}^{+\infty} \frac{n!}{(n-p)!} a_n \, x^{n-p}.$$

On a également le résultat suivant qui concerne l'intégrabilité de la somme d'une série entière réelle.

Proposition 5.9 *Soit $\sum_n a_n x^n$ une série entière réelle de rayon de convergence R non nul et de somme ϕ. Quels que soient les réels α et β dans $]-R, R[$ avec $\alpha < \beta$, l'application ϕ est Riemann-intégrable sur $[\alpha, \beta]$ et*

$$\int_\alpha^\beta \phi(x) \, \mathrm{d}x = \sum_{n=0}^{+\infty} \frac{a_n}{n+1} \, (\beta^{n+1} - \alpha^{n+1}).$$

Démonstration Par hypothèse, la série de fonctions de terme général $f_n : x \mapsto a_n x^n$ converge uniformément sur tout intervalle compact inclus dans $]-R, R[$. Elle converge donc sur $[\alpha, \beta]$ pour tous $\alpha, \beta \in]-R, R[$ avec $\alpha < \beta$. Par ailleurs, les fonctions polynomiales f_n étant Riemann-intégrables sur tout intervalle fermé et borné, il résulte du théorème 3.8 p. 136 que :

$$\int_\alpha^\beta \phi(x) \, \mathrm{d}x = \sum_{n=0}^{+\infty} \int_\alpha^\beta a_n x^n \, \mathrm{d}x = \sum_{n=0}^{+\infty} \frac{a_n}{n+1} (\beta^{n+1} - \alpha^{n+1}).$$

Le résultat est démontré. □

On a le corollaire suivant qui permet de calculer la somme de certaines séries entières.

Corollaire 5.4 *Soit $\sum_n a_n x^n$ une série entière réelle de rayon de convergence R non nul et de somme ϕ. La série $\sum_{n \geqslant 1} \frac{a_{n-1}}{n} x^n$ a pour rayon de convergence R et*

$$\forall x \in]-R, R[\qquad \sum_{n=1}^{+\infty} \frac{a_{n-1}}{n} x^n = \sum_{n=0}^{+\infty} \frac{a_n}{n+1} x^{n+1} = \int_0^x \phi(t) \, \mathrm{d}t.$$

Démonstration On a établi à la proposition 5.2 que les séries $\sum_n \frac{1}{n+1} a_n \, x^n$ et $\sum_n a_n \, x^n$ ont le même rayon de convergence. Pour $x \in]0, R[$ fixé, appliquons la proposition 5.9 avec $\alpha = 0$ et $\beta = x$, on a

$$\int_0^x \phi(t) \, \mathrm{d}t = \sum_{n=0}^{+\infty} \frac{a_n}{n+1} \, x^{n+1} = x \sum_{n=0}^{+\infty} \frac{a_n}{n+1} \, x^n.$$

Pour $x \in]-R, 0[$, le résultat est obtenu de la même manière en prenant $\alpha = x$ et $\beta = 0$. Si $x = 0$, l'égalité a encore lieu car on a $\int_0^x \phi(t)\,\mathrm{d}t = 0$ et $\sum_{n=0}^{+\infty} \frac{a_n}{n+1} x^{n+1} = 0$. □

La série géométrique $\sum_n x^n$ est une série entière de rayon $R = 1$ et de somme $\phi : x \mapsto 1/(1-x)$. On en déduit du corollaire 5.4 que la série entière $\sum_{n\geqslant 1} \frac{1}{n} x^n$ a pour rayon de convergence 1 et que sa somme ψ est donnée sur $]-1, 1[$ par :

$$\psi(x) = \int_0^x \frac{1}{1-t}\,\mathrm{d}t = -\ln(1-x).$$

Du point de vue pratique, on peut présenter les calculs sous la forme suivante :

$$\begin{aligned}\sum_{n=1}^{+\infty} \frac{1}{n} x^n &= \sum_{n=0}^{+\infty} \frac{1}{n+1} x^{n+1} = \sum_{n=0}^{+\infty} \int_0^x t^n\,\mathrm{d}t \\ &= \int_0^x \left(\sum_{n=0}^{+\infty} t^n \right) \mathrm{d}t = \int_0^x \frac{1}{1-t}\,\mathrm{d}t = -\ln(1-x).\end{aligned}$$

On a donc la relation suivante :

$$\boxed{\forall x \in]-1, 1[\qquad \sum_{n=1}^{+\infty} \frac{1}{n} x^n = -\ln(1-x)}$$

et par substitution de x en $-x$, on obtient la relation :

$$\boxed{\forall x \in]-1, 1[\qquad \sum_{n=1}^{+\infty} \frac{(-1)^n}{n} x^n = -\ln(1+x)}$$

De même, la série entière $\sum_n x^{2n}$ a pour rayon de converge 1 et pour somme $\phi : x \in]-1, 1[\mapsto \frac{1}{1-x^2}$. Il s'agit d'une série entière lacunaire dont les coefficients $(a_n)_n$ sont données par $a_n = 0$ si n est impair et $a_n = 1$ si n est pair. Le corollaire 5.4 indique que la série entière $\sum_{n\geqslant 1} \frac{a_{n-1}}{n} x^n$ a pour rayon de convergence 1 et que sa somme ψ est donnée sur $]-1, 1[$ par :

$$\psi(x) = \sum_{n=0}^{+\infty} \frac{a_n}{n+1} x^{n+1} = \sum_{k=0}^{+\infty} \frac{1}{2k+1} x^{2k+1} = \int_0^x \frac{1}{1-t^2}\,\mathrm{d}t = \operatorname{argth}(x).$$

On en déduit la relation :

$$\boxed{\forall x \in]-1, 1[\qquad \sum_{n=0}^{+\infty} \frac{1}{2n+1} x^{2n+1} = \operatorname{argth}(x)}$$

Sur le même principe, la série entière $\sum_n (-1)^n x^{2n}$ a pour rayon de convergence 1 et pour somme $\phi : x \in]-1, 1[\mapsto \frac{1}{1+x^2}$. Le corollaire 5.4 permet d'obtenir la relation :

$$\boxed{\forall x \in]-1, 1[\qquad \sum_{n=0}^{+\infty} \frac{(-1)^n}{2n+1} x^{2n+1} = \arctan(x)}$$

5.3.2 Développement en série entière d'une fonction de la variable réelle

À l'instar de ce qui a été fait avec les fonctions de la variable complexe, on peut se demander si une fonction de la variable réelle définie dans un voisinage de l'origine est la somme d'une série entière de la variable réelle. Le théorème 5.6 indique que toute fonction de la variable complexe holomorphe dans un voisinage de l'origine est développable en série entière dans un voisinage de l'origine. Nous allons voir que dans le cas des fonctions de la variable réelle, une fonction dérivable (et même de classe $\mathcal{C}^\infty$) dans un voisinage de l'origine n'est pas nécessairement développable en série entière.

DÉFINITION 5.8 *Une fonction réelle de la variable réelle f définie dans un voisinage de l'origine est dite* développable en série entière *dans un voisinage de l'origine s'il existe une série entière réelle de rayon de convergence R non nul et un réel $r \in]0, R[$ tel que*

$$\forall x \in]-r, r[\qquad f(x) = \sum_{n=0}^{+\infty} a_n x^n.$$

Exemples

1. Une fonction polynomiale réelle $f_1 : x \mapsto \sum_{n=0}^{d} c_n x^n$ de degré d est développable en série entière dans un voisinage de l'origine car pour tout $x \in \mathbb{R}$ on a $f_1(x) = \sum_{n=0}^{+\infty} a_n x^n$ où les coefficients $(a_n)_n$ sont définis par $a_n = c_n$ pour tout $n \in \{0, \ldots, d\}$ et $a_n = 0$ sinon.

2. La fonction $f_2 : x \mapsto (1-x)^{-1}$ est développable en série entière dans un voisinage de l'origine car pour tout $x \in]-1, 1[$ on a $f_2(x) = \sum_{n=0}^{+\infty} x^n$, voir p. 3.

3. La fonction $f_3 : x \mapsto (1-x^2)^{-1}$ est développable en série entière dans un voisinage de l'origine car, par substitution de x en x^2 dans l'expression de $f_2(x)$, on obtient $f_3(x) = \sum_{n=0}^{+\infty} x^{2n}$ pour tout $x \in]-1, 1[$.

4. La fonction exponentielle est développable en série entière dans un voisinage de l'origine car pour tout réel x on a $\mathrm{e}^x = \sum_{n=0}^{+\infty} \frac{1}{n!} x^n$, voir p. 3.

Comme dans le cas des fonctions complexes de la variable complexe, le résultat suivant indique que si une fonction réelle de la variable réelle est développable en série entière de la variable réelle dans un voisinage de l'origine alors, dans un voisinage de l'origine, f est égale à sa série de Mac-Laurin. La démonstration de ce résultat, qui reprend les idées développées dans celle du théorème 5.5, est omise.

PROPOSITION 5.10 (Série de Mac-Laurin)
Soit f une fonction réelle de la variable réelle développable en série entière dans un voisinage de l'origine. Il existe un réel r strictement positif tel que f est de classe $\mathcal{C}^\infty$ sur $]-r,r[$ et

$$\forall x \in]-r,r[\qquad f(x) = \sum_{n=0}^{+\infty} \frac{1}{n!} f^{(n)}(0)\, x^n.$$

La série entière réelle $\sum_n \frac{1}{n!} f^{(n)}(0)\, x^n$ est appelée série de Mac-Laurin *de f.*

On déduit de la proposition 5.10 le corollaire suivant qui indique que si une fonction est développable en série entière dans un voisinage de l'origine alors ce développement est celui donné par la série de Mac-Laurin.

COROLLAIRE 5.5 *Soit f une fonction réelle de la variable réelle définie dans un voisinage de l'origine. Si f est développable en série entière dans un voisinage de l'origine alors ce développement est unique et est donné par la série de Mac-Laurin.*

Remarque La proposition 5.10 indique que toute fonction réelle de la variable réelle développable en série entière sur un intervalle $]-r,r[$, où $r \in \mathbb{R}_+^*$, est de classe $\mathcal{C}^\infty$ sur cet intervalle. Mais la réciproque est fausse : il existe des fonctions de classe $\mathcal{C}^\infty$ sur un voisinage de l'origine qui ne sont pas développables en série entière dans un voisinage de l'origine. C'est le cas par exemple de l'application f définie sur $\mathbb{R}$ par $f(x) = \exp(-1/x^2)$ si $x \neq 0$ et $f(0) = 0$. Cette application est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$ et pour tout $n \in \mathbb{N}$ on a $f^{(n)}(0) = 0$. La série de Mac-Laurin de f est donc la série nulle mais il n'existe aucun voisinage de 0 sur lequel f coïncide avec la fonction nulle. ◇

Étant donné une application f de classe $\mathcal{C}^\infty$ sur un intervalle $]-r,r[$ avec $r \in \mathbb{R}_+^*$, on peut également se poser la question de savoir à quelle condition il existe une série entière dont elle soit la somme. D'après le corollaire 5.5, si une telle série existe, elle est nécessairement unique et est égale à la série de Mac-Laurin. Le problème est de savoir si la série de Mac-Laurin a un rayon de convergence non nul et, le cas échéant, si sa somme est bien égale à f sur un voisinage de l'origine.

Considérons une application f de classe $\mathcal{C}^\infty$ sur un intervalle $]-r,r[$ et appliquons la formule de Taylor-Mac-Laurin[25] à l'ordre n entre 0 et x pour $x \in]-r,r[$; il existe $\theta \in]0,1[$ tel que :

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!}\, x^k + \frac{f^{(n+1)}(\theta x)}{(n+1)!}\, x^{n+1}.$$

[25] Voir le *Cours de première année* p. 753.

Cette relation indique que $f(x) = S_n(x) + R_n(x)$ où $S_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!} x^k$ est la somme partielle d'ordre n de la série de Mac-Laurin $\sum_n \frac{f^{(n)}(0)}{n!} x^n$ et $R_n(x) = \frac{f^{(n+1)}(\theta x)}{(n+1)!} x^{n+1}$ correspond à un reste d'ordre n.

Il est clair que si pour tout $x \in]-r, r[$, $R_n(x)$ tend vers 0 quand n tend vers $+\infty$ alors $S_n(x)$ tend vers $f(x)$ et cela signifie que la série de Mac-Laurin associée à f est convergente sur $]-r, r[$ et de somme f. Inversement, si la série de Mac-Laurin associée à f est convergente sur $]-r, r[$ alors pour tout $x \in]-r, r[$, la quantité $R_n(x) = f(x) - S_n(x)$ tend vers 0 quand n tend vers $+\infty$. On en déduit qu'une condition nécessaire et suffisante pour qu'une fonction f de classe $\mathcal{C}^\infty$ sur un intervalle $]-r, r[$ soit somme de sa série de Mac-Laurin sur $]-r, r[$ est que pour tout $x \in]-r, r[$ la quantité $R_n(x)$ tende vers 0 quand n tend vers $+\infty$.

D'un point de vue pratique, vérifier que pour tout $x \in]-r, r[$ la quantité $R_n(x)$ tend vers 0 quand n tend vers $+\infty$ peut se révéler ardu, en particulier en raison du fait que θ est généralement difficile à déterminer. Le résultat suivant qui se contente de donner une condition suffisante pour qu'une fonction réelle de la variable réelle soit développable en série entière dans un voisinage de l'origine est d'une utilisation plus aisée.

THÉORÈME 5.9 *Soient r un réel strictement positif et f une application de $]-r, r[$ dans $\mathbb{R}$ de classe $\mathcal{C}^\infty$. Si f a toutes ses dérivées uniformément bornées sur $]-r, r[$, i.e. si*

$$\exists M \in \mathbb{R}^+ \quad \forall n \in \mathbb{N} \quad \forall x \in]-r, r[\quad \left|f^{(n)}(x)\right| \leqslant M$$

alors f est développable en série entière dans un voisinage de l'origine et cette série entière est égale à la série de Mac-Laurin.

Démonstration Sous l'hypothèse

$$\exists M \in \mathbb{R}^+ \quad \forall n \in \mathbb{N} \quad \forall x \in]-r, r[\quad \left|f^{(n)}(x)\right| \leqslant M$$

on a, pour tout $x \in]-r, r[$,

$$|R_n(x)| = \frac{|f^{(n+1)}(\theta x)|}{(n+1)!} |x|^{n+1} \leqslant \frac{M}{(n+1)!} |r|^{n+1}.$$

En utilisant le théorème d'encadrement, on en déduit que $\lim_{n \to +\infty} R_n(x) = 0$ pour tout $x \in]-r, r[$. Or, on a vu qu'une condition nécessaire et suffisante pour qu'une fonction f de classe $\mathcal{C}^\infty$ sur un intervalle $]-r, r[$ soit somme de sa série de Mac-Laurin sur $]-r, r[$, et par conséquent développable en série entière dans un voisinage de l'origine, est que pour tout $x \in]-r, r[$ la quantité $R_n(x)$ tende vers 0 quand n tend vers $+\infty$. □

Exemples

1. La fonction sinus est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$ et pour tout $n \in \mathbb{N}$, la dérivée n-ième de sinus en $x \in \mathbb{R}$ est donnée par[26] $\sin^{(n)}(x) = \sin(x + \frac{n\pi}{2})$. On a donc

$$\forall n \in \mathbb{N} \quad \forall x \in \mathbb{R} \quad \left|\sin^{(n)}(x)\right| \leqslant 1.$$

D'après le théorème 5.9, la fonction sinus est développable en série entière dans un voisinage de l'origine et

$$\boxed{\forall x \in \mathbb{R} \qquad \sin(x) = \sum_{n=0}^{+\infty} \frac{\sin^{(n)}(0)}{n!}\, x^n = \sum_{k=0}^{+\infty} \frac{(-1)^k}{(2k+1)!}\, x^{2k+1}}$$

2. La fonction cosinus est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$ et pour tout $n \in \mathbb{N}$ la dérivée n-ième de sinus en $x \in \mathbb{R}$ est donnée par[26] $\cos^{(n)}(x) = \cos(x + \frac{n\pi}{2})$. On a donc

$$\forall n \in \mathbb{N} \quad \forall x \in \mathbb{R} \quad \left|\cos^{(n)}(x)\right| \leqslant 1.$$

D'après le théorème 5.9, la fonction cosinus est développable en série entière dans un voisinage de l'origine et

$$\boxed{\forall x \in \mathbb{R} \qquad \cos(x) = \sum_{n=0}^{+\infty} \frac{\cos^{(n)}(0)}{n!}\, x^n = \sum_{k=0}^{+\infty} \frac{(-1)^k}{(2k)!}\, x^{2k}}$$

5.3.3 Quelques utilisations des séries entières de la variable réelle

Séries entières et calcul d'intégrale

Intéressons-nous au calcul de l'intégrale $\int_0^1 \frac{t\ln(t)}{t-1}\,\mathrm{d}t$ dont la valeur ne peut être calculée à l'aide des méthodes vues dans le *Cours de première année*[27].

Remarquons tout d'abord que cette intégrale généralisée converge puisque $\lim_{t\to 0^+} \frac{t\ln(t)}{t-1} = 0$ et que $\lim_{t\to 1^-} \frac{t\ln(t)}{t-1} = 1$. Par ailleurs, pour tout $t \in]-1,1[$ on a $\frac{1}{1-t} = \sum_{n=0}^{+\infty} t^n$ et le rayon de convergence de la série entière $\sum_n x^n$ est 1. On déduit de la proposition 5.9 que pour tout $x \in]0,1[$, on a

$$\int_0^x \frac{t\ln(t)}{t-1}\,\mathrm{d}t = \int_0^x \sum_{n=0}^{+\infty} \ln(t)\; t^{n+1}\,\mathrm{d}t = \sum_{n=0}^{+\infty} \int_0^x \ln(t)\; t^{n+1}\,\mathrm{d}t.$$

En intégrant par parties, on obtient :

$$\int_0^x \ln(t)\; t^{n+1}\,\mathrm{d}t = \frac{1}{n+2}\, x^{n+2}\left(\ln(x) - \frac{1}{n+2}\right).$$

[26] Voir le *Cours de première année* p. 726.

[27] Voir le chap. 18 du *Cours de première année*.

On a donc,

$$\begin{aligned}\int_0^x \frac{t\ln(t)}{t-1}\,\mathrm{d}t &= \ln(x)\sum_{n=0}^{+\infty}\frac{1}{n+2}x^{n+2}-\sum_{n=0}^{+\infty}\frac{1}{(n+2)^2}x^{n+2}\\ &= \ln(x)\sum_{n=2}^{+\infty}\frac{1}{n}x^n-\sum_{n=2}^{+\infty}\frac{1}{n^2}x^n\\ &= -\ln(x)(x+\ln(1-x))+1-\sum_{n=1}^{+\infty}\frac{1}{n^2}x^n.\end{aligned}$$

Notons que la série de fonctions $\sum_{n\geqslant 1}\frac{1}{n^2}x^n$ converge normalement sur $]0,1[$ car $\sup_{x\in]0,1[}\left|\frac{1}{n^2}x^n\right|=\frac{1}{n^2}$ et la série de Riemann $\sum_{n\geqslant 1}\frac{1}{n^2}$ converge. Cette série de fonctions converge donc uniformément sur $]0,1[$ et il résulte du corollaire 3.2 p. 135 que

$$\lim_{x\to 1^-}\sum_{n=1}^{+\infty}\frac{1}{n^2}x^n=\sum_{n=1}^{+\infty}\frac{1}{n^2}=\frac{\pi^2}{6}.$$

Comme par ailleurs $\lim_{x\to 1^-}\ln(x)(x+\ln(1-x))=0$, on en déduit que

$$\int_0^1\frac{t\ln(t)}{t-1}\,\mathrm{d}t=\lim_{x\to 1^-}\int_0^x\frac{t\ln(t)}{t-1}\,\mathrm{d}t=1-\frac{\pi^2}{6}.$$

Séries entières et dénombrement

Le BDE de l'école est confronté à la situation suivante : il dispose d'un budget de 100 euros pour organiser une fête et doit pour cela prévoir d'acheter x bouteilles d'une boisson X au coût unitaire de 2 euros et y gâteaux de la marque Y au coût unitaire de 3 euros. L'organisateur de la soirée charge les étudiants ASINSA de déterminer le nombre de possibilités de répartition existant entre les x bouteilles et les y gâteaux pour utiliser exactement le budget alloué (par exemple $50X+0Y$ et $2X+32Y$ sont deux possibilités d'utiliser exactement le budget de 100 euros).

Il s'agit donc de déterminer tous les couples d'entiers naturels (x,y) tels que $2x+3y=100$. Nous allons voir qu'une utilisation astucieuse des séries entières va permettre de répondre à la question.

Considérons les deux séries entières lacunaires $\sum_n x^{2n}$ et $\sum_n x^{3n}$ de rayon de convergence 1. Le produit de ces deux séries entières est une série entière $\sum_n c_n x^n$ de rayon de convergence $R\geqslant 1$ dont les coefficients sont donnés par la relation $c_n=\sum_{k=0}^n a_k\,b_{n-k}$ où $(a_n)_n$ et $(b_n)_n$ désignent respectivement les coefficients des séries entières lacunaires $\sum_n x^{2n}$ et $\sum_n x^{3n}$. On a d'une part $a_n=1$ si n est un entier pair et $a_n=0$ si n est un entier impair et d'autre part $b_n=1$ si n est multiple de 3 et 0 sinon. On en déduit que $a_k\,b_{n-k}=1$ si k est multiple de 2 et $n-k$ est multiple de 3, autrement dit, $a_k\,b_{n-k}=1$ s'il existe $(x,y)\in\mathbb{N}^2$ tel que $k=2x$ et $n-k=3y$. Dans tous les autres cas, on a $a_k\,b_{n-k}=0$. Éliminant k des deux équations $k=2x$ et $n-k=3y$, on obtient

la relation $2x+3y=n$. On en déduit que :

$$c_n=\sum_{k=0}^{n}a_k\,b_{n-k}=\sum_{\substack{(x,y)\in\mathbb{N}^2\\2x+3y=n}}a_{2x}\,b_{3y}=\sum_{\substack{(x,y)\in\mathbb{N}^2\\2x+3y=n}}1=\mathrm{Card}\big\{(x,y)\in\mathbb{N}^2\ ;\ 2x+3y=n\big\}.$$

Le nombre de possibilités de répartition entre x bouteilles à 2 euros et y gâteaux à 3 euros pour utiliser exactement le budget de 100 euros est donc égal à c_{100}. Il nous reste à déterminer la valeur de ce coefficient.

La série entière $\sum_n c_n x^n$ étant le produit des deux séries entières $\sum_n x^{2n}$ et $\sum_n x^{3n}$ de sommes respectives $\phi : x\in]-1,1[\mapsto \frac{1}{1-x^2}$ et $\psi : x\in]-1,1[\mapsto \frac{1}{1-x^3}$, on a pour tout $x\in]-1,1[$

$$\sum_{n=0}^{+\infty}c_n x^n=\frac{1}{1-x^2}\times\frac{1}{1-x^3}=\frac{1}{(1-x)^2(1+x)(x^2+x+1)}.$$

La fraction rationnelle apparaissant dans le membre de droite admet pour décomposition en éléments simples dans $\mathbb{R}(X)$:

$$\frac{1}{(1-X^2)(1-X^3)}=\frac{1}{6(X-1)^2}-\frac{1}{4(X-1)}+\frac{1}{4(X+1)}+\frac{1}{3(X^2+X+1)}$$

Or, pour tout $x\in]-1,1[$ on a

$$\frac{1}{(x-1)^2}=\sum_{n=0}^{+\infty}(n+1)\,x^n,\qquad \frac{1}{x-1}=-\sum_{n=0}^{+\infty}x^n,\qquad \frac{1}{x+1}=\sum_{n=0}^{+\infty}(-1)^n\,x^n,$$

et, on a établi p. 259 que

$$\frac{1}{x^2+x+1}=\sum_{n=0}^{+\infty}\frac{2}{\sqrt{3}}\sin\left(\tfrac{2}{3}(n+1)\pi\right)\,x^n.$$

On en déduit que, pour tout $x\in]-1,1[$,

$$\frac{1}{(1-x^2)(1-x^3)}=\sum_{n=0}^{+\infty}\left(\frac{1}{6}(n+1)+\frac{1}{4}+\frac{(-1)^n}{4}+\frac{2}{3\sqrt{3}}\sin(\tfrac{2}{3}(n+1)\pi)\right)\,x^n$$

et par conséquent que

$$c_n=\frac{1}{6}(n+1)+\frac{1}{4}+\frac{(-1)^n}{4}+\frac{2}{3\sqrt{3}}\sin(\tfrac{2}{3}(n+1)\pi).$$

On obtient finalement : $c_{100}=\dfrac{101}{6}+\dfrac{1}{2}+\dfrac{2}{3\sqrt{3}}\sin(\tfrac{2}{3}(101)\pi)=17.$

Signalons que ce calcul pouvait être également effectué en ayant recours à MAPLE.

```
> P:=convert(series(1/((1-x^2)*(1-x^3)),x,101),polynom):
> coeff(P,x,100);
```

5.3.4 Développement en série entière des fonctions usuelles

Les développements en série entière des fonctions usuelles de la variable réelle sont indiqués dans le tableau suivant :

Développement en série entière	Validité
$\dfrac{1}{1-x} = \sum_{n=0}^{+\infty} x^n$	$x \in]-1,1[$
$\dfrac{1}{(1-x)^2} = \sum_{n=0}^{+\infty} (n+1)\,x^n$	$x \in]-1,1[$
$\forall \alpha \in \mathbb{R} \quad (1+x)^\alpha = 1 + \sum_{n=1}^{+\infty} \dfrac{\alpha(\alpha-1)\ldots(\alpha-n+1)}{n!}\,x^n$	$x \in]-1,1[$
$\forall \alpha \in \mathbb{R} \quad (1-x)^\alpha = 1 + \sum_{n=1}^{+\infty} (-1)^n \dfrac{\alpha(\alpha-1)\ldots(\alpha-n+1)}{n!}\,x^n$	$x \in]-1,1[$
$\ln(1+x) = \sum_{n=1}^{+\infty} \dfrac{(-1)^{n+1}}{n}\,x^n$	$x \in]-1,1[$
$\ln(1-x) = -\sum_{n=1}^{+\infty} \dfrac{1}{n}\,x^n$	$x \in]-1,1[$
$\mathrm{e}^x = \sum_{n=0}^{+\infty} \dfrac{1}{n!}\,x^n$	$x \in \mathbb{R}$
$\mathrm{ch}(x) = \sum_{n=0}^{+\infty} \dfrac{1}{(2n)!}\,x^{2n}$	$x \in \mathbb{R}$
$\mathrm{sh}(x) = \sum_{n=0}^{+\infty} \dfrac{1}{(2n+1)!}\,x^{2n+1}$	$x \in \mathbb{R}$
$\cos(x) = \sum_{n=0}^{+\infty} \dfrac{(-1)^n}{(2n)!}\,x^{2n}$	$x \in \mathbb{R}$
$\sin(x) = \sum_{n=0}^{+\infty} \dfrac{(-1)^n}{(2n+1)!}\,x^{2n+1}$	$x \in \mathbb{R}$

$\arctan(x) = \sum_{n=0}^{+\infty} \frac{(-1)^n}{2n+1} x^{2n+1}$		$x \in]-1,1[$
$\operatorname{argth}(x) = \sum_{n=0}^{+\infty} \frac{1}{2n+1} x^{2n+1}$		$x \in]-1,1[$

Exercice 5.8 En utilisant le développement en série entière dans un voisinage de l'origine des fonctions $x \mapsto (1+x)^\alpha$ et $x \mapsto (1-x)^\alpha$, déterminer ceux des fonctions arc-sinus, arc-cosinus, argument sinus hyperbolique et argument cosinus hyperbolique.

5.4 Séries entières et équations différentielles

5.4.1 Résolution d'une équation différentielle à l'aide de séries entières

Nous avons vu dans le *Cours de première année*[28] des méthodes de résolution d'une équation différentielle linéaire du premier ordre et d'une équation différentielle linéaire du second ordre à coefficients constants. Il n'existe pas de méthode générale pour résoudre une équation différentielle linéaire à coefficients non constants d'ordre supérieur ou égal à 2. Dans certains cas, il est toutefois possible de rechercher des solutions d'une telle équation sous forme de séries entières et c'est cette problématique que nous allons considérer dans cette partie.

Commençons par quelques compléments concernant les équations différentielles linéaires. Une équation différentielle linéaire d'ordre p (où $p \in \mathbb{N}^*$) est une équation de la forme

$$a_0(x)y^{(p)}(x) + a_1(x)\,y^{(p-1)}(x) + \ldots + a_{p-1}(x)\,y'(x) + a_p(x)\,y(x) = b(x) \quad (6)$$

où b et $a_i, i \in \{0,\ldots,p\}$ désignent des applications définies sur un même intervalle I de $\mathbb{R}$ tel que a_0 ne soit pas identiquement nul et y est une fonction réelle de la variables réelle inconnue.

Résoudre l'équation (6) sur un intervalle I de $\mathbb{R}$ consiste à déterminer les applications y de I dans $\mathbb{R}$ de classe $\mathcal{C}^p$ sur I et qui vérifient pour tout $x \in I$ la relation

$$a_0(x)y^{(p)}(x) + a_1(x)\,y^{(p-1)}(x) + \ldots + a_{p-1}(x)\,y'(x) + a_p(x)\,y(x) = b(x).$$

[28] Voir le chap. 20 du *Cours de première année*.

Si a_0 ne s'annule pas sur I alors, après division par a_0, résoudre l'équation (6) sur I revient à résoudre l'équation différentielle dite « normalisée » suivante :

$$y^{(p)}(x) + \alpha_1(x)\, y^{(p-1)}(x) + \ldots + \alpha_{p-1}(x)\, y'(x) + \alpha_p(x)\, y(x) = \beta(x) \qquad (7)$$

où $\beta = b/a_0$ et $\alpha_i = a_i/a_0$ pour tout $i \in \{1, \ldots, p\}$.

Les conditions d'existence et d'unicité d'une solution à l'équation (7) sont données par le théorème suivant que nous admettrons.

THÉORÈME 5.10 (Théorème de Cauchy)
Soient I un intervalle de $\mathbb{R}$ et pour $p \in \mathbb{N}^*$ fixé b, $\alpha_1, \ldots, \alpha_p$ des applications de I dans $\mathbb{R}$ continues. Quels que soient $x_0 \in I$ et $(\gamma_0, \ldots, \gamma_{p-1}) \in \mathbb{R}^p$, l'équation différentielle

$$y^{(p)}(x) + \alpha_1(x)\, y^{(p-1)}(x) + \ldots + \alpha_{p-1}(x)\, y'(x) + \alpha_p(x)\, y(x) = \beta(x)$$

admet une unique solution y de I dans $\mathbb{R}$ vérifiant les conditions

$$y(x_0) = \gamma_0, \quad y'(x_0) = \gamma_1, \quad \ldots, \quad y^{(p-1)}(x_0) = \gamma_{p-1}.$$

La plupart des résultats concernant les équations différentielles d'ordre 1 et 2 s'étendent aux cas des équations différentielles linéaires d'ordre p quelconque. En particulier, on appelle équation homogène associée à l'équation différentielle (7) l'équation

$$y^{(p)}(x) + \alpha_1(x)\, y^{(p-1)}(x) + \ldots + \alpha_{p-1}(x)\, y'(x) + \alpha_p(x)\, y(x) = 0. \qquad (8)$$

Les solutions de l'équation homogène (8) forment un sous-espace vectoriel de dimension p de l'espace vectoriel des applications continues de I dans $\mathbb{R}$. De plus, on obtient les solutions de l'équation différentielle (7) en sommant à l'expression générale des solutions de l'équation homogène (8) une solution particulière de l'équation (7).

Dans le cas d'une équation différentielle linéaire d'ordre 1 homogène, nous avons vu dans le cours de première année qu'il existe une méthodologie générale pour la résoudre. Il n'en existe pas pour résoudre une équation différentielle linéaire d'ordre $p \geqslant 2$ si les fonctions $\alpha_1, \ldots, \alpha_p$ ne sont pas toutes constantes. Nous allons voir qu'il est toutefois possible dans certains cas de trouver une base de l'espace des solutions de l'équation homogène (8) formée par des fonctions qui sont sommes de séries entières. Cette approche est basée sur le résultat suivant que nous admettrons.

PROPOSITION 5.11 *Si les fonctions réelles de la variable réelle $\alpha_1, \ldots, \alpha_n$ et β sont développables en série entière dans un voisinage de l'origine, alors toutes les solutions de l'équation différentielle linéaire d'ordre p*

$$y^{(p)}(x) + \alpha_1(x)\, y^{(p-1)}(x) + \cdots + \alpha_{p-1}(x)\, y'(x) + \alpha_p(x)\, y(x) = \beta(x)$$

sont développables en série entière dans un voisinage de l'origine.

Illustrons par un exemple l'utilisation de la proposition 5.11 ; considérons l'équation différentielle

$$y''(x) + xy(x) = 0$$

connue sous le nom d'équation d'Airy[29], qui apparaît dans la modélisation de nombreux phénomènes physiques comme les phénomènes de diffraction ou encore l'aérodynamique. La fonction $\alpha_1 : x \mapsto x$ est développable en série entière sur $\mathbb{R}$ et la proposition 5.11 nous indique que les solutions de l'équation d'Airy sont toutes développables en série entière sur $\mathbb{R}$. Une solution de l'équation d'Airy s'écrira sur $\mathbb{R}$ sous la forme $y(x) = \sum_{n=0}^{+\infty} a_n x^n$ où les coefficients $(a_n)_{n\in\mathbb{N}}$ vérifieront des relations de récurrence que l'on va obtenir formellement en écrivant que cette série entière est solution de l'équation d'Airy. On a

$$\begin{aligned} y'(x) &= \sum_{n=1}^{+\infty} n\, a_n\, x^{n-1} = \sum_{n=0}^{+\infty} (n+1)\, a_{n+1}\, x^n \\ y''(x) &= \sum_{n=2}^{+\infty} n(n-1)\, a_n\, x^{n-2} = \sum_{n=0}^{+\infty} (n+2)(n+1)\, a_{n+2}\, x^n \end{aligned}$$

et

$$\begin{aligned} y''(x) + xy(x) &= \sum_{n=0}^{+\infty} (n+2)(n+1) a_{n+2}\, x^n + \sum_{n=0}^{+\infty} a_n\, x^{n+1} \\ &= 2a_2 + \sum_{n=1}^{+\infty} (n+2)(n+1) a_{n+2}\, x^n + \sum_{n=1}^{+\infty} a_{n-1}\, x^n \\ &= 2a_2 + \sum_{n=1}^{+\infty} \left((n+2)(n+1) a_{n+2} + a_{n-1}\right) x^n. \end{aligned}$$

On en déduit que pour que y soit solution de l'équation d'Airy, il faut que pour tout $x \in \mathbb{R}$

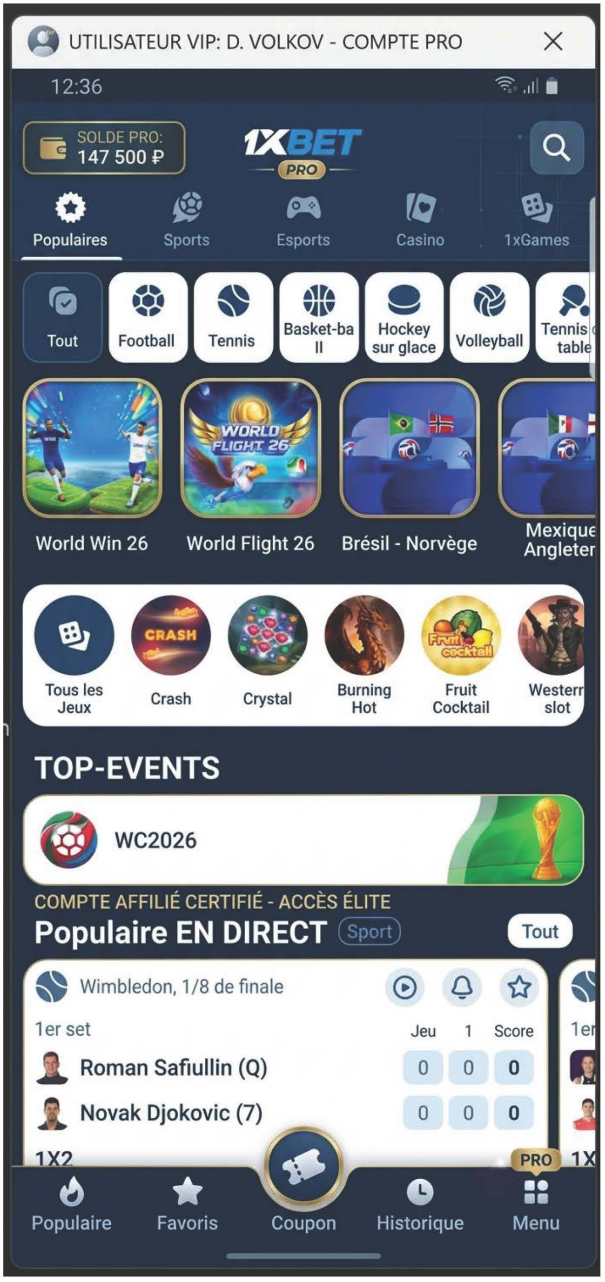
$$2a_2 + \sum_{n=1}^{+\infty} \left((n+2)(n+1)\, a_{n+2} + a_{n-1}\right) x^n = 0.$$

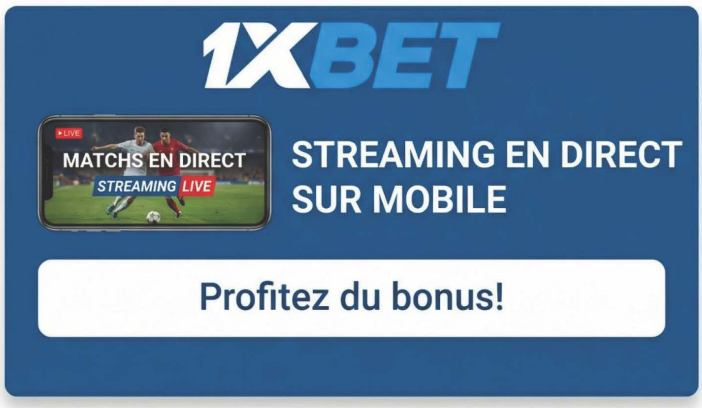
Cette relation indique que la série entière $\sum_n c_n x^n$ où $c_0 = 2a_2$ et $c_n = (n+2)(n+1)a_{n+2} + a_{n-1}$ pour tout $n \in \mathbb{N}^*$ est égale à la série nulle. Or, deux séries entières sont égales si et seulement si elles ont mêmes coefficients : pour tout $n \in \mathbb{N}$ on a donc $c_n = 0$. Les coefficients $(a_n)_{n\in\mathbb{N}}$ vérifient par conséquent les relations : $2a_2 = 0$ et $(n+2)(n+1)a_{n+2} + a_{n-1} = 0$ pour tout $n \in \mathbb{N}^*$. On en déduit que

$$a_2 = 0 \quad \text{et} \quad a_{n+2} = -\frac{1}{(n+2)(n+1)}\, a_{n-1} \quad \forall n \in \mathbb{N}^*.$$

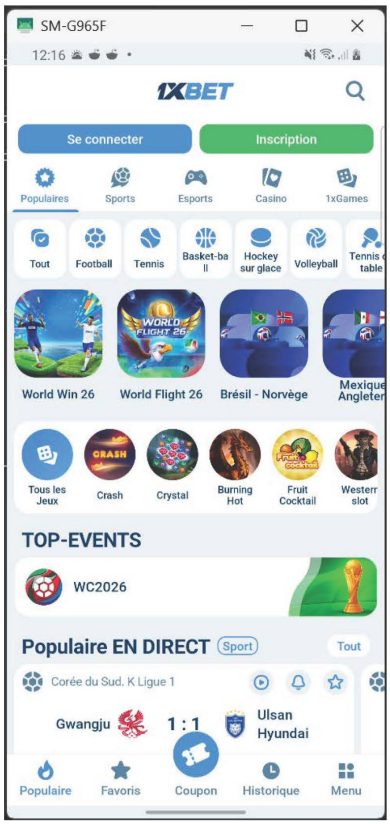
[29] Airy, George Biddell (1801, Alnwick, Angleterre - 1892, Greenwich). Mathématicien et astronome, on lui doit en particulier de nombreuses mesures pendulaires pour déterminer la masse de la Terre et une théorie des arcs-en-ciel. Outre cette équation, Airy a laissé son nom à des fonctions dites « spéciales » liées aux fonctions de Bessel.

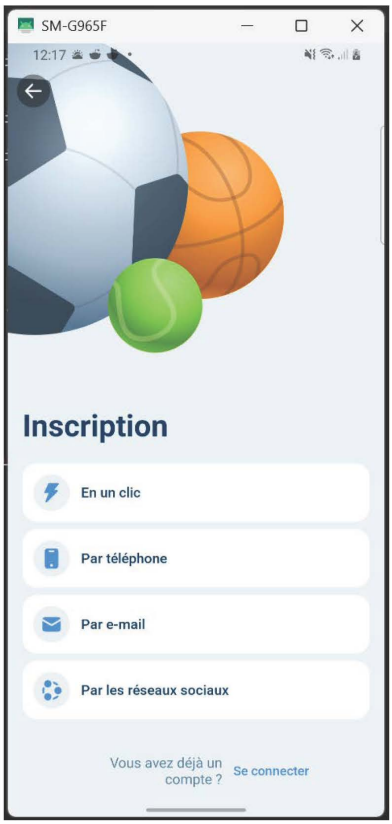

ÉTAPES D'INSCRIPTION
You Tube
pour voir comment creer un compte
CLICK TO WATCH

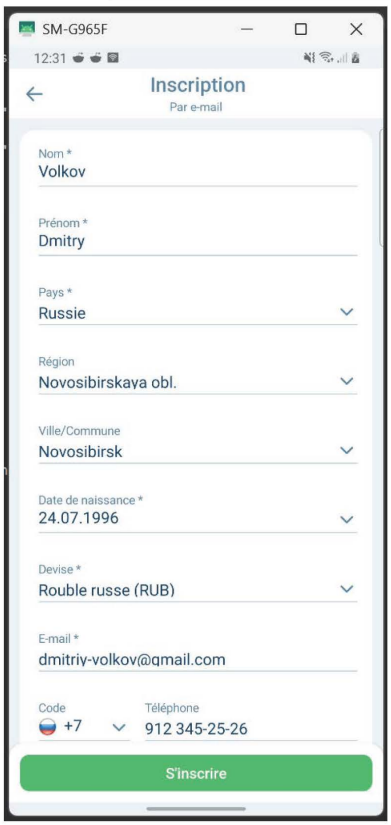

UTILISATEUR VIP: D. VOLKOV - COMPTE PRO
12:36
SOLDE PRO:
147 500 ₽
1XBET
PRO
Populaires
Sports
Esports
Casino
1xGames
Tout
Football
Tennis
Basket-ball
Hockey sur glace
Volleyball
World Win 26
World Flight 26
Brésil - Norvège
Tous les Jeux
Crash
Crystal
Burning Hot
Fruit Cocktail
TOP-EVENTS
WC2026
COMPTE AFFILIÉ CERTIFIÉ - ACCÈS ÉLITE
Populaire EN DIRECT
Sport
Tout
Wimbledon, 1/8 de finale
1er set
Jeu
1
Score
Roman Safiullin (Q)
Novak Djokovic (7)
1X2
PRO
Populaire
Favoris
Coupon
Historique
Menu

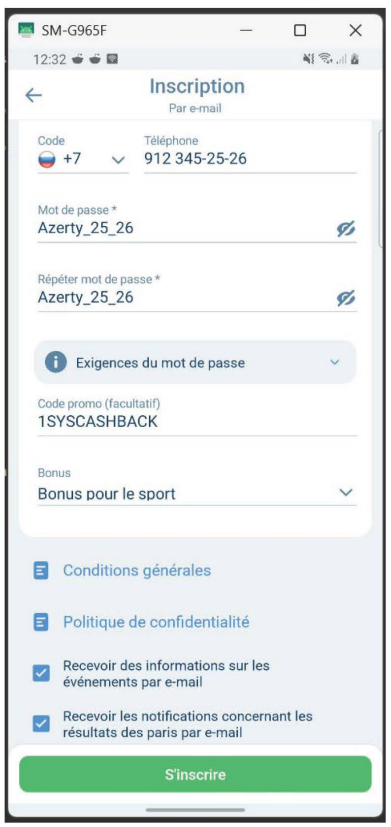

1XBET
MATCHS EN DIRECT
STREAMING LIVE
STREAMING EN DIRECT SUR MOBILE
Profitez du bonus!


CODE PROMO
1SYSCASHBACK


SM-G965F
12:16
1XBET
Se connecter
Inscription
Populaires
Sports
Esports
Casino
1xGames
Tout
Football
Tennis
Basket-ball
Hockey sur glace
Volleyball
World Win 26
World Flight 26
Brésil - Norvège
Tous les Jeux
Crash
Crystal
Burning Hot
Fruit Cocktail
TOP-EVENTS
WC2026
Populaire EN DIRECT
Sport
Tout
Corée du Sud. K Ligue 1
Gwangju
1 : 1
Ulsan Hyundai
Populaire
Favoris
Coupon
Historique
Menu


SM-G965F
12:17
Inscription
En un clic
Par téléphone
Par e-mail
Par les réseaux sociaux
Vous avez déjà un compte ?
Se connecter


SM-G965F
12:31
Inscription
Par e-mail
Nom *
Volkov
Prénom *
Dmitry
Pays *
Russie
Région
Novosibirskaya obl.
Ville/Commune
Novosibirsk
Date de naissance *
24.07.1996
Devise *
Rouble russe (RUB)
E-mail *
dmitriy-volkov@gmail.com
Code
+7
Téléphone
912 345-25-26
S'inscrire


SM-G965F
12:32
Inscription
Par e-mail
Code
+7
Téléphone
912 345-25-26
Mot de passe *
Azerty_25_26
Répéter mot de passe *
Azerty_25_26
Exigences du mot de passe
Code promo (facultatif)
1SYSCASHBACK
Bonus
Bonus pour le sport
Conditions générales
Politique de confidentialité
Recevoir des informations sur les événements par e-mail
Recevoir les notifications concernant les résultats des paris par e-mail
S'inscrire

On peut remarquer qu'aucune condition ne s'applique à a_0 et a_1, ce qui signifie que quelles que soient les valeurs prises pour a_0 et a_1, on obtient une solution de l'équation d'Airy développable en série entière. Cela n'est guère surprenant car l'ensemble des solutions d'une équation différentielle linéaire du second ordre homogène est un $\mathbb{R}$-espace vectoriel de dimension 2. Par ailleurs, on peut remarquer que $a_0 = y(0)$ et $a_1 = y'(0)$, autrement dit que les valeurs de a_0 et a_1 sont déterminées par les conditions initiales ce qui est conforme au théorème 5.10.

On déduit de ce qui précède que les réels a_n s'obtiennent de la manière suivante : a_0 et a_1 sont quelconques, $a_2 = 0$ et pour tout entier $n \geqslant 3$

$$a_n = -\frac{1}{n(n-1)}\, a_{n-3}.$$

Comme $a_2 = 0$, on obtient $a_n = 0$ pour tout $n = 3k+2$ avec $k \in \mathbb{N}$. Si $n = 3k$ avec $k \in \mathbb{N}^*$, on a

$$a_{3k} = -\frac{1}{3k(3k-1)}\, a_{3(k-1)}$$

et on vérifie par récurrence que pour tout $k \in \mathbb{N}^*$,

$$a_{3k} = (-1)^k \prod_{i=1}^{k} \frac{1}{3i(3i-1)}\, a_0 = \frac{(-1)^k a_0}{(3k)!} \prod_{i=1}^{k} (3i-2).$$

Si $n = 3k+1$ avec $k \in \mathbb{N}^*$, on a

$$a_{3k+1} = -\frac{1}{(3k+1)3k}\, a_{3k-2} = -\frac{1}{3k(3k+1)}\, a_{3(k-1)+1}$$

et on vérifie par récurrence que pour tout $k \in \mathbb{N}^*$,

$$a_{3k+1} = (-1)^k \prod_{i=1}^{k} \frac{1}{3i(3i+1)}\, a_1 = \frac{(-1)^k a_1}{(3k+1)!} \prod_{i=1}^{k} (3i-1).$$

Les solutions de l'équation d'Airy sont donc de la forme :

$$\begin{aligned} \sum_{n=0}^{+\infty} a_n x^n &= \sum_{k=0}^{+\infty} a_{3k} x^{3k} + \sum_{k=0}^{+\infty} a_{3k+1} x^{3k+1} \\ &= a_0 \left(1 + \sum_{k=1}^{+\infty} (-1)^k \frac{\prod_{i=1}^{k} (3i-2)}{(3k)!}\, x^{3k} \right) \\ &\quad + a_1 \left(x + \sum_{k=1}^{+\infty} (-1)^k \frac{\prod_{i=1}^{k} (3i-1)}{(3k+1)!}\, x^{3k+1} \right). \end{aligned}$$

Une base du $\mathbb{R}$-espace vectoriel des solutions de l'équation d'Airy est (f_0, f_1) où f_0 et f_1 sont les deux applications de $\mathbb{R}$ dans $\mathbb{R}$, sommes sur $\mathbb{R}$ des séries entières lacunaires

$$1 + \sum_{k=1}^{+\infty} (-1)^k \frac{\prod_{i=1}^{k} (3i-2)}{(3k)!}\, x^{3k} \quad \text{et} \quad x + \sum_{k=1}^{+\infty} (-1)^k \frac{\prod_{i=1}^{k} (3i-1)}{(3k+1)!}\, x^{3k+1}.$$

FROBENIUS, Ferdinand Georg (1849, Charlottenbourg - 1917, Berlin).



Frobenius fut l'élève de Weierstrass à l'Université de Berlin avant d'obtenir un poste de Professeur à l'École polytechnique fédérale de Zurich. Il est l'un des premiers à s'intéresser à la théorie des groupes. Il travaille aussi en algèbre linéaire et donne en 1878 la première démonstration générale du théorème de Cayley-Hamilton. En analyse, il étudie les fonctions elliptiques et les équations aux dérivées partielles, et s'intéresse à la théorie des nombres, en particulier à la fonction[31] Zeta de Riemann.

Plus généralement, on appelle équation d'Airy une équation différentielle de la forme

$$y''(x) + \gamma x\, y(x) = 0$$

où γ désigne un réel strictement positif fixé. Il est aisé de vérifier en reprenant ce qui vient d'être fait qu'une base du $\mathbb{R}$-espace vectoriel des solutions de cette nouvelle équation d'Airy est (f_0, f_1) où f_0 et f_1 sont les deux applications de $\mathbb{R}$ dans $\mathbb{R}$, sommes sur $\mathbb{R}$ des séries entières lacunaires[30]

$$1+\sum_{k=1}^{+\infty}(-1)^k \frac{\gamma^k \prod_{i=1}^{k}(3i-2)}{(3k)!}\, x^{3k} \quad \text{et} \quad x+\sum_{k=1}^{+\infty}(-1)^k \frac{\gamma^k \prod_{i=1}^{k}(3i-1)}{(3k+1)!}\, x^{3k+1}.$$

On remarquera que la méthode présentée ici fournit les solutions d'une équation différentielle comme somme de séries entières et qu'il est en général malaisé (voire impossible) d'en déduire l'expression des solutions à l'aide des fonctions usuelles.

Par ailleurs, signalons que si les fonctions $\alpha_1, \ldots, \alpha_n$ et β sont développables en série entière sur un voisinage V de $x_0 \in \mathbb{R}^*$ alors le changement de variable $t = x - x_0$ permet d'établir que les solutions de l'équation différentielle considérée sont développables en série entière sur V et on peut procéder comme dans l'exemple de l'équation d'Airy présenté.

5.4.2 La méthode de Frobenius

Considérons à présent une équation différentielle linéaire d'ordre p ($p \in \mathbb{N}^*$) de la forme

$$\alpha_0(x)y^{(p)}(x) + \alpha_1(x)\, y^{(p-1)}(x) + \ldots + \alpha_{p-1}(x)\, y'(x) + \alpha_p(x)\, y(x) = \beta(x) \quad (9)$$

[30] On vérifiera en utilisant par exemple le critère de D'Alembert que ces séries entières ont un rayon de convergence infini.

[31] On pourra consulter le chap. 3 p. 129 pour une étude de la fonction Zeta de Riemann.

où les fonctions réelles de la variable réelle $\alpha_0, \alpha_1, \ldots, \alpha_n$ et β sont supposées développables en série entière sur un voisinage $]-R, R[$ $(R \in \mathbb{R}_+^* \cup \{+\infty\})$ de l'origine. Cette équation n'est pas écrite sous forme normalisée et deux situations peuvent se produire selon l'expression de α_0.

1. Si $\alpha_0(0) \neq 0$ alors il existe $R_0 \in \mathbb{R}_+^* \cup \{+\infty\}$ tel que les fonctions $\frac{\alpha_1}{\alpha_0}, \ldots, \frac{\alpha_n}{\alpha_0}$ et $\frac{\beta}{\alpha_0}$ soient développables en série entière sur l'intervalle $]-R_0, R_0[$. On peut appliquer la méthode précédente à l'équation normalisée issue de (9). On dit que 0 est un *point régulier* pour l'équation différentielle.
2. Si $\alpha_0(0) = 0$ alors la méthode précédente ne s'applique plus et il n'est pas certain que l'équation différentielle (9) admette des solutions développables en série entière (elle peut toutefois en admettre). On dit que 0 est un *point singulier* pour l'équation différentielle.

Exemples

1. Le point 0 est un point singulier de l'équation différentielle du second ordre $x^2y''(x)+xy'(x)+(x^2-1)y(x) = 0$. Recherchons une solution de cette équation qui soit développable en série entière dans un voisinage de l'origine. Formellement, les coefficients $(a_n)_{n\in\mathbb{N}}$ de la série entière $\sum_n a_n x^n$ solution vérifient les relations suivantes obtenues par la méthode exposée précédemment :

$$a_0 = 0, \quad 2a_1 = 0 \quad \text{et pour } n \geqslant 2 \quad a_{n+2} + (n+1)(n-1)a_n = 0.$$

On en déduit que $a_n = 0$ pour tout $n \in \mathbb{N}$ et la fonction nulle est la seule solution de l'équation développable en série entière dans un voisinage de l'origine.

2. Le point 0 est un point singulier de l'équation différentielle du second ordre $xy''(x) + (x+1)y'(x) + y(x) = 0$ mais cette équation admet des solutions développables en série entière dans un voisinage de l'origine. On vérifiera, en utilisant la méthode présentée dans la section 5.4.1, que la somme de la série entière $\sum_n a_0 \frac{(-1)^n}{n!} x^n$ (où a_0 désigne une constante réelle) est solution de l'équation différentielle. Dans ce cas particulier, on peut expliciter la somme de la série entière solution ; il s'agit de la fonction $x \mapsto a_0\, \mathrm{e}^{-x}$.

Plus généralement, si les fonctions réelles de la variable réelle $\alpha_0, \alpha_1, \ldots, \alpha_n$ et β sont supposées développables en série entière sur un voisinage V de x_0, le changement de variable $t = x - x_0$ permet de se ramener à la situation d'un point régulier ou singulier situé en 0.

Dans le cas où 0 est un point singulier, la recherche de solutions à l'équation différentielle (9) dans un voisinage de 0 peut s'effectuer en ayant recours à la méthode de Frobenius que nous allons exposer dans le cas d'une équation différentielle d'ordre 2.

Considérons une équation différentielle linéaire du second ordre de la forme

$$\alpha_0(x)y''(x) + \alpha_1(x)y'(x) + \alpha_2(x)y(x) = 0 \tag{10}$$

dont 0 est un point singulier. La méthode de Frobenius consiste à rechercher des solutions de l'équation différentielle (10) qui sont de la forme

$y(x) = x^r \sum_{n=0}^{+\infty} a_n x^n$ où r est un paramètre réel et[32] $a_0 \neq 0$. On montre[33] que sous l'hypothèse que les deux fonctions $\aleph_1 : x \mapsto x\alpha_1(x)/\alpha_0(x)$ et $\aleph_2 : x \mapsto x^2\alpha_2(x)/\alpha_0(x)$ sont développables en série entière dans un voisinage de l'origine[34], il existe au moins un réel r et une fonction f développable en série entière dans un voisinage de l'origine vérifiant $f(0) \neq 0$ tel que l'application $x \in V \cap \mathbb{R}_+^* \mapsto x^r f(x)$ soit solution de l'équation différentielle (10) sur $V \cap \mathbb{R}_+^*$.

Présentons l'utilisation de la méthode de Frobenius en considérant l'équation de Bessel d'ordre ν

$$x^2 y''(x) + x\, y(x) + (x^2 - \nu^2)\, y(x) = 0 \tag{11}$$

où ν désigne un réel positif fixé. La fonction $\aleph_1$ est la fonction constante égale à 1 et $\aleph_2 : x \mapsto x^2 - \nu^2$; ces deux fonctions sont développables en série entière dans un voisinage de l'origine. Recherchons une solution de l'équation de Bessel qui soit de la forme

$$y(x) = x^r \sum_{n=0}^{+\infty} a_n x^n = \sum_{n=0}^{+\infty} a_n x^{n+r} \tag{12}$$

où r désigne un réel fixé à déterminer. Formellement, on a

$$y'(x) = \sum_{n=0}^{+\infty} a_n (n+r) x^{n+r-1} \quad \text{et} \quad y''(x) = \sum_{n=0}^{+\infty} a_n (n+r)(n+r-1) x^{n+r-2}$$

et en substituant ces relations dans le membre de gauche de l'équation (11), on obtient :

$$\begin{aligned} &x^2 y''(x) + x\, y(x) + (x^2 - \nu^2)\, y(x) \\ &\quad = \sum_{n=0}^{+\infty} a_n ((n+r)^2 - \nu^2) x^{r+n} + \sum_{n=0}^{+\infty} a_n x^{r+n+2} \\ &\quad = x^r \Big(a_0(r^2 - \nu^2) + a_1((r+1)^2 - \nu^2)x + \sum_{n=2}^{+\infty} \left(a_n((n+r)^2 - \nu^2) + a_{n-2} \right) x^n \Big). \end{aligned}$$

Pour que l'équation (11) soit satisfaite pour tout $x \in \mathbb{R}$, il faut donc que

$$\begin{cases} a_0(r^2 - \nu^2) &= 0 \\ a_1((r+1)^2 - \nu^2) &= 0 \\ a_n((n+r)^2 - \nu^2) + a_{n-2} &= 0 \quad \text{pour tout entier } n \geqslant 2. \end{cases} \tag{13}$$

---

[32] La situation où $a_0 = 0$ revient à considérer la même forme de solution avec $r+1$ au lieu de r comme paramètre.

[33] Voir par exemple K. Yosida, *Équations Différentielles et Intégrales*, Dunod, Paris, 1971.

[34] Le point 0 est alors appelé *point singulier régulier* de l'équation différentielle (10).

Comme $a_0 \neq 0$, la première équation impose que le paramètre r doit être fixé tel que $r^2 = \nu^2$. Autrement dit, il existe une solution de la forme (12) à l'équation de Bessel (11) pour $r = \nu$ et pour $r = -\nu$ et uniquement dans ces deux cas. La troisième équation quant à elle indique que pour tout entier $n \geqslant 2$ on a

$$a_n = -\frac{1}{n(n+2r)} a_{n-2}. \tag{14}$$

Distinguons les deux situations correspondant à $r = \nu$ et $r = -\nu$.

Cas où $r = \nu$

La seconde équation du système (13) devient $a_1(1+2\nu) = 0$ et comme $\nu \geqslant 0$ on a nécessairement $a_1 = 0$. La relation (14) indique alors que pour tout entier n impair $a_n = 0$. Si n est un entier pair tel que $n = 2k$ avec $k \in \mathbb{N}$, la relation (14) fournit par récurrence l'expression

$$a_{2k} = \frac{(-1)^k a_0}{2^{2k}\, k!\, (1+\nu)(2+\nu)\cdots(k+\nu)} \qquad \forall k \in \mathbb{N}^*.$$

Cas où $r = -\nu$

La seconde équation du système (13) devient $a_1(1-2\nu) = 0$. Si $\nu \neq \frac{1}{2}$ on a nécessairement $a_1 = 0$. La relation (14) indique que si n est un entier impair alors $a_n = 0$ et que si n est un entier pair de la forme $n = 2k$ avec $k \in \mathbb{N}$ alors

$$a_n = a_{2k} = \frac{(-1)^k a_0}{2^{2k}\, k!\, (1-\nu)(2-\nu)\cdots(k-\nu)}.$$

Si $\nu = \frac{1}{2}$, a_1 n'est pas nécessairement nul mais on vérifie que la prise en compte de cette possibilité fournit des solutions déjà exprimées.

On peut conclure que les solutions de l'équation de Bessel d'ordre ν de la forme (12) sont les fonctions

$$J_\nu : x \in \mathbb{R}_+^* \longmapsto a_0 x^\nu S_\nu(x) \quad \text{et} \quad J_{-\nu} : x \in \mathbb{R}_+^* \longmapsto a_0 x^{-\nu} S_{-\nu}(x)$$

où S_ν et $S_{-\nu}$ sont les sommes des séries entières de rayon de convergence $+\infty$ définies par

$$S_\nu : x \in \mathbb{R} \longmapsto 1 + \sum_{k=1}^{+\infty} \frac{(-1)^k}{2^{2k}\, k!\, (1+\nu)(2+\nu)\cdots(k+\nu)} x^{2k}$$

$$S_{-\nu} : x \in \mathbb{R} \longmapsto 1 + \sum_{k=1}^{+\infty} \frac{(-1)^k}{2^{2k}\, k!\, (1-\nu)(2-\nu)\cdots(k-\nu)} x^{2k}.$$

On appelle *fonctions de Bessel de première espèce* les deux fonctions J_ν et $J_{-\nu}$ correspondant au choix

$$a_0 = \frac{1}{2^n\, \Gamma(1+\nu)} \quad \text{où}^{(35)} \quad \Gamma(1+\nu) = \int_0^{+\infty} t^\nu \, \mathrm{e}^{-t} \, \mathrm{d}t.$$

(35) La fonction Γ est connue sous le nom de *fonction Gamma d'Euler*.

On peut montrer que si $\nu \in \mathbb{R}^+ \setminus \mathbb{N}$ alors les deux fonctions J_ν et $J_{-\nu}$ sont linéairement indépendantes et par conséquent que $\{J_\nu, J_{-\nu}\}$ est une base de l'espace vectoriel (de dimension deux) des solutions de l'équation de Bessel d'ordre ν. Autrement dit, si ν n'est pas un entier naturel, les solutions de l'équation de Bessel d'ordre ν sont toutes les fonctions de la forme :

$$x \in \mathbb{R}_+^* \longmapsto C_1 J_\nu(x) + C_2 J_{-\nu}(x),$$

où C_1 et C_2 désignent deux réels quelconques. Par contre, si ν est un entier naturel, $\{J_\nu, J_{-\nu}\}$ ne forme plus une base de l'espace vectoriel des solutions car on montre que dans ce cas $J_{-\nu}(x) = (-1)^\nu J_\nu(x)$ pour tout $x \in \mathbb{R}_+^*$.

Dans le cas où $\nu = \frac{1}{2}$, il est possible d'exprimer $J_{\frac{1}{2}}$ à l'aide des fonctions usuelles. Pour tout $x \in \mathbb{R}_+^*$, on a

$$J_{\frac{1}{2}}(x) = a_0\sqrt{x}\left(1 + \sum_{k=1}^{+\infty} \frac{(-1)^k}{2^{2k}\, k!\, (1+\frac{1}{2})(2+\frac{1}{2})\cdots(k+\frac{1}{2})}\, x^{2k}\right)$$

et

$$2^{2k}\, k!\, (1+\tfrac{1}{2})(2+\tfrac{1}{2})\cdots(k+\tfrac{1}{2}) = 2^k k!\ (3 \times 5 \times \cdots \times (2k+1)) = (2k+1)!.$$

On en déduit que

$$J_{\frac{1}{2}}(x) = \frac{a_0}{\sqrt{x}}\left(x + \sum_{k=1}^{+\infty} \frac{(-1)^k}{(2k+1)!}\, x^{2k+1}\right) = \frac{a_0}{\sqrt{x}}\, \sin(x).$$

L'exercice 5.10 est consacré à une étude détaillée de l'équation de Bessel d'indice 0.

5.4.3 Équations différentielles et développement en série entière

Nous avons vu dans la partie précédente comment résoudre certaines équations différentielles en essayant d'obtenir une représentation des solutions sous forme de série entière.

Cette méthode peut également être utilisée pour obtenir le développement en série entière d'une fonction f donnée. Pour cela, on procède de la manière suivante. On considère une équation différentielle dont f est l'unique solution pour des conditions initiales données. On recherche les solutions de l'équation différentielle développables en série entière. Si une de ces séries entières vérifie les conditions initiales, alors par unicité de la solution elle correspond au développement en série entière de f.

Utilisons cette approche pour obtenir le développement en série entière de l'application $f : x \in]-1, +\infty[\mapsto (1+x)^\alpha$ où $\alpha \in \mathbb{R}$. Cette application est l'unique solution de l'équation différentielle linéaire du premier ordre[36]

$$(1+x)y'(x) - \alpha y(x) = 0 \quad \forall x \in]-1, +\infty[\tag{15}$$

[36] Voir le chap. 20 du *Cours de première année* pour l'étude de ce type d'équations différentielles.

satisfaisant à la condition initiale $y(0) = 1$. On en déduit que f sera développable en série entière si et seulement si il existe une série entière $\sum_n a_n x^n$ de rayon de convergence R dont la somme est solution l'équation différentielle (15) sur $]-R, R[$ et qui satisfait la condition initiale.

On établit par la méthode exposée dans la section 5.4.1 que les coefficients $(a_n)_n$ de la série entière doivent satisfaire les équations

$$(n+1)a_{n+1} + (n-\alpha)a_n = 0 \qquad \forall n \in \mathbb{N}^*$$

et $a_0 = 1$. On obtient par récurrence l'expression suivante pour les coefficients :

$$a_n = \frac{\alpha \times (\alpha-1) \times \cdots \times (\alpha-n+1)}{n!} \qquad \forall n \in \mathbb{N}^*.$$

La règle de D'Alembert indique que le rayon de convergence de la série entière $\sum_n a_n x^n$ vaut 1. On en conclut que :

$$\forall x \in]-1,1[\qquad (1+x)^\alpha = 1 + \sum_{n=1}^{+\infty} \frac{\alpha \times (\alpha-1) \times \cdots \times (\alpha-n+1)}{n!} x^n.$$

Exercice 5.9
1 - Montrer que l'application $f : x \in]-1,1[\mapsto \arcsin(x)^2$ est solution de l'équation différentielle $(\mathcal{E})$: $(1-x^2)y''(x) - xy'(x) = 2$.
2 - Rechercher les solutions de $(\mathcal{E})$ développables en série entière dans un voisinage de l'origine et en déduire le développement en série entière de f.

5.5 Exercices de synthèse

Exercice 5.10 ***Fonction de* BESSEL *d'indice 0***

Considérons *l'équation différentielle de* BESSEL *d'indice* 0 :

$$(\mathcal{E}) \quad x^2 y''(x) + x y'(x) + x^2 y(x) = 0 \qquad \forall x \in \mathbb{R}_+^*.$$

1 - Montrer que $(\mathcal{E})$ admet pour solutions développables en série entière dans un voisinage de l'origine, les fonctions λJ_0 où $\lambda \in \mathbb{R}$ et

$$J_0 : x \in \mathbb{R}_+^* \longmapsto \sum_{n=0}^{+\infty} \frac{(-1)^n}{(n!)^2} \left(\frac{x}{2}\right)^{2n}$$

est la *fonction de* BESSEL *d'indice* 0.

2 - a) Pour $n \in \mathbb{N}^*$, on note $\alpha_n = \sum_{k=1}^n \frac{1}{k}$; montrer que $\alpha_n \underset{+\infty}{\sim} \ln(n)$.

b) Montrer que la série entière $\sum_{n\geqslant 1} \frac{(-1)^n \alpha_n}{(n!)^2} \left(\frac{x}{2}\right)^{2n}$ converge sur $\mathbb{R}$. On note S sa somme.

c) Montrer que : $x^2\,S''(x) + x\,S'(x) + x^2\,S(x) = 2xJ_0'(x)$ pour tout $x \in \mathbb{R}$.

d) Montrer que l'application $T_0 : x \in \mathbb{R}_+^* \mapsto \ln(x)J_0(x) - S(x)$ est une solution de $(\mathcal{E})$.

e) Montrer que J_0 et T_0 sont deux solutions de $(\mathcal{E})$ linéairement indépendantes de J_0. (On pourra considérer les limites en 0 de J_0 et T_0.)

Une base de l'espace vectoriel de dimension 2 constitué des solutions de l'équation différentielle de Bessel d'indice 0 *est donc* $\{J_0, T_0\}$. *Il est de coutume de considérer la fonction fonction de Neumann d'indice* 0 *définie sur* $\mathbb{R}_+^*$ *par :*

$$\begin{aligned} N_0(x) &= \frac{2}{\pi} T_0(x) + \frac{2}{\pi}(\gamma - \ln(2))J_0(x) \\ &= \frac{2}{\pi}\left(\ln(x/2) + \gamma\right) J_0(x) - \frac{2}{\pi} \sum_{n=1}^{+\infty} \frac{(-1)^n \alpha_n}{(n!)^2} \left(\frac{x}{2}\right)^{2n}, \end{aligned}$$

où γ *désigne la constante d'Euler, comme second élément de base à la place de* T_0. *Les solutions de l'équation différentielle de Bessel d'indice* 0 *sont alors toutes les applications de la forme* $x \in \mathbb{R}_+^* \mapsto \lambda J_0(x) + \mu N_0(x)$ *où* λ *et* μ *désignent deux réels.*

3 - a) Pour $p \in \mathbb{N}$, on note $I_p = \int_0^\pi \sin^{2p}(\theta)\,\mathrm{d}\theta$. En utilisant une intégration par parties, montrer que $I_p = (2p-1)I_{p-1} - (2p-1)I_p$ pour tout $p \in \mathbb{N}^*$. En déduire que $I_p = \dfrac{(2p)!\,\pi}{2^{2p}(p!)^2}$.

b) Montrer que la série de fonctions $\sum_n f_n$ de terme général $f_n : \theta \in \mathbb{R} \mapsto (-1)^n \dfrac{x_0^{2n}}{(2n)!} \sin^{2n}(\theta)$ où $x_0 \in \mathbb{R}_+^*$ est fixé, converge uniformément sur $\mathbb{R}$ et que sa somme est $\phi : \theta \in \mathbb{R} \mapsto \cos(x_0 \sin(\theta))$.

c) En déduire que pour tout $x \in \mathbb{R}_+^*$ on a $J_0(x) = \frac{1}{\pi}\int_0^\pi \cos(x\sin(\theta))\,\mathrm{d}\theta$.

Exercice 5.11 *Formule d'*Euler-Abel

L'objet de cet exercice est d'établir la formule d'Euler-Abel et d'étudier deux de ses usages possibles : l'un pour accélérer la convergence d'une série vers sa somme[37] *et l'autre pour calculer la somme des séries entières de la forme* $\sum_n P(n)z^n$ *où* P *désigne un polynôme de* $\mathbb{C}[X]$.

---

[37] Pour une utilisation en électromagnétisme de la formule d'Euler-Abel afin d'accélérer la convergence d'une série vers sa somme, on pourra consulter l'article suivant : Hung-Mou Lee, Electromagnetic Scattering of Tubular Cylindrical Structure-Double Series Formulation and Some Results, *IEEE Transactions on Antennas and Propogation*, vol. 35, n° 4, 1987.

1 - a) Calculer le rayon de convergence de la série entière $\sum_{n\geqslant 1} n^\pi z^n$.

b) On souhaite approcher $S(z) = \sum_{n=1}^{+\infty} n^\pi z^n$ pour $z = -0.99$ par $S_N(-0.99) = \sum_{n=1}^{N} n^\pi (-0.99)^n$. Montrer que la série $\sum_{n\geqslant 1} n^\pi (-0.99)^n$ satisfait au critère d'Abel pour les séries alternées puis donner un majorant de $|S(-0.99) - S_N(-0.99)|$ en fonction de la valeur de N.

c) En utilisant MAPLE, déterminer à partir de quelle valeur de N le réel $S_N(-0.99)$ constitue une valeur approchée de $S(-0.99)$ avec une erreur inférieure à 10^{-6}. Comment obtenir avec MAPLE cette valeur approchée de $S(-0.99)$? Renouveler le calcul avec $z = -0.999$ et commenter les valeurs obtenues.

2 - Soit $\sum_n a_n z^n$ une série entière complexe de rayon de convergence $R = 1$. On désigne par f la somme de cette série entière et par g l'application définie pour $z \in D(0,1) \setminus \{0\}$ par $g(z) = \frac{1}{z}(f(z) - a_0)$.

a) Montrer que g est prolongeable en une application continue sur $D(0,1)$ (que l'on désignera encore par g dans le reste de l'exercice).

b) Montrer que pour tout $z \in D(0,1)$, on a

$$(1-z)\,g(z) = a_0 + \sum_{n=0}^{+\infty} \Delta^{[1]} a_n \; z^n \quad \text{où} \quad \Delta^{[1]} a_n = a_{n+1} - a_n.$$

c) En déduire que pour tout $z \in D(0,1)$, on a :

$$f(z) = \frac{a_0}{1-z} + \frac{z}{1-z} \sum_{n=0}^{+\infty} \Delta^{[1]} a_n \; z^n.$$

d) En utilisant un raisonnement par récurrence, établir la formule d'Euler-Abel à l'ordre p, où $p \in \mathbb{N}^*$:

$$f(z) = \sum_{k=0}^{p-1} \Delta^{[k]} a_0 \; \frac{z^k}{(1-z)^{k+1}} + \frac{z^p}{(1-z)^p} \sum_{n=0}^{+\infty} \Delta^{[p]} a_n \; z^n \qquad \forall z \in D(0,1)$$

où pour tout $n \in \mathbb{N}$ on a défini : $\Delta^{[0]} a_n = a_n$, $\Delta^{[1]} a_n = a_{n+1} - a_n$ et $\Delta^{[k]} a_n = \Delta^{[k-1]} a_{n+1} - \Delta^{[k-1]} a_n$ pour tout entier $k \geqslant 2$.

3 - On suppose que $a_n = n^\alpha$ où $\alpha \in \mathbb{R} \setminus \mathbb{Z}$ pour $n \in \mathbb{N}^*$ et que $a_0 = 0$.

a) Montrer que $(n+1)^\alpha = n^\alpha + \alpha\, n^{\alpha-1} + \frac{1}{2}\alpha(\alpha-1)\, n^{\alpha-2} + \mathcal{O}_{+\infty}(n^{\alpha-2})$.

b) En déduire que $\Delta^{[1]} a_n \underset{+\infty}{\sim} \alpha\, n^{\alpha-1}$ puis que $\Delta^{[2]} a_n \underset{+\infty}{\sim} \alpha(\alpha-1)\, n^{\alpha-2}$.

c) Comment interprétez-vous ces résultats ? Quel est l'intérêt de la formule d'Euler-Abel ?

4 - a) Calculer le rayon de convergence de la série entière $\sum_n n\, z^n$ et montrer que sur le disque de convergence de la série on a $\sum_{n=0}^{+\infty} n z^n = \frac{z}{(1-z)^2}$.

b) En utilisant la formule d'Euler-Abel, en déduire la somme de la série $\sum_n n^2 z^n$.

c) Soit P un polynôme de $\mathbb{C}[X]$ de degré $p-1$ ($p \in \mathbb{N}^*$) ; déterminer le rayon de convergence R de la série entière complexe $\sum_n a_n z^n$ où $a_n = P(n)$.

d) Montrer $\Delta^{[p]} P(n) = 0$ pour tout $n \in \mathbb{N}$. (On pourra utiliser un raisonnement par récurrence portant sur le degré de p.)

e) En utilisant la formule d'Euler-Abel à l'ordre p, montrer que la somme de la série $\sum_n P(n) z^n$ sur $D(0,1)$ est

$$\sum_{n=0}^{+\infty} P(n) z^n = \sum_{k=0}^{p-1} \Delta^{[k]} P(0) \frac{z^k}{(1-z)^{k+1}}.$$

f) Implémenter sous MAPLE un algorithme mettant en oeuvre la formule d'Euler-Abel pour calculer la somme d'une série entière $\sum_n P(n) z^n$ où P désigne un polynôme de $\mathbb{C}[X]$.

5.6 Solutions des exercices

Solution de l'exercice 5.1

1 - La série géométrique alternée $\sum_{n\geqslant 1}\frac{(-1)^n}{n}$ est semi-convergente. On déduit de la proposition 5.3 que le rayon de convergence de la série entière complexe $\sum_{n\geqslant 1}\frac{1}{n}z^n$ est $R=1$. Comme $1=n\times\frac{1}{n}$ et que le rayon de convergence de la série entière $\sum_{n\geqslant 1}\frac{1}{n}z^n$ est $R=1$, on déduit de la proposition 5.2 que le rayon de convergence de la série entière $\sum_{n\geqslant 1}z^n$ est $R=1$. Par ailleurs, comme $\frac{1}{n^2}=\frac{1}{n}\times\frac{1}{n}$ et que le rayon de convergence de la série entière $\sum_{n\geqslant 1}\frac{1}{n}z^n$ est $R=1$, on déduit de la proposition 5.2 que le rayon de convergence de la série entière $\sum_n\frac{1}{n^2}z^n$ est $R=1$.

2 - Soit z un complexe appartenant au cercle de convergence de ces séries. On a $z=\mathrm{e}^{\mathrm{i}\theta}$ avec $\theta\in[0,2\pi[$. Sur le cercle de convergence, la série $\sum_n z^n$ s'exprime donc sous la forme $\sum_n \mathrm{e}^{\mathrm{i}n\theta}$. Il s'agit d'une série géométrique de raison $\rho=\mathrm{e}^{\mathrm{i}\theta}$. Comme ρ est de module 1, cette série diverge quelle que soit la valeur de θ (voir p. 3). On en conclut que la série entière $\sum_n z^n$ ne converge en aucun point du cercle de convergence.

Sur le cercle de convergence, la série $\sum_n\frac{1}{n}z^n$ est une série trigonométrique qui s'exprime sous la forme $\sum_n\frac{1}{n}\mathrm{e}^{\mathrm{i}n\theta}$. Le théorème 1.5 p. 48 indique que cette série converge pour tout $\theta\in]0,2\pi[$. Pour $\theta=0$, cette série est la série harmonique $\sum_n\frac{1}{n}$ qui diverge.

Sur le cercle de convergence, la troisième série est une série trigonométrique de la forme $\sum_n\frac{1}{n^2}\mathrm{e}^{\mathrm{i}n\theta}$. Pour tout $n\in\mathbb{N}^*$ on a $\left|\frac{1}{n^2}\mathrm{e}^{\mathrm{i}n\theta}\right|=\frac{1}{n^2}$ et la série $\sum_n\frac{1}{n^2}$ converge. Cette série est donc absolument convergente sur le cercle de convergence.

Solution de l'exercice 5.2

1 - Pour $n\in\mathbb{N}^*$, posons $a_n=n^{\sqrt{n}}$; on a

$$\sqrt[n]{a_n}=\left(n^{\sqrt{n}}\right)^{1/n}=n^{1/\sqrt{n}}=\exp(\ln(n)/\sqrt{n}).$$

Comme $\lim_{n\to+\infty}\ln(n)/\sqrt{n}=0$, on a $\lim_{n\to+\infty}\sqrt[n]{|a_n|}=1$ et le rayon de convergence de la série entière $\sum_{n\geqslant 1}n^{\sqrt{n}}z^n$ est $R=1$.

2 - Pour $n\in\mathbb{N}^*$, posons $a_n=n^\alpha$; on a $\sqrt[n]{a_n}=n^{\alpha/n}=\exp(\alpha\ln(n)/n)$. Comme $\lim_{n\to+\infty}\ln(n)/n=0$, on a $\lim_{n\to+\infty}\sqrt[n]{|a_n|}=1$ et le rayon de convergence de la série entière $\sum_{n\geqslant 1}n^\alpha z^n$ est $R=1$ quel que soit $\alpha\in\mathbb{N}^*$.

3 - Pour $n\in\mathbb{N}^*$, posons $a_n=n^{\ln(n)}/\ln(n)^n$; on a

$$\sqrt[n]{a_n}=\frac{1}{\ln(n)}n^{\ln(n)/n}=\frac{1}{\ln(n)}\exp(\ln(n)^2/n).$$

Comme $\lim_{n\to+\infty}\ln(n)^2/n=0$, on a $\lim_{n\to+\infty}\sqrt[n]{|a_n|}=0$ et le rayon de convergence de la série entière $\sum_{n\geqslant 1}\frac{n^{\ln(n)}}{\ln(n)^n}z^n$ est $R=+\infty$.

Solution de l'exercice 5.3

1 - Pour $n \in \mathbb{N}^*$, posons $a_n = \mathrm{ch}(n)/n$; en utilisant la formule de trigonométrie hyperbolique $\mathrm{ch}(n+1) = \mathrm{ch}(n)\,\mathrm{ch}(1) + \mathrm{sh}(1)\,\mathrm{sh}(n)$, on obtient

$$\frac{|a_{n+1}|}{|a_n|} = \frac{\mathrm{ch}(n+1)}{\mathrm{ch}(n)}\,\frac{n}{n+1} = (\mathrm{ch}(1) + \mathrm{th}(n)\,\mathrm{sh}(1))\,\frac{n}{n+1}.$$

On a $\lim_{n\to+\infty} \mathrm{th}(n) = 1$ d'où $\lim_{n\to+\infty} |a_{n+1}|/|a_n| = \mathrm{ch}(1) + \mathrm{sh}(1) = \mathrm{e}$. Le rayon de convergence de la série entière $\sum_{n\geqslant 1} \frac{1}{n}\,\mathrm{ch}(n)\,z^n$ est donc $R = 1/\mathrm{e}$.

2 - Pour $n \in \mathbb{N}$ avec $n \geqslant 2$, posons $a_n = \ln(n)$; on a

$$\frac{|a_{n+1}|}{|a_n|} = \frac{\ln(n+1)}{\ln(n)} = \frac{\ln(n) + \ln(1+1/n)}{\ln(n)} = 1 + \frac{\ln(1+1/n)}{\ln(n)}.$$

Or, $\ln(1+1/n) \underset{+\infty}{\sim} 1/n$ et $\lim_{n\to+\infty} \frac{1}{n\ln(n)} = 0$ donc $\lim_{n\to+\infty} |a_{n+1}|/|a_n| = 1$. Le rayon de convergence de la série entière $\sum_{n\geqslant 2} \ln(n)\,z^n$ est donc $R = 1$.

3 - Pour $n \in \mathbb{N}^*$, posons $a_n = (\sqrt{n})^n$; on a

$$\frac{|a_{n+1}|}{|a_n|} = \frac{(\sqrt{n+1})^{n+1}}{(\sqrt{n})^n} = \sqrt{n+1}\,\frac{(\sqrt{n+1})^n}{(\sqrt{n})^n} = \sqrt{n+1}\left(1+\frac{1}{n}\right)^{\frac{n}{2}}.$$

Comme $\lim_{n\to+\infty} \left(1+\frac{1}{n}\right)^n = \mathrm{e}$, on en déduit que $\lim_{n\to+\infty} \left(1+\frac{1}{n}\right)^{\frac{n}{2}} = \sqrt{\mathrm{e}}$ puis que $\lim_{n\to+\infty} |a_{n+1}|/|a_n| = +\infty$. Ainsi[38], le rayon de convergence de la série entière $\sum_{n\geqslant 1} (\sqrt{n})^n\,z^n$ est $R = 0$.

Solution de l'exercice 5.4

La série $\sum_n \frac{(-1)^n}{(n!)^2}\left(\frac{z}{2}\right)^{2n}$ est une série entière lacunaire. Appliquons le critère de D'Alembert pour les séries numériques à la série $\sum_n \frac{(-1)^n}{(n!)^2}\left(\frac{z}{2}\right)^{2n}$ pour z fixé dans $\mathbb{C}$. En notant $u_n = \frac{(-1)^n}{(n!)^2}\left(\frac{z}{2}\right)^{2n}$, on obtient pour tout $n \in \mathbb{N}$:

$$\frac{|u_{n+1}|}{|u_n|} = \frac{(n!)^2}{((n+1)!)^2}\left(\frac{z}{2}\right)^2 = \left(\frac{1}{n+1}\frac{z}{2}\right)^2.$$

On a $\lim_{n\to+\infty} \frac{|u_{n+1}|}{|u_n|} = 0$ donc la série numérique complexe $\sum_n u_n$ converge absolument quel que soit $z \in \mathbb{C}$. On en déduit que la série entière $\sum_n \frac{(-1)^n}{(n!)^2}\left(\frac{z}{2}\right)^{2n}$ converge absolument sur $\mathbb{C}$. Son rayon de convergence est $R = +\infty$. La fonction de Bessel d'ordre 0, qui est la somme de cette série, est donc définie sur $\mathbb{C}$.

---

[38] On peut également obtenir ce résultat en remarquant que pour tout $z \in \mathbb{C}^*$ on a $\lim_{n\to+\infty} |\sqrt{n^n}z^n| = \lim_{n\to+\infty} (\sqrt{n}|z|)^n = +\infty$ de sorte que la série numérique $\sum_{n\geqslant 1} (\sqrt{n}|z|)^n$ diverge grossièrement. Comme une série entière converge absolument sur son disque de convergence, on peut affirmer que $R = 0$.

Solution de l'exercice 5.5

Appliquons le théorème des accroissements finis[39] à la fonction exponentielle complexe entre z et 0 :

$$|\exp(z)-1| = |\exp(z)-\exp(0)| \leqslant |z| \sup_{\zeta\in[0,z]} |\exp(\zeta)| = |z| \sup_{t\in[0,1]} |\exp(tz)|.$$

On a $|\exp(tz)| = \left|e^{t\mathcal{R}e(z)}\, e^{it\mathcal{I}m(z)}\right| = e^{t\mathcal{R}e(z)}$.

- Si $\mathcal{R}e(z) < 0$ la fonction $\phi : t \mapsto e^{t\mathcal{R}e(z)}$ est positive, strictement décroissante sur $[0,1]$ donc $\sup_{t\in[0,1]} e^{t\mathcal{R}e(z)} = 1$;
- si $\mathcal{R}e(z) = 0$ la fonction ϕ est constante égale à 1 ;
- si $\mathcal{R}e(z) > 0$ la fonction ϕ est positive, strictement croissante sur $[0,1]$ donc $\sup_{t\in[0,1]} e^{t\mathcal{R}e(z)} = e^{\mathcal{R}e(z)}$.

Ainsi, dans tous les cas on a $\sup_{t\in[0,1]} e^{t\mathcal{R}e(z)} \leqslant e^{|\mathcal{R}e(z)|}$. Comme pour $z \in D(0,1)$, on a $|\mathcal{R}e(z)| \leqslant |z| \leqslant 1$, on en déduit que $|\exp(z)-1| \leqslant e|z|$.

Solution de l'exercice 5.6

Nous avons établi à l'exercice 5.4 que la série entière $\sum_n \frac{(-1)^n}{(n!)^2}\left(\frac{z}{2}\right)^{2n}$ a pour rayon de convergence $R = +\infty$ et par conséquent que la fonction de Bessel J_0 est définie sur $\mathbb{C}$. Il résulte du théorème 5.4 que J_0 est une application holomorphe sur $\mathbb{C}$.

Pour tout $z \in \mathbb{C}$, on a $J_0(z) = \sum_{p=0}^{+\infty} a_p z^p$ où $a_p = 0$ si p est impair et $a_p = \frac{(-1)^n}{4^n\,(n!)^2}$ si p est pair avec $p = 2n, n \in \mathbb{N}$. D'après le théorème 5.4, pour tout $z \in \mathbb{C}$ on a

$$\begin{aligned} J_0'(z) &= \sum_{p=0}^{+\infty}(p+1)a_{p+1}z^p = \overbrace{\sum_{n=0}^{+\infty}(2n+1)\underbrace{a_{2n+1}}_{=0}z^{2n}}^{\text{indices } p \text{ impairs}} + \overbrace{\sum_{n=0}^{+\infty}(2n+2)a_{2n+2}z^{2n+1}}^{\text{indices } p \text{ pairs}} \\ &= \sum_{n=0}^{+\infty}(2n+2)a_{2n+2}z^{2n+1} = \sum_{n=0}^{+\infty}(2n+2)\frac{(-1)^{n+1}}{4^{n+1}\,((n+1)!)^2}z^{2n+1} \\ &= \sum_{n=0}^{+\infty}\frac{(-1)^{n+1}}{n!\times(n+1)!}\left(\frac{z}{2}\right)^{2n+1}. \end{aligned}$$

On en déduit que pour tout $z \in \mathbb{C}$, on a

$$J_0'(z) = -\sum_{n=0}^{+\infty}\frac{(-1)^n}{n!\times(n+1)!}\left(\frac{z}{2}\right)^{2n+1}.$$

---

[39] Voir le théorème 4.2 p. 201.

Solution de l'exercice 5.7

1 - On a la décomposition en éléments simples[40] suivante :

$$\frac{X}{(X-1)^2(X-2)} = \frac{a}{(X-1)^2} + \frac{b}{X-1} + \frac{c}{X-2}.$$

Le coefficient c s'obtient en multipliant par $X-2$ puis en évaluant cette quantité en 2. On obtient $c = 2$. Le coefficient a s'obtient en multipliant par $(X-1)^2$ puis en évaluant cette quantité en 1. On obtient $a = -1$. Enfin, le coefficient b peut s'obtenir en multipliant par X et en faisant tendre X vers l'infini :

$$\overbrace{\lim_{X\to+\infty} \frac{X^2}{(X-1)^2(X-2)}}^{=0} = \overbrace{\lim_{X\to+\infty} \left(\frac{-X}{(X-1)^2} + \frac{bX}{X-1} + \frac{2X}{X-2}\right)}^{=b+2},$$

d'où $b = -2$. Ainsi

$$\frac{X}{(X-1)^2(X-2)} = \frac{-1}{(X-1)^2} + \frac{-2}{X-1} + \frac{2}{X-2}.$$

2 - On en déduit que pour tout $z \in D(0,1)$, on a

$$f(z) = -\frac{1}{(z-1)^2} - \frac{2}{z-1} + \frac{2}{z-2} = \frac{-1}{(1-z)^2} + \frac{2}{1-z} - \frac{1}{1-\frac{z}{2}}.$$

Or $\dfrac{1}{1-z} = \displaystyle\sum_{n=0}^{+\infty} z^n$ donc $\dfrac{1}{1-\frac{z}{2}} = \displaystyle\sum_{n=0}^{+\infty} \left(\frac{z}{2}\right)^n = \sum_{n=0}^{+\infty} \frac{1}{2^n} z^n$. Par ailleurs, comme le rayon de la série entière $\sum_n z^n$ vaut 1, on a pour tout $z \in D(0,1)$,

$$\frac{1}{(1-z)^2} = \left(\frac{1}{1-z}\right)' = \left(\sum_{n=0}^{+\infty} z^n\right)' = \sum_{n=0}^{+\infty} (z^n)' = \sum_{n=1}^{+\infty} n z^{n-1} = \sum_{n=0}^{+\infty} (n+1) z^n.$$

On en déduit que pour tout $z \in D(0,1)$ on a

$$f(z) = \sum_{n=0}^{+\infty} \left(-(n+1) + 2 - \frac{1}{2^n}\right) z^n = \sum_{n=0}^{+\infty} \left(-n + 1 - \frac{1}{2^n}\right) z^n.$$

Solution de l'exercice 5.8

Pour tous $\alpha \in \mathbb{R}$ et $x \in]-1,1[$, on a

$$(1+x)^\alpha = 1 + \sum_{n=1}^{+\infty} \frac{\alpha(\alpha-1)\dots(\alpha-n+1)}{n!} x^n$$

[40] Voir le chap. 7 du *Cours de première année* pour les techniques de décomposition en éléments simples d'une fraction rationnelle.

et

$$(1-x)^\alpha = 1 + \sum_{n=1}^{+\infty} (-1)^n \, \frac{\alpha(\alpha-1)\dots(\alpha-n+1)}{n!} \, x^n.$$

1 - La fonction arc-sinus est dérivable sur $]-1,1[$ et pour tout $x \in]-1,1[$, on a $\arcsin'(x) = \frac{1}{\sqrt{1-x^2}}$. Or, la fonction $t \mapsto \frac{1}{\sqrt{1-t^2}}$ est développable en série entière dans un voisinage de l'origine et pour tout $t \in]-1,1[$, on a (prendre $\alpha = -\frac{1}{2}$)

$$\frac{1}{\sqrt{1-t^2}} = 1 + \sum_{n=1}^{+\infty} \frac{1 \times 3 \times 5 \times \dots \times (2n-1)}{2^n \, n!} \, t^{2n}.$$

On déduit du corollaire 5.4 que pour tout $x \in]-1,1[$, on a

$$\begin{aligned} \arcsin(x) &= \int_0^x \frac{1}{\sqrt{1-t^2}} \, \mathrm{d}t = x + \sum_{n=1}^{+\infty} \frac{1 \times 3 \times 5 \times \dots \times (2n-1)}{2^n \, n!(2n+1)} \, x^{2n+1} \\ &= x + \sum_{n=1}^{+\infty} \frac{1 \times 3 \times 5 \times \dots \times (2n-1)}{2 \times 4 \times 6 \times \dots \times 2n} \, \frac{x^{2n+1}}{2n+1}. \end{aligned}$$

2 - Comme pour tout $x \in]-1,1[$, on a $\arccos(x) + \arcsin(x) = \frac{\pi}{2}$, on en déduit que la fonction arc-cosinus est développable en série entière dans un voisinage de l'origine et que pour tout $x \in]-1,1[$ on a

$$\arccos(x) = \frac{\pi}{2} - x - \sum_{n=1}^{+\infty} \frac{1 \times 3 \times 5 \times \dots \times (2n-1)}{2 \times 4 \times 6 \times \dots \times 2n} \, \frac{x^{2n+1}}{2n+1}.$$

3 - La fonction argument sinus hyperbolique est dérivable sur $\mathbb{R}$ et pour tout $x \in \mathbb{R}$ on a $\operatorname{argsh}'(x) = \frac{1}{\sqrt{1+x^2}}$. Or, la fonction $t \mapsto \frac{1}{\sqrt{1+t^2}}$ est développable en série entière dans un voisinage de l'origine et pour tout $t \in]-1,1[$, on a (prendre $\alpha = -\frac{1}{2}$)

$$\frac{1}{\sqrt{1+t^2}} = 1 + \sum_{n=1}^{+\infty} (-1)^n \, \frac{1 \times 3 \times 5 \times \dots \times (2n-1)}{2^n \, n!} \, t^{2n}.$$

On déduit du corollaire 5.4 que pour tout $x \in]-1,1[$ on a

$$\begin{aligned} \operatorname{argsh}(x) &= \int_0^x \frac{1}{\sqrt{1+t^2}} \, \mathrm{d}t = x + \sum_{n=1}^{+\infty} (-1)^n \frac{1 \times 3 \times 5 \times \dots \times (2n-1)}{2^n \, n!(2n+1)} \, x^{2n+1} \\ &= x + \sum_{n=1}^{+\infty} (-1)^n \frac{1 \times 3 \times 5 \times \dots \times (2n-1)}{2 \times 4 \times 6 \times \dots \times 2n} \, \frac{x^{2n+1}}{2n+1}. \end{aligned}$$

4 - La fonction argument cosinus hyperbolique n'étant pas définie sur un voisinage de l'origine (son ensemble de définition est $[1,+\infty[$), elle n'est pas développable en série entière dans un voisinage de l'origine.

Solution de l'exercice 5.9

1 - La fonction arc-sinus étant dérivable sur $]-1,1[$ de dérivée l'application $x \in]-1,1[\mapsto (1-x^2)^{-\frac{1}{2}}$, on en déduit que pour $x \in]-1,1[$ on a :

$$f'(x) = 2(1-x^2)^{-\frac{1}{2}} \arcsin(x),$$
$$\text{et} \qquad f''(x) = 2(1-x^2)\left(1 + x(1-x^2)^{-\frac{1}{2}} \arcsin(x)\right).$$

On peut remarquer que $f''(x) = (1-x^2)(xf'(x)+2)$, ce qui indique que f est solution de l'équation différentielle $(1-x^2)y''(x) - xy'(x) = 2$.

2 - Considérons l'équation différentielle normalisée associée à $(\mathcal{E})$ sur $]-1,1[$ qui est

$$(\mathcal{E}_N) \qquad y''(x) - \frac{x}{1-x^2}y'(x) = \frac{2}{1-x^2}.$$

Sur $]-1,1[$, les deux équations $(\mathcal{E})$ et $(\mathcal{E}_N)$ sont équivalentes. Les applications $\alpha_1 : x \in]-1,1[\mapsto \frac{x}{1-x^2}$ et $\beta : x \in]-1,1[\mapsto \frac{2}{1-x^2}$ sont développables en série entière dans un voisinage de l'origine puisque pour tout $x \in]-1,1[$, on a $\frac{1}{1-x^2} = \sum_{n=0}^{+\infty} x^{2n}$. On déduit de la proposition 5.11 que toutes les solutions de $(\mathcal{E}_N)$ (et par conséquent de $(\mathcal{E})$) sont développables en série entière dans un voisinage de l'origine.

Déterminons l'expression des solutions de $(\mathcal{E})$ sur $]-1,1[$ comme sommes de séries entières réelles. Considérons une solution y de $(\mathcal{E})$ sur $]-1,1[$ de la forme $y(x) = \sum_{n=0}^{+\infty} a_n x^n$. On a

$$y'(x) = \sum_{n=1}^{+\infty} n a_n x^{n-1} \quad \text{et} \quad y''(x) = \sum_{n=2}^{+\infty} n(n-1)a_n x^{n-2},$$

ce qui implique que :

$$(1-x^2)y''(x) - xy'(x) = 2a_2 + (6a_3 - a_1)x + \sum_{n=2}^{+\infty}\big((n+2)(n+1)a_{n+2} - n^2 a_n\big)x^n.$$

On en déduit que $2a_2 = 2$, que $6a_3 - a_1 = 0$ et que pour tout entier $n \geqslant 2$ on a $(n+2)(n+1)a_{n+2} - n^2 a_n = 0$. On a donc $a_2 = 1$, $a_3 = a_1$ et pour $n \geqslant 2$, $a_{n+2} = \frac{n^2}{(n+2)(n+1)} a_n$.

Comme $f(0) = 0$ et $f'(0) = 0$, d'après le théorème de Cauchy (théorème 5.10 p. 272) l'application f est l'unique solution de l'équation $(\mathcal{E})$ ou $(\mathcal{E}_N)$ vérifiant les conditions $y(0) = 0$ et $y'(0) = 0$. La condition $y(0) = 0$ implique que $a_0 = 0$ et la condition $y'(0) = 0$ implique que $a_1 = 0$. On déduit de la relation $a_{n+2} = \frac{n^2}{(n+2)(n+1)} a_n$ que pour tout entier n impair on a $a_n = 0$. Si n est pair avec $n = 2k$, $k \geqslant 2$, cette relation devient

$$a_{2k} = \frac{(2k-2)^2}{2k(2k-1)} a_{2k-2},$$

ce qui permet d'établir par récurrence que $a_{2k} = \frac{2^{2k-1}((k-1)!)^2}{(2k)!}$. On en déduit que l'unique série entière solution de l'équation $(\mathcal{E})$ pour les conditions initiales $y(0) = 0$ et $y'(0) = 0$ est $\sum_{k=1}^{+\infty} \frac{2^{2k-1}((k-1)!)^2}{(2k)!} x^{2k}$.

Le rayon de convergence de cette série entière lacunaire est $R = 1$ et pour tout $x \in]-1, 1[$, on a

$$\arcsin(x)^2 = \sum_{k=1}^{+\infty} \frac{2^{2k-1}((k-1)!)^2}{(2k)!} x^{2k}.$$

Solution de l'exercice 5.10

1 - Supposons que $(\mathcal{E})$ admette une solution f développable en série entière dans un voisinage de l'origine : il existe un voisinage V de l'origine tel que pour $x \in \mathbb{R}_+^* \cap V$ on ait $f(x) = \sum_{n=0}^{+\infty} a_n x^n$. On a alors

$$f'(x) = \sum_{n=1}^{+\infty} n a_n x^{n-1}, \quad f''(x) = \sum_{n=2}^{+\infty} n(n-1) a_n x^{n-2}$$

et en substituant ces expressions dans $(\mathcal{E})$, il vient

$$x^2 y''(x) + x y'(x) + x^2 y(x) = a_1 x + \sum_{n=2}^{+\infty} (n(n-1)a_n + n a_n + a_{n-2}) x^n = 0.$$

On en déduit que pour que y soit solution de $(\mathcal{E})$, on doit avoir

$$a_1 = 0 \quad \text{et} \quad n^2 a_n + a_{n-2} = 0 \text{ pour tout entier } n \geqslant 2.$$

La condition $a_1 = 0$ combinée à la relation $a_n = -\frac{1}{n^2} a_{n-2}$ permet d'établir par récurrence que $a_n = 0$ pour tout entier n impair. Par ailleurs, si n est un entier pair tel que $n = 2k$ avec $k \in \mathbb{N}^*$, on établit par récurrence que

$$a_n = a_{2k} = \frac{(-1)^k}{2^{2k}(k!)^2} a_0.$$

On en déduit que les solutions de $(\mathcal{E})$, développables en série entière dans un voisinage de l'origine, sont les fonctions f définies par

$$f(x) = \sum_{n=0}^{+\infty} a_n x^n = \sum_{k=0}^{+\infty} a_{2k} x^{2k} = \sum_{k=0}^{+\infty} \frac{(-1)^k a_0}{2^{2k}(k!)^2} x^{2k} = a_0 \sum_{k=0}^{+\infty} \frac{(-1)^k}{(k!)^2} \left(\frac{x}{2}\right)^{2k}.$$

Pour déterminer le rayon de convergence de la série entière lacunaire $\sum_n \frac{(-1)^n}{(n!)^2} \left(\frac{x}{2}\right)^{2n}$, utilisons le critère de D'Alembert pour les séries numériques. Pour $x \in \mathbb{R}$ et $k \in \mathbb{N}$, posons $u_k = \frac{(-1)^k}{(k!)^2} \left(\frac{x}{2}\right)^{2k}$. On a

$$\left|\frac{u_{k+1}}{u_k}\right| = \frac{(k!)^2}{((k+1)!)^2} \left(\frac{x}{2}\right)^2 = \left(\frac{x}{2(k+1)}\right)^2$$

et $\lim_{k\to+\infty} |u_{k+1}|/|u_k| = 0$. On en déduit que la série numérique $\sum_n \frac{(-1)^n}{(n!)^2}\left(\frac{x}{2}\right)^{2n}$ converge pour tout $x \in \mathbb{R}$. Les fonctions $a_0 J_0$ sont donc solutions sur $\mathbb{R}$ de l'équation de Bessel $(\mathcal{E})$.

2 - a) Pour montrer que $\alpha_n \underset{+\infty}{\sim} \ln(n)$, il faut montrer que $\lim_{n\to+\infty} \alpha_n/\ln(n) = 1$. Pour tout $n \in \mathbb{N}^*$, on a

$$\begin{aligned}\alpha_n &= \sum_{k=1}^{n} \frac{1}{k} = \sum_{k=1}^{n} \int_k^{k+1} \frac{1}{k}\,\mathrm{d}x \geqslant \sum_{k=1}^{n} \int_k^{k+1} \frac{1}{x}\,\mathrm{d}x \\ &= \sum_{k=1}^{n} \big(\ln(k+1) - \ln(k)\big) = \ln(n+1)\end{aligned}$$

et pour tout entier $n \geqslant 2$

$$\begin{aligned}\alpha_n &= \sum_{k=1}^{n} \frac{1}{k} = 1 + \sum_{k=2}^{n} \int_{k-1}^{k} \frac{1}{k}\,\mathrm{d}x \leqslant 1 + \sum_{k=2}^{n} \int_{k-1}^{k} \frac{1}{x}\,\mathrm{d}x \\ &= 1 + \sum_{k=2}^{n} \big(\ln(k-1) - \ln(k)\big) = 1 + \ln(n).\end{aligned}$$

On en déduit que pour tout entier $n \geqslant 2$, on a

$$\frac{\ln(n+1)}{\ln(n)} \leqslant \frac{\alpha_n}{\ln(n)} \leqslant 1 + \frac{1}{\ln(n)}$$

et on conclut par le théorème d'encadrement que $\lim_{n\to+\infty} \alpha_n/\ln(n) = 1$.

b) Pour déterminer le rayon de convergence de la série entière lacunaire $\sum_{n\geqslant 1} \frac{(-1)^n \alpha_n}{(n!)^2}\left(\frac{x}{2}\right)^{2n}$, utilisons le critère de D'Alembert pour les séries numériques. Pour $x \in \mathbb{R}$ et $n \in \mathbb{N}$, posons $u_n = \frac{(-1)^n \alpha_n}{(n!)^2}\left(\frac{x}{2}\right)^{2n}$. On a

$$\left|\frac{u_{n+1}}{u_n}\right| = \frac{\alpha_{n+1}}{\alpha_n}\left(\frac{x}{2(n+1)}\right)^2 \underset{+\infty}{\sim} \left(\frac{x}{2(n+1)}\right)^2$$

car d'après la question précédente $\alpha_{n+1} \underset{+\infty}{\sim} \ln(n+1) \underset{+\infty}{\sim} \ln(n)$ et $\alpha_n \underset{+\infty}{\sim} \ln(n)$. Pour tout $x \in \mathbb{R}$ on a $\lim_{n\to+\infty} |u_{n+1}|/|u_n| = 0$ donc la série numérique $\sum_{n\geqslant 1} \frac{(-1)^n \alpha_n}{(n!)^2}\left(\frac{x}{2}\right)^{2n}$ converge pour tout $x \in \mathbb{R}$.

c) Comme la série entière $\sum_{n\geqslant 1} \frac{(-1)^n \alpha_n}{(n!)^2}\left(\frac{x}{2}\right)^{2n}$ converge sur $\mathbb{R}$, elle a un rayon de convergence infini et elle est indéfiniment dérivable sur $\mathbb{R}$, voir le corollaire 5.3 p. 262. Pour tout $x \in \mathbb{R}$, on a

$$S'(x) = \sum_{n=1}^{+\infty} \frac{(-1)^n n \alpha_n}{(n!)^2}\left(\frac{x}{2}\right)^{2n-1}, \quad S''(x) = \sum_{n=1}^{+\infty} \frac{(-1)^n \alpha_n}{(n!)^2}\frac{(2n-1)n}{2}\left(\frac{x}{2}\right)^{2n-2}$$

On en déduit que

$$x^2\,S''(x) = \sum_{n=1}^{+\infty} \frac{(-1)^n 2n(2n-1)\alpha_n}{(n!)^2} \left(\frac{x}{2}\right)^{2n},$$

$$x\,S'(x) = \sum_{n=1}^{+\infty} \frac{(-1)^n 2n\alpha_n}{(n!)^2} \left(\frac{x}{2}\right)^{2n},$$

$$x^2\,S(x) = 4\sum_{n=1}^{+\infty} \frac{(-1)^n \alpha_n}{(n!)^2} \left(\frac{x}{2}\right)^{2n+2} = -4\sum_{n=2}^{+\infty} \frac{(-1)^n \alpha_{n-1}}{((n-1)!)^2} \left(\frac{x}{2}\right)^{2n}.$$

On a donc

$$\begin{aligned} x^2\,S''(x) + x\,S'(x) + x^2\,S(x) &= -4\left(\frac{x}{2}\right)^2 + 4\sum_{n=2}^{+\infty} \frac{(-1)^n}{((n-1)!)^2} (\alpha_n - \alpha_{n-1}) \left(\frac{x}{2}\right)^{2n} \\ &= -4\left(\frac{x}{2}\right)^2 + 4\sum_{n=2}^{+\infty} \frac{(-1)^n}{n!(n-1)!} \left(\frac{x}{2}\right)^{2n}. \end{aligned}$$

Par ailleurs,

$$\begin{aligned} 2xJ_0'(x) &= 2x\sum_{n=1}^{+\infty} \frac{(-1)^n}{n!(n-1)!} \left(\frac{x}{2}\right)^{2n-1} = 4\sum_{n=1}^{+\infty} \frac{(-1)^n}{n!(n-1)!} \left(\frac{x}{2}\right)^{2n} \\ &= -4\left(\frac{x}{2}\right)^2 + 4\sum_{n=2}^{+\infty} \frac{(-1)^n}{n!(n-1)!} \left(\frac{x}{2}\right)^{2n}. \end{aligned}$$

On a donc bien $x^2\,S''(x) + x\,S'(x) + x^2\,S(x) = 2xJ_0'(x)$.

d) Pour montrer que l'application $T_0 : x \in \mathbb{R}_+^* \mapsto \ln(x)J_0(x) - S(x)$ est une solution de $(\mathcal{E})$, il faut montrer que $x^2\,T''(x) + x\,T'(x) + x^2\,T(x) = 0$ pour tout $x \in \mathbb{R}_+^*$. D'après la question précédente, on a

$$\begin{aligned} x^2\,T''(x) + x\,T'(x) + x^2\,T(x) &= x^2\,(\ln(x)J_0(x))'' + x\,(\ln(x)J_0(x))' \\ &\quad + x^2\,\ln(x)J_0(x) - \underbrace{\left(x^2\,S''(x) + x\,S'(x) + x^2\,S(x)\right)}_{=2xJ_0'(x)}. \end{aligned}$$

Or, $(\ln(x)J_0(x))' = \frac{1}{x}J_0(x) + \ln(x)J_0'(x)$ et

$$(\ln(x)J_0(x))'' = -\frac{1}{x^2}J_0(x) + \frac{2}{x}J_0'(x) + \ln(x)J_0''(x)$$

d'où

$$\begin{aligned} &x^2\,(\ln(x)J_0(x))'' + x\,(\ln(x)J_0(x))' + x^2\,\ln(x)J_0(x) \\ &\quad = 2xJ_0'(x) + \ln(x)\underbrace{\left(x^2\,J_0''(x) + x\,J_0'(x) + x^2\,J_0(x)\right)}_{=0} = 2xJ_0'(x) \end{aligned}$$

car J_0 est solution de $(\mathcal{E})$.

e) Montrons que J_0 et T_0 sont deux solutions de $(\mathcal{E})$ linéairement indépendantes. Supposons[41] qu'il existe $(\lambda, \mu) \in \mathbb{R}^2$ tel que $\lambda J_0 + \mu T_0 = 0$ *i.e.* tel que pour tout $x \in \mathbb{R}_+^*$ on ait $\lambda J_0(x) + \mu T_0(x) = 0$ et montrons que nécessairement cela implique que $\lambda = \mu = 0$.

Remarquons que l'on a $\lim_{n\to 0+} J_0(x) = 1$ et $\lim_{n\to 0+} S(x) = 0$ ce qui implique que $\lim_{n\to 0+} T_0(x) = -\infty$. Raisonnons par l'absurde et supposons que $\lambda \neq 0$. Pour tout $x \in \mathbb{R}_+^*$ on a $J_0(x) = -\frac{\mu}{\lambda} T_0(x)$ et par conséquent

$$\frac{\mu}{\lambda} \lim_{n\to 0+} T_0(x) = -\lim_{n\to 0+} J_0(x) = -1.$$

Comme $\lim_{n\to 0+} T_0(x) = -\infty$, cette égalité est impossible et cela quelle que soit la valeur de μ. On a donc nécessairement $\lambda = 0$ et par conséquent $\mu T_0(x) = 0$ pour tout $x \in \mathbb{R}_+^*$. Comme $T_0(1) = S(1) \neq 0$, on a obligatoirement $\mu = 0$. On en conclut que J_0 et T_0 sont deux solutions de $(\mathcal{E})$ linéairement indépendantes.

3 - a) Intégrons par parties : pour tout $p \in \mathbb{N}$, on a

$$\begin{aligned} I_p &= \int_0^\pi \sin(\theta)\, \sin(\theta)^{2p-1}\, \mathrm{d}\theta \\ &= -\Big[\cos(\theta)\sin(\theta)^{2p-1}\Big]_0^\pi + (2p-1)\int_0^\pi \cos(\theta)^2\, \sin(\theta)^{2p-2}\, \mathrm{d}\theta \\ &= (2p-1)\int_0^\pi (1-\sin(\theta)^2)\, \sin(\theta)^{2p-2}\, \mathrm{d}\theta \;=\; (2p-1)(I_{p-1} - I_p). \end{aligned}$$

On en déduit que $I_p = \frac{2p-1}{2p} I_{p-1}$.

Montrons par récurrence que pour tout $p \in \mathbb{N}$ on a $I_p = \dfrac{(2p)!\,\pi}{2^{2p}(p!)^2}$.

- Pour $p = 0$ comme $0! = 1$ et $2^0 = 1$, on a $\frac{(2p)!\,\pi}{2^{2p}(p!)^2} = \pi$ et $I_0 = \int_0^\pi \mathrm{d}\theta = \pi$.
- Supposons, c'est notre hypothèse de récurrence, que la relation est vraie pour l'entier $p-1$, *i.e.* supposons que $I_{p-1} = \frac{(2p-2)!\,\pi}{2^{2p-2}((p-1)!)^2}$ et montrons que la relation varie pour l'entier p, *i.e.* montrons que $I_p = \frac{(2p)!\,\pi}{2^{2p}(p!)^2}$. On a

$$\begin{aligned} I_p &= \frac{2p-1}{2p} I_{p-1} = \frac{2p-1}{2p}\, \frac{(2p-2)!\,\pi}{2^{2p-2}((p-1)!)^2} = \frac{(2p-1)!\,\pi}{2^{2p-1}(p!(p-1)!)} \\ &= \frac{(2p)!\,\pi}{2p\, 2^{2p-1}(p!(p-1)!)} = \frac{(2p)!\,\pi}{2^{2p}(p!)^2}. \end{aligned}$$

La récurrence est achevée : pour tout $p \in \mathbb{N}$ on a $I_p = \frac{(2p)!\,\pi}{2^{2p}(p!)^2}$.

[41] Voir la définition 8.8 p. 317 du *Cours de première année*.

b) Montrons que la série de fonctions $\sum_n f_n$ de terme général $f_n : \theta \in \mathbb{R} \mapsto (-1)^n \frac{x_0^{2n}}{(2n)!} \sin^{2n}(\theta)$ où $x_0 \in \mathbb{R}_+^*$ est fixé, converge normalement sur $\mathbb{R}$. On a $\sup_{\theta\in\mathbb{R}} |f_n(\theta)| = \frac{x_0^{2n}}{(2n)!}$. Intéressons-nous à la nature de la série de terme général $u_n = \frac{x_0^{2n}}{(2n)!}$ pour $x_0 \in \mathbb{R}_+^*$ fixé. La quantité

$$\left|\frac{u_{n+1}}{u_n}\right| = \frac{x_0^2}{(2n+1)(2n+2)}$$

tendant vers 0 quand n tend vers $+\infty$, le critère de D'Alembert indique que la série $\sum_n u_n$ converge. On en déduit que la série de fonctions $\sum_n f_n$ converge normalement et par conséquent uniformément sur $\mathbb{R}$.

Calculons la somme de la série de fonctions $\sum_n f_n$. Pour tout $x \in \mathbb{R}$, on a

$$\sum_{n=0}^{+\infty} f_n(x) = \sum_{n=0}^{+\infty} (-1)^n \frac{x_0^{2n}}{(2n)!} \sin^{2n}(\theta) = \sum_{n=0}^{+\infty} (-1)^n \frac{(x_0 \sin(\theta))^{2n}}{(2n)!}.$$

Comme $\cos(u) = \sum_{n=0}^{+\infty} (-1)^n \frac{u^{2n}}{(2n)!}$, on en déduit que $\sum_{n=0}^{+\infty} f_n(x) = \cos(x_0 \sin(\theta))$.

c) D'après le théorème 3.8 p. 136, la convergence uniforme sur $\mathbb{R}$ et donc sur $[0, \pi]$ de la série de fonctions $\sum_n f_n$ permet d'écrire pour tout $x_0 \in \mathbb{R}_+^*$ fixé :

$$\int_0^\pi \cos(x \sin(\theta))\, \mathrm{d}\theta = \int_0^\pi \sum_{n=0}^{+\infty} f_n(\theta)\, \mathrm{d}\theta = \sum_{n=0}^{+\infty} \int_0^\pi f_n(\theta)\, \mathrm{d}\theta.$$

Or,

$$\begin{aligned}\int_0^\pi f_n(\theta)\, \mathrm{d}\theta &= (-1)^n \frac{x_0^{2n}}{(2n)!} \int_0^\pi \sin^{2n}(\theta)\, \mathrm{d}\theta = (-1)^n \frac{x_0^{2n}}{(2n)!} I_{2n} \\ &= (-1)^n \frac{x_0^{2n}}{(2n)!} \frac{(2n)!\,\pi}{2^{2n}(n!)^2}.\end{aligned}$$

On en déduit que pour tout $x_0 \in \mathbb{R}_+^*$, on a

$$\int_0^\pi \cos(x \sin(\theta))\, \mathrm{d}\theta = \pi \sum_{n=0}^{+\infty} (-1)^n \frac{x_0^{2n}}{(2n)!} \frac{(2n)!}{2^{2n}(n!)^2},$$

autrement dit, que $J_0(x_0) = \frac{1}{\pi} \int_0^\pi \cos(x_0 \sin(\theta))\, \mathrm{d}\theta$.

Solution de l'exercice 5.11

1 - a) Pour calculer le rayon de convergence de la série entière $\sum_{n\geqslant 1} n^\pi z^n$ on peut utiliser le critère de D'Alembert. Pour tout $n \in \mathbb{N}^*$ on a

$$\left|\frac{(n+1)^\pi}{n^\pi}\right| = (1 + 1/n)^\pi = \exp\left(\pi \ln(1 + 1/n)\right).$$

Cette quantité tend vers 1 quand n tend vers $+\infty$ ce qui indique que le rayon de convergence de la série $\sum_n n^\pi z^n$ est 1.

b) On a $S(-0.99) = \sum_{n=1}^{+\infty}(-1)^n n^\pi (0.99)^n$. Pour tout $n \in \mathbb{N}^*$, posons $\alpha_n = n^\pi (0.999)^n = \exp(\pi \ln(n) + n\ln(0.99))$. La suite $(\alpha_n)_n$ est clairement une suite à termes positifs. De plus, la suite $(\alpha_n)_n$ converge vers 0 car $\pi \ln(n) = o_{+\infty}(n\ln(0.99))$ et $\lim_{n\to+\infty} n\ln(0.99) = -\infty$ puisque $\ln(0.99) < 0$. Pour pouvoir appliquer le critère spécial pour les séries alternées à la série $\sum_n (-1)^n n^\pi (0.999)^n$, il reste à vérifier que la suite $(\alpha_n)_n$ est décroissante.

Pour tout $n \in \mathbb{N}^*$ on a

$$\beta_n = \frac{\alpha_{n+1}}{\alpha_n} = 0.99\,(1+1/n)^\pi = 0.99\exp\left(\pi \ln\left(1+1/n\right)\right).$$

Il est difficile d'établir que $\beta_n \leqslant 1$ directement. Toutefois, la suite $(\beta_n)_n$ admet pour limite 1. Elle est strictement croissante car les fonctions exponentielle et logarithme étant strictement croissantes, pour tout $(n,p) \in \mathbb{N}^* \times \mathbb{N}^*$ on a

$$\begin{aligned} n < p &\iff \ln(1+1/n) < \ln(1+1/p) \\ &\iff \exp(\pi\ln(1+1/n)) < \exp(\pi\ln(1+1/p)) \iff \beta_n < \beta_p \end{aligned}$$

Ceci implique que pour tout $n \in \mathbb{N}^*$ on a $\beta_n = \frac{\alpha_{n+1}}{\alpha_n} < 1$. On en déduit que la suite $(\alpha_n)_n$ est décroissante.

On déduit du critère spécial pour les séries alternées[42] que la série numérique $\sum_n (-1)^n n^\pi (0.99)^n$ est une série alternée convergente. Par conséquent, son reste d'ordre N est majoré par le premier terme négligé[43], *i.e.*

$$|S(-0.99) - S_N(-0.99)| = \left|\sum_{n=N+1}^{+\infty}(-1)^n\alpha_n\right| \leqslant \alpha_{N+1} = (N+1)^\pi (0.99)^{N+1}.$$

c) On peut obtenir une valeur approchée à 10^{-6} près de la valeur de $S(-0.99)$ avec MAPLE de la manière suivante :

```
> z0:=-0.99:
> alpha:=(n,z)->n**Pi*abs(z)**n:
> N:=solve(alpha(n+1,z0)=1E-6,n);
                N := -0.9876928418, 3963.443118
> SP:= (N,z)-> sum(evalf(n^Pi*z**n),n=0..N):
> SP(ceil(N[2]),z0);
                          0.21856
```

Avec la valeur $z = -0.99$, le rang à partir duquel la somme partielle $S_N(z)$ constitue une valeur approchée de $S(z)$ est 3964. Avec la valeur $z = -0.999$, le rang à partir duquel la somme partielle $S_N(z)$ constitue une valeur approchée de $S(z)$ est 47630 comme l'indique le calcul suivant :

---

[42] Voir le théorème 1.3 p. 42.

[43] Voir la proposition 1.15 p. 44.

```
> z0:=-0.999:
> N:=solve(alpha(n+1,z0)=1E-6,n);
                    N := -0.9876932781, 47629.49147
```

Il est peu raisonnable de devoir sommer environ cinquante mille termes pour obtenir une valeur approchée de $S(-0.999)$ avec une erreur de l'ordre de 10^{-6}. La transformée d'Euler-Abel, comme nous allons le voir, va permettre d'accélérer la convergence de la série vers sa somme.

2 - a) Comme pour tout $z \in D(0,1)$ on a $f(z) = \sum_{n=0}^{+\infty} a_n z^n$, la fonction g est définie sur $D(0,1) \setminus \{0\}$ et pour tout $z \in D(0,1) \setminus \{0\}$, on a :

$$g(z) = \frac{1}{z} \sum_{n=1}^{+\infty} a_n z^n = \sum_{n=1}^{+\infty} a_n z^{n-1} = \sum_{k=0}^{+\infty} a_{k+1} z^k = a_1 + \sum_{k=1}^{+\infty} a_{k+1} z^k.$$

On en déduit que $\lim_{z \to 0} g(z) = a_1$ et par conséquent que g est prolongeable par continuité en 0 en posant $g(0) = a_1$.

b) D'après le calcul précédent, il vient pour tout $z \in D(0,1)$

$$\begin{aligned}
(1-z)g(z) &= (1-z) \sum_{n=0}^{+\infty} a_{n+1} z^n = \sum_{n=0}^{+\infty} a_{n+1} z^n - \sum_{n=0}^{+\infty} a_{n+1} z^{n+1} \\
&= a_1 + \sum_{n=1}^{+\infty} (a_{n+1} - a_n) z^n = a_1 + \sum_{n=1}^{+\infty} \Delta^{[1]} a_n z^n \\
&= a_0 + \underbrace{(a_1 - a_0)}_{\Delta^{[1]} a_0} + \sum_{n=1}^{+\infty} \Delta^{[1]} a_n z^n = a_0 + \sum_{n=0}^{+\infty} \Delta^{[1]} a_n z^n.
\end{aligned}$$

c) Puisque pour tout $z \in D(0,1) \setminus \{0\}$ on a $g(z) = \frac{1}{z}(f(z) - a_0)$, on en déduit que :

$$\begin{aligned}
f(z) &= a_0 + z g(z) = a_0 + \frac{z}{1-z} (1-z) g(z) \\
&= a_0 + \frac{z}{1-z} \left(a_0 + \sum_{n=0}^{+\infty} \Delta^{[1]} a_n z^n \right) = \frac{a_0}{1-z} + \frac{z}{1-z} \sum_{n=0}^{+\infty} \Delta^{[1]} a_n z^n.
\end{aligned}$$

Pour $z = 0$, la relation est encore vraie car $f(0) = 0$ et pour $z = 0$ on a $\frac{a_0}{1-z} + \frac{z}{1-z} \sum_{n=0}^{+\infty} \Delta^{[1]} a_n z^n = a_0$.

d) Utilisons un raisonnement par récurrence pour établir la formule d'Euler-Abel à l'ordre p. Elle s'écrit :

$$(\mathcal{E}_p) \quad f(z) = \sum_{k=0}^{p-1} \Delta^{[k]} a_0 \frac{z^k}{(1-z)^{k+1}} + \frac{z^p}{(1-z)^p} \sum_{n=0}^{+\infty} \Delta^{[p]} a_n z^n \quad \forall z \in D(0,1).$$

Pour $p = 1$ la relation $(\mathcal{E}_1)$ est vraie, elle a été établie à la question précédente. Supposons pour un entier $p \geqslant 1$ fixé la relation $(\mathcal{E}_p)$ vraie et montrons que la relation $(\mathcal{E}_{p+1})$ est vraie.

Comme $\Delta^{[p+1]}a_n = \Delta^{[p]}a_{n+1} - \Delta^{[p]}a_n$ on a pour tout $z \in D(0,1)$,

$$\sum_{k=0}^{p} \Delta^{[k]}a_0 \frac{z^k}{(1-z)^{k+1}} + \frac{z^{p+1}}{(1-z)^{p+1}} \sum_{n=0}^{+\infty} \Delta^{[p+1]}a_n z^n$$
$$= \sum_{k=0}^{p-1} \Delta^{[k]}a_0 \frac{z^k}{(1-z)^{k+1}} + \frac{z^{p+1}}{(1-z)^{p+1}} \left(\frac{\Delta^{[p]}a_0}{z} \right. \tag{16}$$
$$\left. + \sum_{n=0}^{+\infty} \Delta^{[p]}a_{n+1} z^n - \sum_{n=0}^{+\infty} \Delta^{[p]}a_n z^n \right).$$

De plus, on a :

$$\sum_{n=0}^{+\infty} \Delta^{[p]}a_{n+1} z^n = \frac{1}{z} \sum_{n=0}^{+\infty} \Delta^{[p]}a_{n+1} z^{n+1} = \frac{1}{z} \sum_{n=1}^{+\infty} \Delta^{[p]}a_n z^n.$$

On en déduit que :

$$\frac{\Delta^{[p]}a_0}{z} + \sum_{n=0}^{+\infty} \Delta^{[p]}a_{n+1} z^n - \sum_{n=0}^{+\infty} \Delta^{[p]}a_n z^n \tag{17}$$
$$= \frac{\Delta^{[p]}a_0}{z} + \Delta^{[p]}a_0 + \left(\frac{1}{z} - 1\right) \sum_{n=1}^{+\infty} \Delta^{[p]}a_n z^n = \left(\frac{1}{z} - 1\right) \sum_{n=0}^{+\infty} \Delta^{[p]}a_n z^n.$$

Or, $\frac{z^{p+1}}{(1-z)^{p+1}} \left(\frac{1}{z} - 1\right) = \frac{z^p}{(1-z)^p}$ donc (16) combiné avec (17) donne :

$$\sum_{k=0}^{p} \Delta^{[k]}a_0 \frac{z^k}{(1-z)^{k+1}} + \frac{z^{p+1}}{(1-z)^{p+1}} \sum_{n=0}^{+\infty} \Delta^{[p+1]}a_n z^n$$
$$= \sum_{k=0}^{p-1} \Delta^{[k]}a_0 \frac{z^k}{(1-z)^{k+1}} + \frac{z^p}{(1-z)^p} \sum_{n=0}^{+\infty} \Delta^{[p]}a_n z^n = f(z).$$

On en conclut que la relation $(\mathcal{E}_p)$ est vraie pour tout $p \in \mathbb{N}^*$.

3 - a) On a $(n+1)^\alpha = n^\alpha (1 + 1/n)^\alpha$ et $(1+u)^\alpha = 1 + \alpha u + \frac{1}{2}\alpha(\alpha-1)u^2 + o_0(u^2)$.
On en déduit que :

$$(n+1)^\alpha = n^\alpha \left(1 + \frac{\alpha}{n} + \alpha(\alpha-1)\frac{1}{2n^2} + o_{+\infty}\left(n^{-2}\right)\right)$$
$$= n^\alpha + \alpha n^{\alpha-1} + \tfrac{1}{2}\alpha(\alpha-1) n^{\alpha-2} + o_{+\infty}(n^{\alpha-2})$$

b) On a,

$$\Delta^{[1]}a_n = a_{n+1} - a_n = n^{\alpha+1} - n^\alpha$$
$$= \alpha n^{\alpha-1} + \tfrac{1}{2}\alpha(\alpha-1) n^{\alpha-2} + o_{+\infty}(n^{\alpha-2}) \underset{+\infty}{\sim} \alpha n^{\alpha-1}.$$

Par ailleurs,

$$\begin{aligned}\Delta^{[2]}a_n &= \Delta^{[1]}a_{n+1} - \Delta^{[1]}a_n \\ &= \alpha\,(n+1)^{\alpha-1} + \tfrac{1}{2}\alpha(\alpha-1)\,(n+1)^{\alpha-2} - \alpha\, n^{\alpha-1} \\ &\quad -\tfrac{1}{2}\alpha(\alpha-1)\,n^{\alpha-2} + o_{+\infty}(n^{\alpha-2})\end{aligned}$$

car $(n+1)^{\alpha-2} \underset{+\infty}{\sim} n^{\alpha-2}$.
Or, d'après la question 3-a $(n+1)^{\alpha-1} = n^{\alpha-1} + (\alpha-1)\,n^{\alpha-2} + o_{+\infty}(n^{\alpha-2})$ et $(n+1)^{\alpha-2} = n^{\alpha-2} + o_{+\infty}(n^{\alpha-2})$. On en déduit finalement que

$$\begin{aligned}\Delta^{[2]}a_n &= \alpha n^{\alpha-1} + \alpha(\alpha-1)\,n^{\alpha-2} + \tfrac{1}{2}\alpha(\alpha-1)\,n^{\alpha-2} \\ &\qquad -\alpha n^{\alpha-1} - \alpha(\alpha-1)\,n^{\alpha-2} + o_{+\infty}(n^{\alpha-2}) \\ &= \alpha(\alpha-1)\,n^{\alpha-2} + o_{+\infty}(n^{\alpha-2}) \underset{+\infty}{\sim} \alpha(\alpha-1)\,n^{\alpha-2}.\end{aligned}$$

c) On a vu que le calcul d'une approximation de la somme de la série $\sum_n n^\pi z^n$ pour $z = -0.999$ à partir des sommes partielles requérait le calcul d'un très grand nombre de termes car $n^\pi(-0.999)^n$ est une quantité qui tend très lentement vers 0. En utilisant la formule d'Euler-Abel, on améliore nettement la vitesse de convergence et le calcul d'une approximation de la somme de la série requiert beaucoup moins de termes à calculer. En effet, alors que $a_n \underset{+\infty}{\sim} n^\alpha$ on a $\Delta^{[1]}a_n \underset{+\infty}{\sim} \alpha\, n^{\alpha-1}$ et $\Delta^{[2]}a_n \underset{+\infty}{\sim} \alpha(\alpha-1)\,n^{\alpha-2}$.

4 - a) Pour calculer le rayon de convergence de la série entière $\sum_n n\,z^n$, on peut utiliser le critère de D'Alembert. On a $\lim_{n\to+\infty}\left|\frac{n+1}{n}\right| = 1$ ce qui indique que le rayon de convergence de la série $\sum_n n\,z^n$ est 1.

Une série entière étant holomorphe sur son disque de convergence et la série entière $\sum_n z^n$ ayant pour rayon de convergence 1, on déduit du théorème 5.4 p. 249 que pour tout $z \in D(0,1)$ on a

$$\sum_{n=0}^{+\infty} nz^n = \sum_{n=1}^{+\infty} nz^n = z\sum_{n=1}^{+\infty} nz^{n-1} = z\sum_{n=1}^{+\infty}(z^n)' = z\left(\sum_{n=1}^{+\infty} z^n\right)' = z\left(\sum_{n=0}^{+\infty} z^n\right)'.$$

Comme $\displaystyle\sum_{n=0}^{+\infty} z^n = \frac{1}{1-z}$, on en déduit que $\displaystyle\sum_{n=0}^{+\infty} nz^n = z\left(\frac{1}{1-z}\right)' = \frac{z}{(1-z)^2}$.

b) Le rayon de convergence de la série entière $\sum_n n^2 z^n$ est 1. Utilisons la formule d'Euler-Abel à l'ordre 1 pour calculer la somme f de cette série. Pour tout $n \in \mathbb{N}$, on a $\Delta^{[1]}a_n = (n+1)^2 - n^2 = 2n+1$. On en déduit que pour tout $z \in D(0,1)$,

$$f(z) = \sum_{n=0}^{+\infty} \Delta^{[1]}a_n z^n = 2\sum_{n=0}^{+\infty} nz^n + \sum_{n=0}^{+\infty} z^n = \frac{2z}{(1-z)^2} + \frac{1}{1-z} = \frac{z(z+1)}{(1-z)^3}.$$

c) Soit $P = \sum_{k=0}^{p-1} \alpha_k X^k$ un polynôme de $\mathbb{C}[X]$ de degré $p-1$ ($p \in \mathbb{N}^*$). Pour déterminer le rayon de convergence R de la série entière complexe $\sum_n P(n)\, z^n$, utilisons le critère de D'Alembert. Comme P est un polynôme de $\mathbb{C}[X]$ de degré $p-1$, il admet exactement $p-1$ racines complexes comptées avec multiplicité et au plus $p-1$ racines réelles[44]. Pour tout entier n supérieur à la valeur absolue de la racine réelle de P ayant la plus grande valeur absolue, on a $P(n) \neq 0$ et

$$\left|\frac{P(n+1)}{P(n)}\right| = \left|\frac{\sum_{k=0}^{p-1} \alpha_k (n+1)^k}{\sum_{k=0}^{p-1} \alpha_k n^k}\right|.$$

Or,

$$\sum_{k=0}^{p-1} \alpha_k (n+1)^k \underset{n\to+\infty}{\sim} \alpha_{p-1} n^{p-1} \quad \text{et} \quad \sum_{k=0}^{p-1} \alpha_k n^k \underset{n\to+\infty}{\sim} \alpha_{p-1} n^{p-1}$$

donc $\lim_{n\to+\infty} \left|\frac{P(n+1)}{P(n)}\right| = 1$. On en conclut que le rayon de convergence R de la série entière complexe $\sum_n P(n)\, z^n$ vaut 1.

d) Montrons en utilisant un raisonnement par récurrence que pour tout polynôme $P = \sum_{k=0}^{p-1} \alpha_k X^k$ de $\mathbb{C}[X]$ de degré $p-1$ ($p \in \mathbb{N}^*$) on a $\Delta^{[p]} P(n) = 0$ pour tout $n \in \mathbb{N}$.

Pour $p = 1$ la relation est vraie : on a $P = \alpha_0$ donc pour tout $z \in \mathbb{C}$, on a $\Delta^{[1]} P(n) = P(n+1) - P(n) = \alpha_0 - \alpha_0 = 0$. Supposons, c'est notre hypothèse de récurrence, que pour tout polynôme P de $\mathbb{C}[X]$ de degré $p-1$ fixé ($p \in \mathbb{N}^*$) on a $\Delta^{[p]} P(n) = 0$ pour tout $n \in \mathbb{N}$ et montrons que pour tout polynôme $Q = \sum_{k=0}^{p} \alpha_k X^k$ de $\mathbb{C}[X]$ de degré p on a $\Delta^{[p+1]} Q(n) = 0$ pour tout $n \in \mathbb{N}$. On a $Q = Q_1 + \alpha_p X^p$ où Q_1 est le polynôme de $\mathbb{C}[X]$ de degré $p-1$ donné par $Q_1 = \sum_{k=0}^{p-1} \alpha_k X^k$. On en déduit que pour tout $m \in \mathbb{N}$, on a

$$\begin{aligned}
\Delta^{[p+1]} Q(m) &= \Delta^{[p]} Q(m+1) - \Delta^{[p]} Q(m) \\
&= \Delta^{[p]} \left(Q_1(m+1) + \alpha_p (m+1)^p\right) - \Delta^{[p]} \left(Q_1(m) + \alpha_p m^p\right) \\
&= \Delta^{[p]} Q_1(m+1) - \Delta^{[p]} Q_1(m) + \alpha_p \Delta^{[p]} (m+1)^p - \alpha_p \Delta^{[p]} m^p.
\end{aligned}$$

Comme Q_1 est de degré $p-1$, d'après l'hypothèse de récurrence $\Delta^{[p]} Q_1(m) = 0$ et $\Delta^{[p]} Q_1(m+1) = 0$. On en déduit que

$$\Delta^{[p+1]} Q(m) = \alpha_p \Delta^{[p]} (m+1)^p - \alpha_p \Delta^{[p]} m^p = \alpha_p \Delta^{[p]} \left((m+1)^p - m^p\right).$$

Or, la formule du binôme de Newton donne

$$(m+1)^p - m^p = \sum_{k=0}^{p} C_p^k m^k - m^p = \sum_{k=0}^{p-1} C_p^k m^k = Q_2(m)$$

[44] Voir le *Cours de première année* p. 248.

où $Q_2 = \sum_{k=0}^{p-1} C_p^k X^k$ est un polynôme de $\mathbb{R}[X]$ de degré $p-1$. On a donc d'après l'hypothèse de récurrence $\Delta^{[p]} Q_2(m) = 0$ et on en conclut que $\Delta^{[p+1]} Q(m) = 0$.

On a donc établi que pour tout polynôme P de $\mathbb{C}[X]$ de degré $p-1$ ($p \in \mathbb{N}^*$) on a $\Delta^{[p]} P(n) = 0$ pour tout $n \in \mathbb{N}$.

e) Pour $p \in \mathbb{N}^*$ fixé, en appliquant la formule d'Euler-Abel à l'ordre p à la série entière $\sum_n P(n) z^n$ où P est un polynôme de $\mathbb{C}[X]$ de degré $p-1$, on obtient pour tout $z \in D(0,1)$

$$\sum_{n=0}^{+\infty} P(n)\, z^n = \sum_{k=0}^{p-1} \Delta^{[k]} P(0)\, \frac{z^k}{(1-z)^{k+1}} + \frac{z^p}{(1-z)^p} \sum_{n=0}^{+\infty} \Delta^{[p]} P(n)\, z^n.$$

D'après la question précédente, comme P est un polynôme de $\mathbb{C}[X]$ de degré $p-1$ on a $\Delta^{[p]} P(n) = 0$ pour tout $n \in \mathbb{N}$. On en déduit que pour tout $z \in D(0,1)$ on a

$$\sum_{n=0}^{+\infty} P(n)\, z^n = \sum_{k=0}^{p-1} \Delta^{[k]} P(0)\, \frac{z^k}{(1-z)^{k+1}}. \tag{18}$$

f) Pour implémenter sous MAPLE un algorithme mettant en œuvre la formule d'Euler-Abel pour calculer la somme d'une série entière de la forme $\sum_n P(n) z^n$ où P désigne un polynôme de $\mathbb{C}[X]$, il convient dans un premier temps d'expliciter la relation permettant de calculer $\Delta^{[k]} P(0)$. On montre par récurrence que pour tout $k \in \mathbb{N}$,

$$\Delta^{[k]} P(0) = \sum_{i=0}^{k} (-1)^i C_k^i\, \Delta^{[0]} P(k-i) = \sum_{i=0}^{k} (-1)^i C_k^i\, P(k-i).$$

La programmation avec MAPLE de la formule d'Euler-Abel pour calculer la somme de la série $\sum_n P(n) z^n$ est alors la suivante.

```
> P:=z-> z^6+1: p:=degree(P(z),z)+1:
> Delta:=k-> add((-1)^i*binomial(k,i)*P(k-i),i=0..k):
> f:=z->simplify(add(Delta(k)*z^k/(1-z)^(k+1),k=0..p-1)): f(z);
```

$$-\frac{1 + 2z^6 - 5z + 72z^2 + 282z^3 + 317z^4 + 51z^5}{(-1+z)^7}$$

Bien entendu cette formule d'Euler-Abel est connue des concepteurs de MAPLE et on obtient le même résultat directement avec la commande `sum` :

```
> sum(P(n)*z^n,n=0..infinity);
```

$$-\frac{1 + 2z^6 - 5z + 72z^2 + 282z^3 + 317z^4 + 51z^5}{(-1+z)^7}$$

On sera attentif au fait que la relation (18) n'a de sens que pour $z \in D(0,1)$ même si l'expression $\sum_{k=0}^{p-1} \Delta^{[k]} P(0)\, \frac{z^k}{(1-z)^{k+1}}$ est définie pour tout $z \neq 1$.

CHAPITRE 6

Séries de Fourier

En analyse, les séries de Fourier sont un outil essentiel pour l'étude des fonctions périodiques. Elles sont également largement utilisées dans de nombreux domaines pour l'étude de signaux périodiques et interviennent par exemple en filtrage numérique, en synthèse sonore ou en traitement d'images. On fait habituellement remonter l'origine des séries de Fourier à l'ouvrage de Joseph Fourier *Théorie analytique de la chaleur* paru en 1822 où est exposée l'utilisation formelle de sommes trigonométriques infinies pour la résolution de l'équation de la chaleur. À cette époque toutefois, les notions d'intégration ou de continuité ainsi que les problèmes de convergence des séries de Fourier ne sont pas clairement maîtrisés et les séries de Fourier vont se trouver au cœur du développement de l'analyse au 19^{e} siècle.

6.1 Séries trigonométriques

6.1.1 Convergence d'une série trigonométrique

DÉFINITION 6.1 *On appelle* série trigonométrique réelle *toute série d'applications de terme général*

$$f_n : x \in \mathbb{R} \longmapsto a_n \cos(nx) + b_n \sin(nx)$$

où $(a_n)_n$ et $(b_n)_n$ désignent deux suites numériques réelles.

Une application de la forme $x \in \mathbb{R} \mapsto \sum_{k=0}^{n} a_k \cos(kx) + b_k \sin(kx)$, où $(a_k)_{k\in\{0,\dots,n\}}$ et $(b_k)_{k\in\{0,\dots,n\}}$ désignent des réels, porte le nom de *polynôme trigonométrique.* Les sommes partielles associées à une série trigonométrique sont donc des polynômes trigonométriques.

PROPOSITION 6.1 *Si les séries numériques $\sum_n a_n$ et $\sum_n b_n$ convergent absolument alors la série trigonométrique dont le terme général est $f_n : x \mapsto a_n \cos(nx) + b_n \sin(nx)$ converge normalement (et par conséquent simplement, absolument et uniformément) sur $\mathbb{R}$.*

Démonstration Les fonctions sinus et cosinus étant majorées par 1 et minorées par -1, pour tout $n \in \mathbb{N}$ et tout $x \in \mathbb{R}$ on a

$$|a_n \cos(nx) + b_n \sin(nx)| \leqslant |a_n|\,|\cos(nx)| + |b_n|\,|\sin(nx)| \leqslant |a_n| + |b_n|.$$

On en déduit que $\sup_{x\in\mathbb{R}} |a_n \cos(nx) + b_n \sin(nx)| \leqslant |a_n| + |b_n|$. Si les séries numériques $\sum_n a_n$ et $\sum_n b_n$ convergent absolument alors la série de terme général $\sum_n |a_n| + |b_n|$ converge. D'après les critères de convergence pour les séries à termes positifs[1], la série de terme général $\sup_{x\in\mathbb{R}} |a_n \cos(nx) + b_n \sin(nx)|$ converge. Par définition, cela signifie que la série $\sum_n a_n \cos(nx) + b_n \sin(nx)$ converge normalement sur $\mathbb{R}$. □

Proposition 6.2 (Règle d'Abel pour les séries trigonométriques) *Soient $(a_n)_n$ et $(b_n)_n$ deux suites de nombres réels positifs, décroissantes et convergeant vers 0. La série trigonométrique de terme général $f_n : x \mapsto a_n \cos(nx) + b_n \sin(nx)$,*

- *converge simplement sur $\mathbb{R} \setminus 2\pi\mathbb{Z}$;*
- *converge uniformément sur tout intervalle $[2k\pi+\alpha, 2(k+1)\pi-\alpha]$ où $k \in \mathbb{Z}$ et $\alpha \in]0, \pi[$.*

Démonstration Ce résultat se démontre en appliquant le théorème d'Abel pour les séries numériques (théorème 1.4 p. 46) et le théorème d'Abel uniforme pour les séries de fonctions (théorème 3.6 p. 133) à la série trigonométrique $\sum_n a_n \cos(nx) + b_n \sin(nx)$. Le détail de la justification se situe à la p. 48 pour la première partie du théorème et à la p. 133 pour la seconde. □

6.1.2 Propriétés de la somme d'une série trigonométrique

On dit qu'une application définie sur un sous-ensemble D de $\mathbb{R}$ est 2π-périodique si :

$$\forall x \in D \quad \big(x + 2\pi \in D \quad \text{et} \quad f(x + 2\pi) = f(x)\big).$$

La période d'une telle application peut être inférieure à 2π ; c'est le cas par exemple de l'application $x \in \mathbb{R} \mapsto \cos(2x)$ qui est 2π-périodique et qui admet pour *période fondamentale* π.

Proposition 6.3 *La somme d'une série trigonométrique est une application 2π-périodique.*

Démonstration Soit $\sum_n a_n \cos(nx) + b_n \sin(nx)$ une série trigonométrique qui converge en $x_0 \in \mathbb{R}$. Pour tout $n \in \mathbb{N}^*$, on a

$$\cos(n(x_0 + 2\pi)) = \cos(nx_0 + 2\pi n) = \cos(nx_0)$$
$$\text{et} \qquad \sin(n(x_0 + 2\pi)) = \sin(nx_0 + 2\pi n) = \sin(nx_0).$$

[1] Voir le théorème 1.2 p. 14.

Ces relations sont encore valables si $n = 0$. On en déduit que les deux séries $\sum_n a_n \cos(nx_0) + b_n \sin(nx_0)$ et $\sum_n a_n \cos(n(x_0 + 2\pi)) + b_n \sin(n(x_0 + 2\pi))$ sont égales. Elles ont donc même somme. □

Une conséquence de la proposition 6.3 est qu'en pratique on peut se contenter d'étudier une série trigonométrique sur un intervalle de période, par exemple sur $[0, 2\pi]$.

PROPOSITION 6.4 *Si une série trigonométrique converge uniformément sur un intervalle ouvert I alors sa somme est une application continue sur I.*

Démonstration L'application $f_n : x \in \mathbb{R} \mapsto a_n \cos(nx) + b_n \sin(nx)$ est continue quel que soit $n \in \mathbb{N}$. S'il y a convergence uniforme de la série $\sum_n f_n$ sur un intervalle I de $\mathbb{R}$, la somme de cette série est, d'après le théorème 3.7 p. 135, continue sur I. □

PROPOSITION 6.5 *Soit $\sum_n f_n$ une série trigonométrique qui converge sur un intervalle ouvert I de $\mathbb{R}$ et soit ϕ sa somme. Si la série trigonométrique $\sum_n f'_n$ converge uniformément sur I alors ϕ est dérivable sur I et ϕ' est la somme de la série $\sum_n f'_n$*

Démonstration Soit $\sum_n f_n$ une série trigonométrique de terme général $f_n : x \in \mathbb{R} \mapsto a_n \cos(nx) + b_n \sin(nx)$. Pour tout $n \in \mathbb{N}$, l'application f_n est dérivable sur $\mathbb{R}$ et la série $\sum_n f'_n$ est une série trigonométrique. Si $\sum_n f_n$ converge simplement sur un intervalle I et que la série $\sum_n f'_n$ converge uniformément sur I alors il résulte du théorème 3.9 p. 139 que l'application ϕ, somme de la série $\sum_n f_n$, est dérivable sur I et que $\phi' = \psi$ où ψ est la somme de la série $\sum_n f'_n$. □

Dans l'énoncé des propositions 6.4 et 6.5, l'hypothèse de convergence uniforme de la série est une condition suffisante pour que la propriété indiquée soit vraie. Cette condition n'est toutefois pas nécessaire mais dans les cas où la convergence uniforme n'a pas lieu une étude spécifique doit être menée pour établir si la propriété est vraie ou pas.

Liens entre somme d'une série trigonométrique et coefficients de la série

Soit $\sum_n f_n$ une série trigonométrique de terme général $f_n : x \mapsto a_n \cos(nx) + b_n \sin(nx)$, convergeant uniformément sur l'intervalle $[0, 2\pi]$. D'après les propositions 6.3 et 6.4, la somme ϕ de cette série trigonométrique est une application continue sur $\mathbb{R}$ et 2π-périodique. Nous allons montrer que les coefficients a_n et b_n de la série trigonométrique $\sum_n f_n$ sont liés à ϕ de la série par des relations très simples.

Pour tout $p \in \mathbb{N}$, la série de terme général $g_n : x \in \mathbb{R} \mapsto f_n(x)\cos(px)$ converge uniformément sur $\mathbb{R}$ et a pour somme l'application $\psi : x \in \mathbb{R} \mapsto \phi(x)\cos(px)$. En effet pour tout $n \in \mathbb{N}$, on a

$$\begin{aligned}\sup_{x\in\mathbb{R}}\left|\psi(x) - \sum_{k=0}^{n} g_k(x)\right| &= \sup_{x\in[0,2\pi]}\left|\underbrace{\cos(px)}_{\leqslant 1}\left(\phi(x) - \sum_{k=0}^{n} f_k(x)\right)\right| \\ &\leqslant \sup_{x\in[0,2\pi]}\left|\phi(x) - \sum_{k=0}^{n} f_k(x)\right|,\end{aligned}$$

donc la convergence uniforme sur $[0, 2\pi]$ de la série $\sum_n f_n$ implique la convergence uniforme sur $[0, 2\pi]$ de la série $\sum_n g_n$. Par un argument analogue, on montre que pour tout $p \in \mathbb{N}$, la série de terme général $x \in \mathbb{R} \mapsto f_n(x)\sin(px)$ converge uniformément sur $\mathbb{R}$ et a pour somme $x \in \mathbb{R} \mapsto \phi(x)\sin(px)$.

Le théorème 3.8 p. 136 indique que l'on peut échanger l'ordre dans lequel sont effectuées la sommation et l'intégration : pour tout $p \in \mathbb{N}$, on a

$$\begin{aligned}\int_0^{2\pi} \phi(x)\cos(px)\,\mathrm{d}x &= \sum_{n=0}^{+\infty}\int_0^{2\pi} f_n(x)\cos(px)\,\mathrm{d}x \\ &= \sum_{n=0}^{+\infty}\left(a_n\int_0^{2\pi}\cos(nx)\cos(px)\,\mathrm{d}x + b_n\int_0^{2\pi}\sin(nx)\cos(px)\,\mathrm{d}x\right)\end{aligned}$$

et

$$\begin{aligned}\int_0^{2\pi} \phi(x)\sin(px)\,\mathrm{d}x &= \sum_{n=0}^{+\infty}\int_0^{2\pi} f_n(x)\sin(px)\,\mathrm{d}x \\ &= \sum_{n=0}^{+\infty}\left(a_n\int_0^{2\pi}\cos(nx)\sin(px)\,\mathrm{d}x + b_n\int_0^{2\pi}\sin(nx)\sin(px)\,\mathrm{d}x\right).\end{aligned}$$

De la relation[(2)] $\cos(nx)\cos(px) = \frac{1}{2}(\cos((n-p)x) + \cos((n+p)x))$, on déduit que :

$$\begin{aligned}\int_0^{2\pi}\cos(nx)\cos(px)\,\mathrm{d}x &= \frac{1}{2}\int_0^{2\pi}\cos((n-p)x)\,\mathrm{d}x + \frac{1}{2}\int_0^{2\pi}\cos((n+p)x)\,\mathrm{d}x \\ &= \begin{cases} 0 & \text{si } n \neq p \\ \pi & \text{si } n = p \text{ avec } n \neq 0 \\ 2\pi & \text{si } n = p = 0 \end{cases}.\end{aligned}$$

On a aussi

$$\begin{aligned}\int_0^{2\pi}\sin(nx)\cos(px)\,\mathrm{d}x &= \frac{1}{2}\int_0^{2\pi}\sin((n-p)x)\,\mathrm{d}x + \frac{1}{2}\int_0^{2\pi}\sin((n+p)x)\,\mathrm{d}x \\ &= 0\end{aligned}$$

(2) Voir le *Cours de première année* p. 155.

et

$$\int_0^{2\pi} \sin(nx)\sin(px)\ \mathrm{d}x = \frac{1}{2}\int_0^{2\pi} \cos((n-p)x)\ \mathrm{d}x - \frac{1}{2}\int_0^{2\pi} \cos((n+p)x)\ \mathrm{d}x$$
$$= \begin{cases} 0 & \text{si } n \neq p \\ \pi & \text{si } n = p \text{ avec } n \neq 0 \\ 0 & \text{si } n = p = 0 \end{cases}.$$

Ainsi, pour tout $p \in \mathbb{N}$ on a :

$$\int_0^{2\pi} \phi(x)\cos(px)\ \mathrm{d}x = \begin{cases} 2\pi a_0 & \text{si } p = 0 \\ \pi a_p & \text{si } p \neq 0 \end{cases}$$

et $$\int_0^{2\pi} \phi(x)\sin(px)\ \mathrm{d}x = \pi b_p.$$

On en déduit que lorsqu'une série trigonométrique $\sum_n a_n \cos(nx) + b_n \sin(nx)$ converge uniformément sur $[0, 2\pi]$, les coefficients de la série et la somme ϕ de cette série sont liés par les relations :

$$a_0 = \frac{1}{2\pi}\int_0^{2\pi} \phi(x)\ \mathrm{d}x,$$

$$a_p = \frac{1}{\pi}\int_0^{2\pi} \phi(x)\cos(px)\ \mathrm{d}x \qquad \text{pour tout } p \in \mathbb{N}^* \qquad (1)$$

et $$b_p = \frac{1}{\pi}\int_0^{2\pi} \phi(x)\sin(px)\ \mathrm{d}x \qquad \text{pour tout } p \in \mathbb{N}. \qquad (2)$$

On pourra retenir les formules suivantes.

Lemme 6.1 *Quels que soient les entiers naturels n et p on a*

$$\int_0^{2\pi} \cos(nx)\cos(px)\ \mathrm{d}x = \begin{cases} 0 & \textit{si } n \neq p \\ \pi & \textit{si } n = p \textit{ avec } n \neq 0 \\ 0 & \textit{si } n = p = 0 \end{cases}$$

$$\int_0^{2\pi} \sin(nx)\sin(px)\ \mathrm{d}x = \begin{cases} 0 & \textit{si } n \neq p \\ \pi & \textit{si } n = p \textit{ avec } n \neq 0 \\ 0 & \textit{si } n = p = 0 \end{cases}$$

$$\int_0^{2\pi} \sin(nx)\cos(px)\ \mathrm{d}x = 0$$

6.2 Décomposition en série de Fourier

6.2.1 Coefficients de Fourier

La décomposition d'une fonction f en série de Fourier consiste à trouver une série trigonométrique dont f est la somme. Il résulte des propriétés des séries trigonométriques que seule une fonction 2π-périodique peut être la somme

d'une série trigonométrique (voir la proposition 6.3). Par ailleurs, les coefficients d'une série trigonométrique qui converge uniformément sur $\mathbb{R}$ sont liés à la somme de cette série par les relations (1) et (2) p. 305. Ces considérations motivent la définition suivante.

DÉFINITION 6.2 (Série de Fourier)
Soit f une application définie sur $\mathbb{R}$, 2π-périodique et Riemann-intégrable sur $[0, 2\pi]$.

✗ *On appelle* coefficients de Fourier *de f les réels*[3]

$$a_0 = \frac{1}{2\pi}\int_0^{2\pi} f(x)\,\mathrm{d}x, \tag{3}$$

$$a_n = \frac{1}{\pi}\int_0^{2\pi} f(x)\cos(nx)\,\mathrm{d}x \qquad \textit{pour tout } n \in \mathbb{N}^* \tag{4}$$

$$\textit{et} \qquad b_n = \frac{1}{\pi}\int_0^{2\pi} f(x)\sin(nx)\,\mathrm{d}x \qquad \textit{pour tout } n \in \mathbb{N}. \tag{5}$$

✗ *On appelle* série de Fourier *de f la série trigonométrique de terme général*

$$x \in \mathbb{R} \mapsto a_n\cos(nx) + b_n\sin(nx).$$

Exemples

1. Considérons l'application $f : x \in \mathbb{R} \mapsto \cos^4(\frac{1}{2}x)$. Cette application est continue sur $\mathbb{R}$ et 2π-périodique. Pour calculer ses coefficients de Fourier, on peut linéariser[4] $\cos^4(\frac{1}{2}x)$. Pour tout réel θ, on a $\cos^4(\theta) = \frac{1}{8}(\cos(4\theta) + 4\cos(2\theta) + 3)$; donc pour tout réel x, on a $f(x) = \frac{1}{8}(\cos(2x) + 4\cos(x) + 3)$. On en déduit que :

$$a_0 = \frac{1}{8}\left(3 + \frac{1}{2\pi}\int_0^{2\pi}\cos(2x)\,\mathrm{d}x + \frac{2}{\pi}\int_0^{2\pi}\cos(x)\,\mathrm{d}x\right) = \frac{3}{8}.$$

De plus, pour tout $n \in \mathbb{N}^*$ on a :

$$a_n = \frac{1}{8}\Big(\frac{3}{\pi}\int_0^{2\pi}\cos(nx)\,\mathrm{d}x + \frac{1}{\pi}\int_0^{2\pi}\cos(2x)\cos(nx)\,\mathrm{d}x + \frac{4}{\pi}\int_0^{2\pi}\cos(x)\cos(nx)\,\mathrm{d}x\Big).$$

On déduit du lemme 6.1 que a_n est nul si $n \geqslant 3$ et que $a_1 = \frac{1}{2}$, $a_2 = \frac{1}{8}$. De même, le lemme 6.1 indique que b_n est nul quel que soit $n \in \mathbb{N}$. Finalement, la série de Fourier de f est $\frac{3}{8} + \frac{1}{2}\cos(x) + \frac{1}{8}\cos(2x)$, expression que nous avions

[3] On notera que b_0 est toujours nul.

[4] Voir la section 4.5.3 p. 158 du *Cours de première année* pour une présentation de la méthode générale et pour le détail de la linéarisation de $\cos^4(\theta)$.

obtenue par la linéarisation[5]. On remarque que f est égale sur $\mathbb{R}$ à sa série de Fourier.

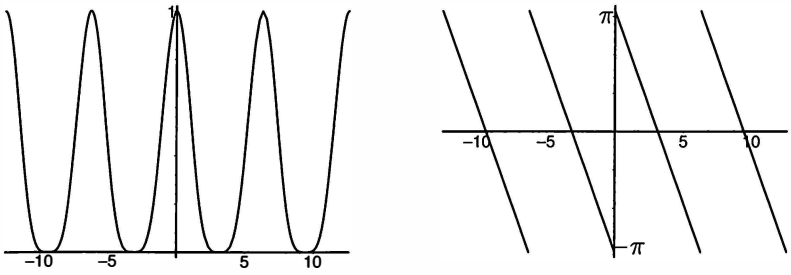


Fig. 1 Représentation graphique des fonctions 2π-périodiques $x \mapsto \cos(\frac{1}{2}x)^4$ (à gauche) et $x \in [0, 2\pi[\mapsto \pi - x$ (à droite).

2. Considérons l'application 2π-périodique f qui en restriction à l'intervalle $[0, 2\pi[$ est $x \mapsto \pi - x$. On a $a_0 = 0$ et en effectuant le changement de variable $y = \pi - x$, on obtient pour tout $n \in \mathbb{N}^*$:

$$\begin{aligned} a_n &= \frac{1}{\pi}\int_0^{2\pi} (\pi - x)\cos(nx)\ \mathrm{d}x = -\frac{1}{\pi}\int_{-\pi}^{\pi} y\cos(ny - n\pi)\ \mathrm{d}y \\ &= \frac{1}{\pi}(-1)^{n+1}\int_{-\pi}^{\pi} y\cos(ny)\ \mathrm{d}y. \end{aligned}$$

Comme la fonction $x \mapsto x\cos(nx)$ est impaire, on a $a_n = 0$. Pour tout $n \in \mathbb{N}^*$, on obtient en intégrant par parties :

$$\begin{aligned} b_n &= \frac{1}{\pi}\int_0^{2\pi} (\pi - x)\sin(nx)\ \mathrm{d}x \\ &= -\frac{1}{\pi}\Big[(\pi - x)\frac{\cos(nx)}{n}\Big]_0^{2\pi} - \frac{1}{n\pi}\int_0^{2\pi}\cos(nx)\ \mathrm{d}x = \frac{2}{n}. \end{aligned}$$

La série de Fourier de f est donc $\sum_{n\geqslant 1}\frac{2}{n}\sin(nx)$. La règle d'Abel pour les séries trigonométriques (voir la proposition 6.2) permet de conclure à la convergence simple de la série de Fourier de f sur $\mathbb{R} \setminus 2\pi\mathbb{Z}$. Il y a en fait convergence simple sur $\mathbb{R}$ car si $x \in 2\pi\mathbb{Z}$, la série de Fourier de f est la série nulle. Comme $f(x) = \pi$ pour $x \in 2\pi\mathbb{Z}$, on pourra noter qu'en $x \in 2\pi\mathbb{Z}$ la série de Fourier de f ne converge pas vers $f(x)$.

[5] Cela est lié au fait que la série de Fourier d'une application 2π-périodique et intégrable est unique.

[6] Voir : `www.academie-sciences.fr/MEMBRES/in_memoriam/Fourier/Fourier_oeuvre.htm`

FOURIER, Joseph (1768, Auxerre - 1830, Paris).



Joseph Fourier se montra durant la Révolution un révolutionnaire actif, animateur du comité local révolutionnaire d'Auxerre. En 1794, il fut de la première promotion de l'École Normale Supérieure. Il fit partie en 1798 de l'expédition napoléonienne en Égypte. Fourier y séjourna jusqu'en 1801 où il mena des travaux scientifiques, historiques, administratifs et diplomatiques. À son retour en France, il fut nommé préfet du département de l'Isère. C'est à Grenoble que Fourier réalisa l'essentiel de ses travaux scientifiques consacrés à l'étude de l'évolution de la température d'un corps au cours du temps. En 1808, il fut nommé baron par Napoléon et en 1817 il fut élu à l'Académie des Sciences[6] dont il devint le secrétaire perpétuel pour les sciences mathématiques.

Exercice 6.1 Calculer la série de Fourier de la fonction $x \mapsto |\sin(x/2)|$.

Les exemples précédents permettent de faire les constatations suivantes. La série de Fourier associée à une fonction f donnée n'est pas nécessairement convergente sur l'ensemble de définition de f. Par ailleurs, même si cette série converge, sa somme n'est pas nécessairement la fonction f.

Sans préjuger de la somme de la série de Fourier associée à une fonction f nous écrirons :

$$\mathcal{S}_F(f)(x) = \sum_{n=0}^{+\infty} a_n \cos(nx) + b_n \sin(nx).$$

Les conditions de convergence d'une série de Fourier selon la régularité de l'application f seront étudiées à la section 6.2.3.

Remarques

1. On peut étendre la définition des coefficients de Fourier donnée par les relations de la définition 6.2 au cas où f est localement Riemann-intégrable[7] sur $[0, 2\pi]$ et admet une intégrale généralisée absolument convergente sur $[0, 2\pi]$. Dans ce cas, pour tout entier naturel n, les intégrales $\int_0^{2\pi} f(x)\cos(nx)\,\mathrm{d}x$ et $\int_0^{2\pi} f(x)\sin(nx)\,\mathrm{d}x$ sont des intégrales généralisées absolument convergentes.

2. Il résulte des propriétés de l'intégrale de Riemann[8] que les coefficients de Fourier d'une fonction f ne sont pas modifiés si on modifie les valeurs de la fonction f un nombre fini de points. ◇

[7] Voir la définition 19.1 p. 917 du *Cours de première année*.

[8] Voir le *Cours de première année* p. 845.

Il existe une variante dans la façon de définir la série de Fourier d'une fonction f 2π-périodique et Riemann-intégrable sur $[0, 2\pi]$ qui consiste à poser pour tout $n \in \mathbb{N}$

$$a_n = \frac{1}{\pi}\int_0^{2\pi} f(x)\cos(nx)\,\mathrm{d}x, \qquad b_n = \frac{1}{\pi}\int_0^{2\pi} f(x)\sin(nx)\,\mathrm{d}x,$$

et

$$\mathcal{S}_F(f)(x) = \frac{a_0}{2} + \sum_{n=1}^{+\infty} a_n\cos(nx) + b_n\sin(nx).$$

Seule la valeur du coefficient a_0 diffère entre ces deux approches ; l'expression de la série de Fourier associée à la fonction f reste inchangée.

Écriture complexe d'une série de Fourier

En utilisant les formules d'Euler[9], on peut représenter la série de Fourier d'une fonction réelle de la variable réelle sous forme complexe. Considérons une application f de $\mathbb{R}$ dans $\mathbb{R}$, 2π-périodique et Riemann-intégrable sur $[0, 2\pi]$ dont la série de Fourier est la série trigonométrique $\sum_n f_n$ de terme général $f_n : x \in \mathbb{R} \mapsto a_n\cos(nx) + b_n\sin(nx)$. Les formules d'Euler permettent d'écrire pour tout $x \in \mathbb{R}$ et tout $n \in \mathbb{N}$,

$$\begin{aligned} f_n(x) &= \tfrac{1}{2}a_n\left(\mathrm{e}^{\mathrm{i}nx} + \mathrm{e}^{-\mathrm{i}nx}\right) + \tfrac{1}{2\mathrm{i}}b_n\left(\mathrm{e}^{\mathrm{i}nx} - \mathrm{e}^{-\mathrm{i}nx}\right) \\ &= \tfrac{1}{2}(a_n - \mathrm{i}b_n)\,\mathrm{e}^{\mathrm{i}nx} + \tfrac{1}{2}(a_n + \mathrm{i}b_n)\,\mathrm{e}^{-\mathrm{i}nx}. \end{aligned}$$

La série de Fourier de f s'écrit donc sous la forme $\sum_{k\in\mathbb{Z}} c_k\,\mathrm{e}^{\mathrm{i}kx}$ où on a posé pour $k \in \mathbb{Z}$,

$$\begin{cases} c_k = \tfrac{1}{2}(a_k - \mathrm{i}b_k) & \text{si } k > 0 \\ c_0 = a_0 \\ c_k = \tfrac{1}{2}(a_{-k} + \mathrm{i}b_{-k}) & \text{si } k < 0 \end{cases}$$

Pour tout $n \in \mathbb{N}$, les coefficients de Fourier réels a_n et b_n de f sont donnés par les relations (3), (4) et (5) p. 306. On déduit de ces relations, l'expression des coefficients de Fourier complexes de f ; pour tout $k \in \mathbb{N}^*$, on a[10]

$$c_k = \frac{1}{2\pi}\int_0^{2\pi} f(x)\left(\cos(kx) - \mathrm{i}\sin(kx)\right)\mathrm{d}x = \frac{1}{2\pi}\int_0^{2\pi} f(x)\,\mathrm{e}^{-\mathrm{i}kx}\,\mathrm{d}x;$$

pour $k = 0$, on a

$$c_0 = \frac{1}{2\pi}\int_0^{2\pi} f(x)\,\mathrm{d}x$$

[9] Voir la proposition 4.6 p. 138 du *Cours de première année*.

[10] Dans cette expression, le sens à donner à l'intégrale entre 0 et 2π de la fonction réelle à valeurs complexes $x \in \mathbb{R} \mapsto f(x)\,\mathrm{e}^{-\mathrm{i}kx}$ est celui indiqué p. 162.

et pour $k \in \mathbb{Z} \setminus \mathbb{N}$ on a

$$\begin{aligned} c_k &= \frac{1}{2\pi}\int_0^{2\pi} f(x)\left(\cos(-kx)+\mathrm{i}\sin(-kx)\right)\,\mathrm{d}x \\ &= \frac{1}{2\pi}\int_0^{2\pi} f(x)\left(\cos(kx)-\mathrm{i}\sin(kx)\right)\,\mathrm{d}x \;=\; \frac{1}{2\pi}\int_0^{2\pi} f(x)\,\mathrm{e}^{-\mathrm{i}kx}\,\mathrm{d}x. \end{aligned}$$

On remarquera que ces trois expressions des coefficients c_k pour $k > 0, k = 0$ et $k < 0$ prennent une même forme :

$$\boxed{\forall k \in \mathbb{Z} \qquad c_k = \frac{1}{2\pi}\int_0^{2\pi} f(x)\,\mathrm{e}^{-\mathrm{i}kx}\,\mathrm{d}x.}$$

La série de Fourier associée à une application f définie sur $\mathbb{R}$, 2π-périodique et Riemann-intégrable sur $[0, 2\pi]$ s'écrit sous la forme complexe

$$\sum_{k\in\mathbb{Z}} c_k\,\mathrm{e}^{\mathrm{i}kx}.$$

Dans la manipulation des séries de Fourier données sous forme complexe, on sera attentif au fait que la sommation est effectuée sur les entiers relatifs et non pas sur les entiers naturels comme cela a été le cas jusqu'à présent. La série de Fourier donnée sous forme complexe $\sum_{k\in\mathbb{Z}} c_k\,\mathrm{e}^{\mathrm{i}kx}$ converge si la suite $(S_n)_n$ des sommes partielles de terme général $S_n = \sum_{k=-n}^{n} c_k\,\mathrm{e}^{\mathrm{i}kx}$ converge.

Lemme 6.2 (Formule de Dirichlet)
Soit f une application définie sur $\mathbb{R}$, 2π-périodique et Riemann-intégrable sur $[0, 2\pi]$. La suite des sommes partielles $(S_n)_n$ associée à la série de Fourier de f vérifie pour tous $n \in \mathbb{N}$ et $x \in \mathbb{R}$

$$S_n(x) = \frac{1}{2\pi}\int_0^{\pi} \frac{\sin((2n+1)\frac{t}{2})}{\sin(\frac{t}{2})}\left(f(x+t)+f(x-t)\right)\,\mathrm{d}t.$$

Démonstration ⊵ Commençons par établir le résultat préliminaire suivant : pour tout $t \in \mathbb{R} \setminus 2\pi\mathbb{Z}$ et $n \in \mathbb{N}$, on a

$$\boxed{\frac{1}{2} + \sum_{k=1}^{n} \cos(kt) = \frac{\sin((2n+1)\frac{t}{2})}{2\sin(\frac{t}{2})}.} \tag{6}$$

D'après les formules d'Euler[11] on a : $\cos(kt) = \frac{1}{2}(\mathrm{e}^{\mathrm{i}kt} + \mathrm{e}^{-\mathrm{i}kt})$; par conséquent, en utilisant les propriétés des progressions géométriques, on obtient :

$$\sum_{k=1}^{n}\cos(kt) = \frac{1}{2}\sum_{k=1}^{n}\mathrm{e}^{\mathrm{i}kt} + \frac{1}{2}\sum_{k=1}^{n}\mathrm{e}^{-\mathrm{i}kt} = \frac{\mathrm{e}^{\mathrm{i}t}}{2}\frac{1-\mathrm{e}^{\mathrm{i}nt}}{1-\mathrm{e}^{\mathrm{i}t}} + \frac{\mathrm{e}^{-\mathrm{i}t}}{2}\frac{1-\mathrm{e}^{-\mathrm{i}nt}}{1-\mathrm{e}^{-\mathrm{i}t}}.$$

[11] Voir la proposition 4.6 p. 138 du *Cours de première année*.

On vérifie aisément en utilisant les formules d'Euler que $1-\mathrm{e}^{\mathrm{i}t} = -2\mathrm{i}\mathrm{e}^{\mathrm{i}\frac{t}{2}}\sin(\frac{t}{2})$ et que $1-\mathrm{e}^{-\mathrm{i}t} = 2\mathrm{i}\mathrm{e}^{-\mathrm{i}\frac{t}{2}}\sin(\frac{t}{2})$. On en déduit que :

$$\begin{aligned}\sum_{k=1}^{n}\cos(kt) &= \frac{1}{4\mathrm{i}\sin(\frac{t}{2})}\left(-\mathrm{e}^{\mathrm{i}\frac{t}{2}}+\mathrm{e}^{\mathrm{i}(nt+\frac{t}{2})}+\mathrm{e}^{-\mathrm{i}\frac{t}{2}}-\mathrm{e}^{-\mathrm{i}(nt+\frac{t}{2})}\right)\\ &= \frac{1}{2\sin(\frac{t}{2})}\left(\sin((2n+1)\tfrac{t}{2})-\sin(\tfrac{t}{2})\right) = \frac{\sin((2n+1)\frac{t}{2})}{2\sin(\frac{t}{2})}-\frac{1}{2}.\end{aligned}$$

La relation (6) se déduit de la dernière égalité. Remarquons que si $t \in 2\pi\mathbb{Z}$ alors $\frac{1}{2}+\sum_{k=1}^{n}\cos(kt) = \frac{1}{2}(2n+1)$, ce qui nous amène à définir l'application :

$$K_n : t \in \mathbb{R} \mapsto \begin{cases} \dfrac{\sin((2n+1)\frac{t}{2})}{2\sin(\frac{t}{2})} & \text{si } t \in \mathbb{R}\setminus 2\pi\mathbb{Z} \\ \dfrac{1}{2}(2n+1) & \text{sinon} \end{cases} \tag{7}$$

Pour tout $t \in \mathbb{R}$, on a la relation $\frac{1}{2}+\sum_{k=1}^{n}\cos(kt) = K_n(t)$ et il est aisé d'établir que K_n est continue sur $\mathbb{R}$ en utilisant le fait que $\sin(x) \underset{0}{\sim} x$ et que K_n est 2π-périodique.

$\trianglerighteq$ La suite des sommes partielles $(S_n)_n$ associée à la série de Fourier de f a pour terme général $S_n : x \in \mathbb{R} \mapsto \sum_{k=0}^{n} a_k\cos(kx)+b_k\sin(kx)$ où pour $k \in \{0,\dots,n\}$ les réels a_k et b_k sont donnés par les relations (3) à (5) p. 306. Ces relations permettent d'écrire pour tout $x \in \mathbb{R}$:

$$\begin{aligned}S_n(x) &= \frac{1}{2\pi}\int_0^{2\pi} f(t)\,\mathrm{d}t + \frac{1}{\pi}\sum_{k=1}^{n}\int_0^{2\pi} f(t)\big(\cos(kt)\cos(kx)+\sin(kt)\sin(kx)\big)\,\mathrm{d}t\\ &= \frac{1}{2\pi}\int_0^{2\pi} f(t)\,\mathrm{d}t + \frac{1}{\pi}\sum_{k=1}^{n}\int_0^{2\pi} f(t)\,\cos(k(t-x))\,\mathrm{d}t\\ &= \frac{1}{\pi}\int_0^{2\pi} f(t)\left(\frac{1}{2}+\sum_{k=1}^{n}\cos(k(t-x))\right).\mathrm{d}t.\end{aligned}$$

On déduit alors de la relation (6) que : $S_n(x) = \dfrac{1}{\pi}\displaystyle\int_0^{2\pi} f(t)\,K_n(t-x)\,\mathrm{d}t$. En effectuant le changement de variables $s = t-x$ et en remarquant que la fonction intégrée est 2π-périodique, il est indifférent d'intégrer sur l'intervalle de période $[-x, 2\pi-x]$ ou $[-\pi,\pi]$, on obtient :

$$S_n(x) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(s+x)\,K_n(s)\,\mathrm{d}s,$$

En utilisant le fait que la restriction de K_n à $[-\pi, \pi]$ est paire, on peut écrire :

$$\begin{aligned} S_n(x) &= \frac{1}{\pi}\int_{-\pi}^{0} f(s+x)\,K_n(s)\,\mathrm{d}s + \frac{1}{\pi}\int_{0}^{\pi} f(s+x)\,K_n(s)\,\mathrm{d}s \\ &= \frac{1}{\pi}\int_{0}^{\pi} f(-s+x)\,K_n(-s)\,\mathrm{d}s + \frac{1}{\pi}\int_{0}^{\pi} f(s+x)\,K_n(s)\,\mathrm{d}s \\ &= \frac{1}{\pi}\int_{0}^{\pi} \big(f(x+s)+f(x-s)\big)\,K_n(s)\,\mathrm{d}s. \end{aligned}$$

Le résultat est démontré ; d'après (7) on a en effet :

$$S_n(x) = \frac{1}{2\pi}\int_{0}^{\pi} \frac{\sin((2n+1)\frac{t}{2})}{\sin(\frac{t}{2})}\big(f(x+t)+f(x-t)\big)\,\mathrm{d}t.$$

□

Exercice 6.2 En utilisant la formule de Dirichlet, montrer que pour tout $n \in \mathbb{N}$ on a

$$\int_{0}^{\pi} \frac{\sin((2n+1)\frac{t}{2})}{\sin(\frac{t}{2})}\,\mathrm{d}t = \pi.$$

Calcul des coefficients de Fourier avec MAPLE

MAPLE peut se révéler utile pour calculer les coefficients de Fourier d'une fonction. Dans certains cas, MAPLE parvient à déterminer l'expression générique des coefficients a_n et b_n en fonction de n. Considérons par exemple le calcul des coefficients de Fourier de la fonction $x \mapsto |\sin(x/2)|$ effectué à l'exercice 6.1.

```
> f:= x-> abs(sin(x/2)):
> assume(n::integer); additionally(n>0);
> a_n:=(1/Pi)*int(f(x)*cos(n*x),x=0..2*Pi);
```

$$a_n := -4\frac{1}{\pi(-1+4n^2)}$$

```
> b_n:= (1/Pi)*int(f(x)*sin(n*x),x=0..2*Pi);
```

$$b_n := 0$$

On remarquera, en effectuant le calcul avec MAPLE des coefficients de Fourier de différentes fonctions, que l'on obtient assez fréquemment une expression des coefficients de Fourier à l'aide de fonctions « spéciales ».

Dans d'autres cas, MAPLE ne parvient pas à calculer explicitement les coefficients de Fourier de la fonction considérée. C'est le cas par exemple avec la fonction $x \mapsto \exp(\sin(x))$.

La procédure suivante permet d'obtenir une valeur numérique approchée des N premiers coefficients de Fourier d'une fonction f donnée. Les coefficients de Fourier $a_0, \ldots, a_N$ et $b_0, \ldots, b_N$ sont retournés dans les tableaux `a` et `b`

```
SFcoef:=proc(f,N,a,b)
  local n;
  a[0]:= (2*Pi)^(-1)*int(f(t),t=0..2*Pi):
  b[0]:= 0;
  for n from 1 to N do
    a[n]:= evalf((Pi)^(-1)*Int(f(t)*cos(n*t),t=0..2*Pi)):
    b[n]:= evalf((Pi)^(-1)*Int(f(t)*sin(n*t),t=0..2*Pi)):
  od:
end proc:
```

On peut ainsi obtenir des valeurs approchées pour les premiers coefficients de Fourier de la fonction $x \mapsto \exp(\sin(x))$.

```
> f:=x-> exp(sin(x));
> SFcoef(f,4,a,b);
> seq(printf("n = %d | a_n = %+.6e | b_n = %+.6e \n",n,a[n],b[n]),
      n=0..4);
n = 0  |  a_n = +1.266066e+00  |  b_n = +0.000000e-01
n = 1  |  a_n = +1.319448e-13  |  b_n = +1.130318e+00
n = 2  |  a_n = -2.714953e-01  |  b_n = +1.186598e-17
n = 3  |  a_n = +1.317857e-13  |  b_n = -4.433685e-02
n = 4  |  a_n = +5.474240e-03  |  b_n = -9.111633e-17
```

Coefficients de Fourier d'une fonction périodique de période T

Il est possible de définir la notion de série de Fourier pour des fonctions périodiques dont la période est différente de 2π. Il est aisé de vérifier que si f est une application périodique de période $T \in \mathbb{R}_+^*$ alors $\phi : x \in \mathbb{R} \mapsto f(x/\omega)$ où $\omega = 2\pi/T$ est une application périodique de période 2π. En effectuant le changement de variable $t = x/\omega$, on obtient pour tout $n \in \mathbb{N}$:

$$\begin{aligned} b_n &= \frac{1}{\pi}\int_0^{2\pi} \phi(x)\sin(nx)\ \mathrm{d}x = \frac{1}{\pi}\int_0^{2\pi} f(x/\omega)\ \sin(nx)\ \mathrm{d}x \\ &= \frac{2}{T}\int_0^{T} f(t)\ \sin(n\omega t)\ \mathrm{d}t. \end{aligned}$$

De même, pour tout $n \in \mathbb{N}^*$ on obtient :

$$a_n = \frac{1}{\pi}\int_0^{2\pi} \phi(x)\ \cos(nx)\ \mathrm{d}x = \frac{2}{T}\int_0^{T} f(t)\ \cos(n\omega t)\ \mathrm{d}t$$

$$\text{et} \quad a_0 = \frac{1}{2\pi}\int_0^{2\pi} \phi(x)\ \mathrm{d}x = \frac{1}{T}\int_0^{T} f(t)\ \mathrm{d}t.$$

Les réels a_n et b_n sont appelés coefficients de Fourier de rang n de f et la série

$$\sum_n a_n \cos(n\omega x) + b_n \sin(n\omega x)$$

est appelée série de Fourier de f.

Tous les résultats liés aux propriétés des séries de Fourier pour les fonctions 2π-périodiques donnés dans ce chapitre s'étendent aux fonctions périodiques de période T en exploitant le changement de variable indiqué.

6.2.2 Propriétés des coefficients de Fourier

La proposition suivante indique que l'on peut choisir n'importe quel intervalle de période pour calculer les coefficients de Fourier d'une application 2π-périodique.

PROPOSITION 6.6 *Soit f une application définie sur $\mathbb{R}$, 2π-périodique et Riemann-intégrable sur $[0, 2\pi]$. Pour tout réel α on a*

$$a_0 = \frac{1}{2\pi}\int_\alpha^{\alpha+2\pi} f(x)\,\mathrm{d}x,$$

$$a_n = \frac{1}{\pi}\int_\alpha^{\alpha+2\pi} f(x)\cos(nx)\,\mathrm{d}x \qquad \textit{pour tout } n \in \mathbb{N}^*,$$

$$\textit{et} \qquad b_n = \frac{1}{\pi}\int_\alpha^{\alpha+2\pi} f(x)\sin(nx)\,\mathrm{d}x \qquad \textit{pour tout } n \in \mathbb{N}.$$

Démonstration Ce résultat découle des propriétés générales de l'intégrale des fonctions périodiques[12]. Soit g une application définie sur $\mathbb{R}$, 2π-périodique et intégrable sur $[0, 2\pi]$. En utilisant la relation de Chasles on a :

$$\int_\alpha^{\alpha+2\pi} g(t)\,\mathrm{d}t = \int_\alpha^0 g(t)\,\mathrm{d}t + \int_0^{2\pi} g(t)\,\mathrm{d}t + \int_{2\pi}^{\alpha+2\pi} g(t)\,\mathrm{d}t.$$

Or, $\displaystyle\int_{2\pi}^{\alpha+2\pi} g(t)\,\mathrm{d}t = \int_{\phi(0)}^{\phi(\alpha)} g(t)\,\mathrm{d}t$ où $\phi : x \in [0, \alpha] \mapsto x + 2\pi$. En utilisant la formule du changement de variable[13], on obtient :

$$\int_{2\pi}^{\alpha+2\pi} g(t)\,\mathrm{d}t = \int_0^\alpha g(\phi(t))\,\phi'(t)\,\mathrm{d}t = \int_0^\alpha g(t+2\pi)\,\mathrm{d}t = \int_0^\alpha g(t)\,\mathrm{d}t,$$

la dernière égalité résultant du fait que g est 2π-périodique. On en conclut que :

$$\int_\alpha^{\alpha+2\pi} g(t)\,\mathrm{d}t = \int_\alpha^0 g(t)\,\mathrm{d}t + \int_0^{2\pi} g(t)\,\mathrm{d}t + \int_0^\alpha g(t)\,\mathrm{d}t = \int_0^{2\pi} g(t)\,\mathrm{d}t.$$

Les relations proposées découlent de l'utilisation de ce résultat avec les fonctions $x \mapsto f(x)\cos(nx)$ et $x \mapsto f(x)\sin(nx)$ pour $n \in \mathbb{N}$. □

[12] Voir la proposition 18.15 p. 807 du *Cours de première année*.

[13] Voir le théorème 18.2 p. 803 du *Cours de première année*.

Corollaire 6.1 *Soit f une application définie sur $\mathbb{R}$, 2π-périodique et intégrable sur $[0, 2\pi]$.*

- *Si f est paire alors pour tout $n \in \mathbb{N}$ on a $b_n = 0$ et*

$$a_0 = \frac{1}{\pi}\int_0^\pi f(x)\ \mathrm{d}x, \qquad a_n = \frac{2}{\pi}\int_0^\pi f(x)\cos(nx)\ \mathrm{d}x.$$

- *Si f est impaire alors pour tout $n \in \mathbb{N}$ on a $a_n = 0$ et*

$$b_n = \frac{2}{\pi}\int_0^\pi f(x)\sin(nx)\ \mathrm{d}x.$$

Démonstration Supposons que f est paire et appliquons la proposition 6.6 avec $\alpha = -\pi$. En utilisant la relation de Chasles, on obtient pour tout $n \in \mathbb{N}$:

$$b_n = \frac{1}{\pi}\int_{-\pi}^\pi f(x)\sin(nx)\ \mathrm{d}x = \frac{1}{\pi}\int_{-\pi}^0 f(x)\sin(nx)\ \mathrm{d}x + \frac{1}{\pi}\int_0^\pi f(x)\sin(nx)\ \mathrm{d}x.$$

L'application f étant paire, le changement de variable $x \mapsto -x$ indique que $\int_{-\pi}^0 f(x)\sin(nx)\ \mathrm{d}x = -\int_0^\pi f(x)\sin(nx)\ \mathrm{d}x$, ce qui implique que $b_n = 0$.
On obtient aussi en utilisant le changement de variable $x \mapsto -x$, compte tenu de la parité de f, que pour tout $n \in \mathbb{N}$

$$\int_{-\pi}^0 f(x)\cos(nx)\ \mathrm{d}x = \int_0^\pi f(x)\cos(nx)\ \mathrm{d}x.$$

Cela implique que pour $n \in \mathbb{N}^*$,

$$\begin{aligned} a_n &= \frac{1}{\pi}\int_{-\pi}^\pi f(x)\cos(nx)\ \mathrm{d}x \\ &= \frac{1}{\pi}\int_{-\pi}^0 f(x)\cos(nx)\ \mathrm{d}x + \frac{1}{\pi}\int_0^\pi f(x)\cos(nx)\ \mathrm{d}x \\ &= \frac{2}{\pi}\int_0^\pi f(x)\cos(nx)\ \mathrm{d}x, \end{aligned}$$

et que

$$a_0 = \frac{1}{2\pi}\int_{-\pi}^\pi f(x)\ \mathrm{d}x = \frac{1}{2\pi}\int_{-\pi}^0 f(x)\ \mathrm{d}x + \frac{1}{2\pi}\int_0^\pi f(x)\ \mathrm{d}x = \frac{1}{\pi}\int_0^\pi f(x)\ \mathrm{d}x.$$

Dans le cas où f est impaire, les relations données s'obtiennent de façon similaire. □

Exercice 6.3 Soit f une application 2π-périodique de classe $\mathcal{C}^1$ sur $\mathbb{R}$. En utilisant une intégration par parties, montrer que les suites $(a_n)_n$ et $(b_n)_n$ des coefficients de Fourier de f convergent vers 0.

Le résultat suivant indique que la propriété établie à l'exercice 6.3 est valable non seulement pour des fonctions qui sont de classe $\mathcal{C}^1$ sur $\mathbb{R}$ mais plus généralement pour toute fonction Riemann-intégrable sur $[0, 2\pi]$. Bien entendu, la

démonstration requiert une justification différente de celle utilisée pour l'exercice 6.3 ; f n'étant plus nécessairement dérivable et à dérivée intégrable, il n'est plus possible d'avoir recours à la formule d'intégration par parties.

PROPOSITION 6.7 *Soit f une application définie sur $\mathbb{R}$, 2π-périodique et Riemann-intégrable sur $[0, 2\pi]$. Les suites $(a_n)_n$ et $(b_n)_n$ de ses coefficients de Fourier convergent vers 0.*

Démonstration Dans un premier temps, nous montrerons le résultat dans le cas où f est une fonction en escalier[14], puis nous établirons le résultat dans le cas général d'une fonction Riemann-intégrable en ayant recours à la définition de la Riemann-intégrabilité.

$\trianglerighteq$ Supposons que l'application f est en escalier sur $[0, 2\pi]$ et considérons une subdivision $\sigma = (x_i)_{i=0,\dots,p}$ de $[0, 2\pi]$ adaptée[14] à f. Pour $i \in \{1, \dots, p\}$, désignons par λ_i la valeur prise par f sur l'intervalle $]x_{i-1}, x_i[$. Pour tout $n \in \mathbb{N}^*$, on a

$$\begin{aligned} \pi a_n &= \int_0^{2\pi} f(x)\cos(nx)\,\mathrm{d}x = \sum_{i=1}^{p} \int_{x_{i-1}}^{x_i} f(x)\cos(nx)\,\mathrm{d}x \\ &= \sum_{i=1}^{p} \lambda_i \int_{x_{i-1}}^{x_i} \cos(nx)\,\mathrm{d}x = \frac{1}{n}\sum_{i=1}^{p} \lambda_i\big(\sin(nx_i) - \sin(nx_{i-1})\big). \end{aligned}$$

Puisque $|\sin(nx_i) - \sin(nx_{i-1})| \leqslant |\sin(nx_i)| + |\sin(nx_{i-1})| \leqslant 2$, on en déduit que $|\pi a_n| \leqslant \frac{2}{n}\sum_{i=1}^{p} |\lambda_i|$ puis par le théorème d'encadrement[15] que $\lim\limits_{n\to+\infty} a_n = 0$. On procède de la même façon pour montrer que $\lim\limits_{n\to+\infty} b_n = 0$.

$\trianglerighteq$ Considérons à présent le cas d'une application f Riemann-intégrable sur $[0, 2\pi]$ et montrons que la suite $(a_n)_n$ de terme général $a_n = \frac{1}{\pi}\int_0^{2\pi} f(x)\cos(nx)\,\mathrm{d}x$ converge vers 0. Pour cela, montrons que[16] :

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad (n \geqslant N \Longrightarrow |a_n| \leqslant \varepsilon). \tag{8}$$

Soit ε un réel strictement positif. Comme f est Riemann-intégrable sur $[0, 2\pi]$, il existe[17] deux fonctions ϕ_ε et ψ_ε en escalier sur $[0, 2\pi]$ telles que $\psi_\varepsilon \leqslant f \leqslant \phi_\varepsilon$ et $\int_0^{2\pi}(\phi_\varepsilon(x) - \psi_\varepsilon(x))\,\mathrm{d}x < \pi\varepsilon/2$, voir la figure 2. On a donc

$$0 \leqslant \int_0^{2\pi} (\phi_\varepsilon(x) - f(x))\,\mathrm{d}x \leqslant \int_0^{2\pi} (\phi_\varepsilon(x) - \psi_\varepsilon(x))\,\mathrm{d}x < \pi\varepsilon/2. \tag{9}$$

[14] Voir la définition 18.2 p. 832 du *Cours de première année*.
[15] Voir le théorème 5.1 p. 179 du *Cours de première année*.
[16] Voir la définition 5.1 p. 168 du *Cours de première année*.
[17] Voir la définition 18.4 p. 836 du *Cours de première année*.

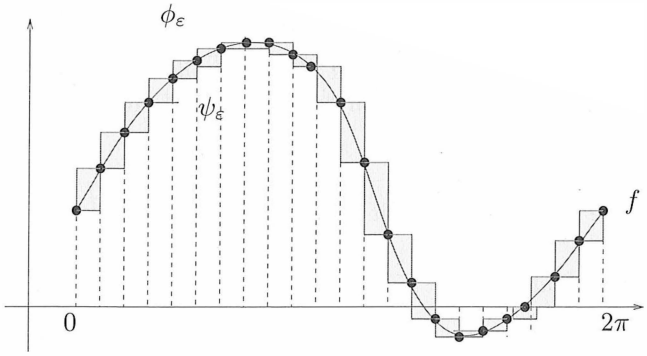


Fig. 2 Deux fonctions en escaliers vérifiant $\psi_\varepsilon \leqslant f \leqslant \phi_\varepsilon$; la surface de la partie grisée correspond à $\int_0^{2\pi} (\phi_\varepsilon(x) - \psi_\varepsilon(x))\ \mathrm{d}x$.

Par ailleurs, pour tout $n \in \mathbb{N}^*$, on a

$$\begin{aligned} \pi a_n &= \int_0^{2\pi} f(x)\cos(nx)\ \mathrm{d}x \\ &= \int_0^{2\pi} (f(x) - \phi_\varepsilon(x))\cos(nx)\ \mathrm{d}x + \int_0^{2\pi} \phi_\varepsilon(x)\cos(nx)\ \mathrm{d}x. \end{aligned}$$

Notons $\alpha_n = \frac{1}{\pi}\int_0^{2\pi} \phi_\varepsilon(x)\cos(nx)\ \mathrm{d}x$; il s'agit d'un des deux coefficients de Fourier de rang n de ϕ_ε. Comme ϕ_ε est une fonction 2π-périodique et en escalier sur $[0, 2\pi]$, d'après ce qui a été établi dans la première partie de la démonstration la suite $(\alpha_n)_n$ converge vers 0. On a donc[17]

$$\exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad (n \geqslant N \Longrightarrow |\alpha_n| \leqslant \tfrac{1}{2}\varepsilon). \tag{10}$$

On déduit des assertions (9) et (10) que pour tout entier $n \geqslant N$ on a

$$\begin{aligned} |\pi a_n| &= \left| \int_0^{2\pi} (f(x) - \phi_\varepsilon(x))\cos(nx)\ \mathrm{d}x + \pi\alpha_n \right| \\ &\leqslant \left| \int_0^{2\pi} (f(x) - \phi_\varepsilon(x))\ \mathrm{d}x \right| + |\pi\alpha_n| \ \leqslant\ \tfrac{\pi}{2}\varepsilon + \tfrac{\pi}{2}\varepsilon = \pi\varepsilon. \end{aligned}$$

Pour tout entier $n \geqslant N$ on a donc $|a_n| \leqslant \varepsilon$. L'assertion (8) est démontrée. On procède de la même manière pour montrer que $\lim_{n\to+\infty} b_n = 0$. □

Remarques

1. Le résultat de la proposition 6.8 demeure vrai si f est localement Riemann-intégrable sur $[0, 2\pi]$.

2. On déduit de la proposition 6.8 en considérant le changement de variable $x \mapsto 2\pi(x-a)/(b-a)$ que si f est une application de $[a, b]$ dans $\mathbb{R}$, Riemann-intégrable sur $[a, b]$ alors les suites $(\alpha_n)_n$ et $(\beta_n)_n$ de terme général

$$\alpha_n = \int_a^b f(t)\cos(n\omega t)\ \mathrm{d}t \quad \text{et} \quad \beta_n = \int_a^b f(t)\sin(n\omega t)\ \mathrm{d}t$$

où $\omega = 2\pi/(b-a)$, convergent vers 0.

3. On a également le résultat suivant appelé *Lemme de Lebesgue* que nous admettrons : si f est une application de $[a,b]$ dans $\mathbb{R}$, Riemann-intégrable sur $[a,b]$ alors les applications

$$A : x \in \mathbb{R} \longmapsto \int_a^b f(t)\cos(tx)\,\mathrm{d}t \quad \text{et} \quad B : x \in \mathbb{R} \longmapsto \int_a^b f(t)\sin(tx)\,\mathrm{d}t$$

tendent vers 0 quand le réel x tend vers $+\infty$. ◇

PROPOSITION 6.8 *Soit f une application 2π-périodique de classe $\mathcal{C}^k$ sur $\mathbb{R}$. Pour $n \in \mathbb{N}^*$, les coefficients de Fourier $a_n(f)$ et $b_n(f)$ de f et les coefficients de Fourier $a_n(f^{(k)})$ et $b_n(f^{(k)})$ de $f^{(k)}$ sont liés par les relations suivantes :*

1. si $k \in \mathbb{N}^*$ est pair alors :

$$a_n(f) = \frac{(-1)^{\frac{k}{2}}}{n^k}\, a_n(f^{(k)}) \quad \text{et} \quad b_n(f) = \frac{(-1)^{\frac{k}{2}}}{n^k}\, b_n(f^{(k)});$$

2. si $k \in \mathbb{N}^*$ est impair alors :

$$a_n(f) = \frac{(-1)^{\frac{k+1}{2}}}{n^k}\, b_n(f^{(k)}) \quad \text{et} \quad b_n(f) = -\frac{(-1)^{\frac{k+1}{2}}}{n^k}\, a_n(f^{(k)}).$$

Démonstration Démontrons le résultat en utilisant un raisonnement par récurrence portant sur l'entier k.

▷ Considérons une application 2π-périodique f de classe $\mathcal{C}^1$ sur $\mathbb{R}$. Comme f est une application dérivable et 2π-périodique, sa dérivée f' est définie sur $\mathbb{R}$ et est elle aussi 2π-périodique. De plus, comme f est de classe $\mathcal{C}^1$ sur $\mathbb{R}$, f' est continue sur $[0,2\pi]$ donc intégrable sur $[0,2\pi]$. En intégrant par parties, on obtient pour $n \in \mathbb{N}^*$:

$$\pi b_n(f) = \int_0^{2\pi} f(x)\sin(nx)\,\mathrm{d}x = \Big[-\frac{1}{n}\cos(nx)\,f(x)\Big]_0^{2\pi} + \frac{1}{n}\int_0^{2\pi} f'(x)\cos(nx)\,\mathrm{d}x.$$

Comme f est 2π-périodique, on a $\Big[\cos(nx)f(x)\Big]_0^{2\pi} = f(2\pi) - f(0) = 0$, ce qui implique que

$$b_n(f) = \frac{1}{n\pi}\int_0^{2\pi} f'(x)\cos(nx)\,\mathrm{d}x = \frac{1}{n}\,a_n(f').$$

Par un calcul analogue, on obtient

$$a_n(f) = -\frac{1}{n}\,b_n(f').$$

La propriété est établie pour $k = 1$.

$\trianglerighteq$ Supposons la propriété vraie pour un entier $k \geqslant 1$ donné et montrons qu'elle est alors également vraie pour l'entier suivant $k+1$. Soit f une application 2π-périodique, de classe $\mathcal{C}^{k+1}$ sur $\mathbb{R}$. L'application $f^{(k)}$ est de classe $\mathcal{C}^1$ sur $[0, 2\pi]$ et en intégrant par parties, on obtient

$$\begin{aligned} a_n(f^{(k)}) &= \frac{1}{\pi}\int_0^{2\pi} f^{(k)}(x)\cos(nx)\,\mathrm{d}x \\ &= \Big[\frac{1}{\pi n}\sin(nx)\,f^{(k)}(x)\Big]_0^{2\pi} - \frac{1}{\pi n}\int_0^{2\pi} f^{(k+1)}(x)\sin(nx)\,\mathrm{d}x. \end{aligned}$$

Comme f est 2π-périodique et de classe $\mathcal{C}^{k+1}$ sur $\mathbb{R}$, ses dérivées successives jusqu'à l'ordre $k+1$ sont également 2π-périodiques[18], ce qui implique que le premier terme du membre de droite est nul. L'intégrale dans le second terme du membre de droite étant égale à $\pi b_n(f^{(k+1)})$, on en déduit que

$$a_n(f^{(k)}) = -\frac{1}{n}\,b_n(f^{(k+1)}).$$

Un calcul similaire permet d'établir que

$$b_n(f^{(k)}) = \frac{1}{n}\,a_n(f^{(k+1)}).$$

$\triangleright$ Une application de classe $\mathcal{C}^{k+1}$ étant de classe $\mathcal{C}^k$, il résulte de l'hypothèse de récurrence que si k est impair (autrement dit si $\ell = k+1$ est pair) alors

$$a_n(f) = \frac{(-1)^{\frac{k+1}{2}}}{n^k}\,b_n(f^{(k)}) = \frac{(-1)^{\frac{k+1}{2}}}{n^{k+1}}\,a_n(f^{(k+1)}) = \frac{(-1)^{\frac{\ell}{2}}}{n^\ell}\,a_n(f^{(\ell)})$$

et

$$b_n(f) = -\frac{(-1)^{\frac{k+1}{2}}}{n^k}\,a_n(f^{(k)}) = \frac{(-1)^{\frac{k+1}{2}}}{n^{k+1}}\,b_n(f^{(k+1)}) = \frac{(-1)^{\frac{\ell}{2}}}{n^\ell}\,b_n(f^{(\ell)}).$$

$\triangleright$ Par ailleurs, il résulte de l'hypothèse de récurrence que si k est pair (autrement dit si $\ell = k+1$ est impair) alors

$$a_n(f) = \frac{(-1)^{\frac{k}{2}}}{n^k}\,a_n(f^{(k)}) = -\frac{(-1)^{\frac{k}{2}}}{n^{k+1}}\,b_n(f^{(k+1)}) = \frac{(-1)^{\frac{\ell+1}{2}}}{n^\ell}\,b_n(f^{(\ell)})$$

et

$$b_n(f) = \frac{(-1)^{\frac{k}{2}}}{n^k}\,b_n(f^{(k)}) = \frac{(-1)^{\frac{k}{2}}}{n^{k+1}}\,a_n(f^{(k+1)}) = -\frac{(-1)^{\frac{\ell+1}{2}}}{n^\ell}\,a_n(f^{(\ell)}).$$

Dans les deux cas (k pair et k impair) les relations sont vraies pour l'entier $\ell = k+1$. La propriété est donc bien héréditaire. □

[18] On notera que pour effectuer l'intégration par parties, il suffit que f soit de classe $\mathcal{C}^{k+1}$ sur $[0, 2\pi]$ (en fait que f admette une dérivée $(k+1)$-ième qui soit Riemann-intégrable). Par contre, il est indispensable de supposer que f est de classe $\mathcal{C}^{k+1}$ sur $\mathbb{R}$ pour que $f^{(k+1)}$ soit une fonction périodique définie sur $\mathbb{R}$.

Il résulte de la proposition 6.8 que plus une fonction est régulière et plus les suites $(a_n)_n$ et $(b_n)_n$ de ses coefficients de Fourier tendent rapidement vers 0. Si f est une application 2π-périodique de classe $\mathcal{C}^k$ sur $\mathbb{R}$, la proposition 6.7 indique que les coefficients de Fourier de l'application $f^{(k)}$ tendent vers 0, ce qui compte tenu de la proposition 6.8 implique que $a_n = o_{+\infty}(n^{-k})$ et que $b_n = o_{+\infty}(n^{-k})$.

6.2.3 Convergence d'une série de Fourier

Proposition 6.9 *Soit f une application 2π-périodique, de classe $\mathcal{C}^2$ sur $\mathbb{R}$. La série de Fourier de f converge normalement sur $\mathbb{R}$.*

Démonstration En appliquant les relations de la proposition 6.8 pour $n = 2$, on obtient :

$$a_n(f) = -\frac{1}{n^2}\, a_n(f'') \quad \text{et} \quad b_n(f) = -\frac{1}{n^2}\, b_n(f'').$$

Puisque f est de classe $\mathcal{C}^2$ sur $\mathbb{R}$, f'' est une application continue sur $[0, 2\pi]$. Par conséquent, sur l'intervalle fermé borné $[0, 2\pi]$, l'application f'' est bornée[19] et atteint ses bornes. En désignant par M la borne supérieure de $|f''|$ sur $[0, 2\pi]$, on obtient

$$|a_n(f'')| = \left|\frac{1}{\pi}\int_0^{2\pi} f''(x)\cos(nx)\,\mathrm{d}x\right| \leqslant \frac{1}{\pi}\int_0^{2\pi} \underbrace{|f''(x)|}_{\leqslant M}\,\underbrace{|\cos(nx)|}_{\leqslant 1}\,\mathrm{d}x \leqslant 2M.$$

Par un calcul analogue, on obtient également $|b_n(f'')| \leqslant 2M$. On en conclut que pour tout $n \in \mathbb{N}^*$,

$$|a_n(f)| \leqslant \frac{2M}{n^2} \quad \text{et} \quad |b_n(f)| \leqslant \frac{2M}{n^2}.$$

Les critères de comparaison pour les séries numériques[20] permettent d'affirmer que les séries $\sum_n a_n(f)$ et $\sum_n b_n(f)$ convergent absolument. La proposition 6.1 permet de conclure que la série de Fourier de f convergence normalement sur $\mathbb{R}$. □

Définition 6.3 (Dérivabilité par morceaux)
Une application f définie sur un intervalle $[a, b]$ et à valeurs réelles est dite dérivable par morceaux *sur $[a, b]$ s'il existe une subdivision $\sigma = (x_i)_{i\in\{0,\dots,n\}}$ de $[a, b]$ telle que pour tout $i \in \{1, \dots, n\}$ la restriction de f à l'intervalle ouvert $]x_{i-1}, x_i[$ coïncide avec la restriction d'une fonction dérivable sur l'intervalle fermé $[x_{i-1}, x_i]$.*

[19] Voir le théorème 13.4 p. 596 du *Cours de première année*.
[20] Voir le théorème 1.2 p. 14.

Exemples

1. La fonction partie entière est dérivable par morceaux sur $[0,5]$ (on peut dans ce cas prendre la subdivision $\sigma = \{0,1,2,3,4,5\}$ de l'intervalle $[0,5]$).

2. Les fonctions $f_1 : x \mapsto \sqrt{|x|}$, $f_2 : x \mapsto |\sin(x)|$ et $f_3 : x \mapsto \sin(x)/|x|$ sont dérivables sur $[-\pi, 0[\cup]0, \pi]$. Elles ne sont pas dérivables en 0. Les applications f_2 et f_3 sont toutefois dérivables par morceaux sur $[-\pi, \pi]$ (prendre $\sigma = \{-\pi, 0, \pi\}$). L'application f_1 n'est pas dérivable par morceaux sur $[-\pi, \pi]$, voir la figure 3.

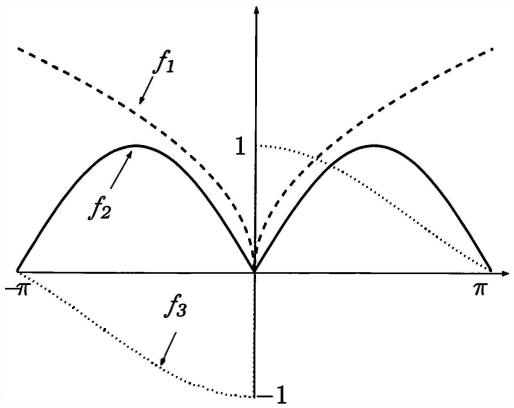


Fig. 3 Exemple d'applications dérivables par morceaux (f_2 et f_3) et d'une application qui ne l'est pas (f_1).

On définit de la même manière la notion d'application continue par morceaux (resp. de classe $\mathcal{C}^1$ par morceaux). Une application f de $[a,b]$ dans $\mathbb{R}$ est dite *continue par morceaux* (resp. *de classe $\mathcal{C}^1$ par morceaux*) sur $[a,b]$ s'il existe une subdivision $\sigma = (x_i)_{i\in\{0,\ldots,n\}}$ de $[a,b]$ telle que pour tout $i \in \{1,\ldots,n\}$ la restriction de f à l'intervalle ouvert $]x_{i-1}, x_i[$ coïncide avec la restriction d'une fonction continue (resp. de classe $\mathcal{C}^1$) sur l'intervalle fermé $[x_{i-1}, x_i]$.

Il résulte de manière directe de ces définitions que toute application continue (resp. dérivable, resp. de classe $\mathcal{C}^1$) sur $[a,b]$ est continue (resp. dérivable, resp. de classe $\mathcal{C}^1$) par morceaux sur $[a,b]$ (considérer la subdivision $\sigma = \{a,b\}$ de $[a,b]$). Par contre, une application continue (resp. dérivable, resp. de classe $\mathcal{C}^1$) sur un intervalle ouvert $]a,b[$ n'est pas nécessairement continue (resp. dérivable, resp. de classe $\mathcal{C}^1$) par morceaux sur $[a,b]$ (considérer par exemple la fonction logarithme sur $]0,1[$). On notera également que toute fonction dérivable sur un intervalle étant continue sur cet intervalle, une application dérivable par morceaux sur un intervalle $[a,b]$ est continue par morceaux sur cet intervalle. De même, une application de classe $\mathcal{C}^1$ par morceaux sur $[a,b]$ est dérivable par morceaux et par conséquent continue par morceaux sur $[a,b]$.

Sur l'exemple présenté figure 3, les applications f_1 et f_2 sont continues sur $[-\pi,\pi]$ et l'application f_3 est continue par morceaux sur $[-\pi,\pi]$. L'application f_2 est de classe $\mathcal{C}^1$ par morceaux $[-\pi,\pi]$ car sur $[-\pi,0[$ (resp. sur $]0,\pi]$)

elle est la restriction de la fonction $x \mapsto -\sin(x)$ (resp. de la fonction sinus) qui est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$ donc qui est de classe $\mathcal{C}^1$ sur $[-\pi, 0]$ (resp. sur $[0, \pi]$). De même l'application f_3 est de classe $\mathcal{C}^1$ par morceaux $[-\pi, \pi]$ car sur $[-\pi, 0[$ elle est la restriction de la fonction

$$\phi : x \in \mathbb{R} \longmapsto \begin{cases} -\sin(x)/x & \text{si } x \neq 0 \\ -1 & \text{sinon} \end{cases}$$

qui est de classe $\mathcal{C}^1$ sur $\mathbb{R}$ donc sur $[-\pi, 0]$. Sur $]0, \pi]$, f_3 est la restriction de la fonction $-\phi$ qui est de classe $\mathcal{C}^1$ sur $[0, \pi]$.

Une application continue (resp. dérivable) par morceaux sur un intervalle $[a, b]$ admet une limite (resp. une dérivée) à gauche et à droite en chaque nœud de la subdivision $\sigma = (x_i)_{i \in \{0,\dots,n\}}$ de $[a, b]$ qui lui est associée.

THÉORÈME 6.1 (Théorème de Dirichlet)
Si f est une application 2π-périodique, dérivable par morceaux sur $[0, 2\pi]$ alors pour tout $x_0 \in \mathbb{R}$, la série de Fourier de f converge. Sa somme en x_0 :

- *vaut $f(x_0)$ si x_0 est un point où f est continue ;*
- *vaut $\frac{1}{2}\Big(\lim\limits_{x \to x_0^+} f(x) + \lim\limits_{x \to x_0^-} f(x)\Big)$ si x_0 est un point où f est discontinue.*

Démonstration Pour tout réel x_0, notons $y_0^+ = \lim_{x \to x_0^+} f(x)$ la limite de f à gauche en x_0, $y_0^- = \lim_{x \to x_0^-} f(x)$ la limite de f à droite en x_0 et $y_0 = \frac{1}{2}(y_0^+ + y_0^-)$. Notons que si x_0 est un point où f est continue alors $y_0 = f(x_0)$. Pour montrer que pour tout $x_0 \in \mathbb{R}$, la série de Fourier de f en x_0 converge et a pour somme y_0, nous allons montrer que la suite $(S_n(x_0))_n$ des sommes partielles associée à la série de Fourier de f converge et a pour limite y_0.

D'après la formule de Dirichlet, voir le lemme 6.2, la somme partielle d'ordre n en x_0 est donnée par :

$$\begin{aligned} S_n(x_0) &= \sum_{k=0}^{n} \big(a_k \cos(kx_0) + b_k \sin(kx_0)\big) \\ &= \frac{1}{2\pi} \int_0^\pi K_n(t) \big(f(x_0 + t) + f(x_0 - t)\big)\, \mathrm{d}t, \end{aligned}$$

où K_n est l'application 2π-périodique définie sur $[0, 2\pi[$ par

$$K_n(x) = \frac{\sin(\frac{1}{2}(2n+1)x)}{2\sin(\frac{1}{2}x)} \quad \text{si } x \neq 0 \quad \text{et} \quad K_n(0) = \frac{1}{2}(2n+1).$$

Puisque $\int_0^\pi K_n(t)\, \mathrm{d}t = \pi$ (voir l'exercice 6.2) et que $y_0 = \frac{1}{2\pi} \int_0^\pi 2y_0\, \mathrm{d}t$, on a

$$\begin{aligned} S_n(x_0) - y_0 &= \frac{1}{2\pi} \int_0^\pi \big(f(x_0 + t) + f(x_0 - t) - 2y_0\big) K_n(t)\, \mathrm{d}t \\ &= \frac{1}{2\pi} \int_0^\pi \frac{f(x_0 + t) - y_0^+ + f(x_0 - t) - y_0^-}{2\sin(\frac{1}{2}t)} \sin(\tfrac{1}{2}(2n+1)t)\, \mathrm{d}t. \end{aligned}$$

Considérons l'application g de $]0,\pi]$ dans $\mathbb{R}$ définie par

$$g(t) = \frac{f(x_0+t) - y_0^+ + f(x_0-t) - y_0^-}{2\sin(\frac{1}{2}t)}.$$

Comme par hypothèse f est dérivable par morceaux, les deux limites suivantes existent :

$$\alpha = \lim_{t\to 0^+} \frac{f(x_0+t)-y_0^+}{t} \quad \text{et} \quad \beta = \lim_{t\to 0^+} \frac{f(x_0-t)-y_0^-}{t}.$$

De plus, comme $2\sin(\frac{1}{2}t) \underset{0}{\sim} t$, on a

$$\begin{aligned}\lim_{t\to 0} \frac{f(x_0+t) - y_0^+ + f(x_0-t) - y_0^-}{2\sin(\frac{1}{2}t)} &= \lim_{t\to 0} \frac{f(x_0+t) - y_0^+ + f(x_0-t) - y_0^-}{t} \\ &= \alpha + \beta.\end{aligned}$$

Ainsi l'application g est prolongeable par continuité en 0 en posant $g(0) = \alpha+\beta$. Comme f est 2π-périodique et dérivable par morceaux sur $[0,2\pi]$, l'application g est continue par morceaux sur $[0,\pi]$ et par conséquent[21] Riemann-intégrable sur $[0,\pi]$. On déduit du Lemme de Lebesgue (voir p. 317) que

$$\lim_{n\to+\infty} S_n(x_0) - y_0 = \frac{1}{2\pi} \lim_{n\to+\infty} \int_0^\pi g(t)\sin((2n+1)\tfrac{1}{2}t)\,\mathrm{d}t = 0.$$

Ainsi, la suite $(S_n(x_0))_n$ converge et a pour limite y_0 ; le théorème de Dirichlet est démontré. □

Remarque Il existe de nombreux autres résultats de convergence concernant les séries de Fourier. Mentionnons par exemple le résultat suivant (théorème de Jordan[22]) : si f est une application 2π-périodique, monotone par morceaux[23] sur $[0,2\pi]$ alors la série de Fourier de f converge en $x_0 \in \mathbb{R}$ et a pour somme le réel $\frac{1}{2}\big(\lim_{x\to x_0^+} f(x) + \lim_{x\to x_0^-} f(x)\big)$. ◇

Corollaire 6.2 *Soit f une application 2π-périodique, de classe $\mathcal{C}^2$ sur $\mathbb{R}$. La série de Fourier de f converge normalement sur $\mathbb{R}$ et a pour somme f.*

Démonstration D'après la proposition 6.9, la série de Fourier de f converge normalement sur $\mathbb{R}$ et d'après le théorème de Dirichlet puisque f est continue la série de Fourier de f converge simplement sur $\mathbb{R}$ et a pour somme f. □

---

[21] Voir la proposition 18.3 p. 843 du *Cours de première année*.

[22] Jordan, Camille (1838, Lyon - 1922, Paris).

[23] Une application f de $[a,b]$ dans $\mathbb{R}$ est dite *monotone par morceaux* sur $[a,b]$ s'il existe une subdivision $\sigma = (x_i)_{i\in\{0,\dots,n\}}$ de $[a,b]$ telle que pour tout $i \in \{1,\dots,n\}$ la restriction de f à l'intervalle ouvert $]x_{i-1},x_i[$ coïncide avec la restriction d'une fonction monotone sur l'intervalle fermé $[x_{i-1},x_i]$.

Remarque On peut montrer que la conclusion du corollaire 6.2 est vraie avec des hypothèses plus faibles. Il suffit que f soit une application 2π-périodique, continue sur $\mathbb{R}$ et de classe C^1 par morceaux sur $\mathbb{R}$ pour que l'on puisse conclure que la série de Fourier de f converge normalement sur $\mathbb{R}$ et ait pour somme f. ◇

Le résultat suivant, que nous admettrons[24], est d'une grande importance en traitement du signal. La formule de Parseval indique que l'énergie totale d'un signal périodique s'obtient en sommant l'énergie des différentes « harmoniques » du signal.

PROPOSITION 6.10 (Égalité de Parseval[25])
Soit f une application 2π-périodique, Riemann-intégrable sur $[0, 2\pi]$, de coefficients de Fourier $(a_n)_n$ et $(b_n)_n$. Les séries numériques $\sum_n a_n^2$ et $\sum_n b_n^2$ convergent et on a

$$\frac{1}{\pi}\int_0^{2\pi} f(x)^2 \, \mathrm{d}x = 2a_0^2 + \sum_{n=1}^{+\infty} \left(a_n^2 + b_n^2\right).$$

Remarque L'égalité de Parseval permet d'affirmer qu'il existe des séries trigonométriques qui ne sont série de Fourier d'aucune fonction 2π-périodique et intégrable. C'est le cas par exemple de la série trigonométrique $\sum_n \frac{1}{\sqrt{n}} \sin(nx)$. D'après la règle d'Abel pour les séries trigonométriques, voir la proposition 6.2 p. 302, cette série converge simplement sur $\mathbb{R} \setminus 2\pi\mathbb{Z}$. Comme par ailleurs pour $x \in 2\pi\mathbb{Z}$ cette série est la série nulle, elle converge simplement sur $\mathbb{R}$. Si cette série trigonométrique était la série de Fourier d'une fonction f 2π-périodique et Riemann-intégrable sur $[0, 2\pi]$, alors on aurait :

$$\frac{1}{\pi}\int_0^{2\pi} f(x)^2 \, \mathrm{d}x = \sum_{n=1}^{+\infty} \frac{1}{n}.$$

Cette égalité conduit à une absurdité car comme f est supposée Riemann-intégrable sur $[0, 2\pi]$, l'intégrale $\int_0^{2\pi} f(x)^2 \, \mathrm{d}x$ est bien définie alors que la série $\sum_{n \geqslant 1} \frac{1}{n}$ diverge. ◇

COROLLAIRE 6.3 *L'unique application continue et 2π-périodique qui a tous ses coefficients de Fourier nuls est l'application nulle.*

[24] Pour une démonstration, on pourra consulter : J. LELONG-FERRAND, J.M. ARNAUDIÈS, Cours de Mathématiques, t. 2, p. 555.

[25] DE PARSEVAL DES CHÊSNES, Marc-Antoine (1755, Rosières-aux-Salines -1836, Paris).

Démonstration Il est clair que les coefficients de Fourier de l'application nulle sont tous nuls. Soit f une application 2π-périodique et Riemann-intégrable sur $[0, 2\pi]$ de coefficients de Fourier tous nuls. D'après l'égalité de Parseval, on a

$$\frac{1}{\pi}\int_0^{2\pi} f(x)^2 \, \mathrm{d}x = 0.$$

Comme f est continue sur $[0, 2\pi]$, l'application f^2 est elle aussi continue sur $[0, 2\pi]$. On en déduit[26] que f^2 est l'application nulle ce qui implique que f est également l'application nulle. □

On déduit du corollaire 6.3 qu'une fonction continue est entièrement déterminée par la donnée de ses coefficients de Fourier. En d'autres termes, on a le résultat suivant.

Corollaire 6.4 *Si deux applications continues sur $\mathbb{R}$ et 2π-périodiques ont les même coefficients de Fourier alors ces deux applications sont égales.*

Démonstration Si f et g sont deux applications continues sur $\mathbb{R}$ et 2π-périodiques ayant mêmes coefficients de Fourier alors l'application $h = f - g$ est une application continue et 2π-périodique qui a tous ses coefficients de Fourier nuls. D'après le corollaire 6.3, h est l'application nulle. On a donc $f = g$. □

6.2.4 Quelques exemples

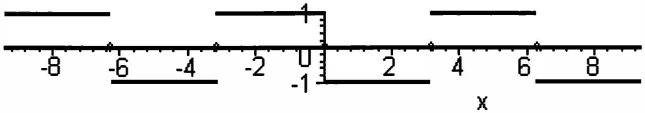
En mathématiques, les séries de Fourier constituent un moyen de déterminer la somme de certaines séries numériques. À titre d'illustration, nous allons voir comment nous pouvons obtenir la somme de la série numérique $\sum_{n\geqslant 1} \frac{1}{n^2}$. On comparera avec la méthode utilisée dans le chap. 1 (voir l'exercice 1.15 p. 53) pour déterminer cette somme.

Un premier exemple

Considérons l'application 2π-périodique f définie sur $[0, 2\pi[$ par :

$f(x) = 1$ pour $x \in]0, \pi[$, $f(x) = -1$ pour $x \in]\pi, 2\pi[$ et $f(0) = f(\pi) = 0$.

```
> f:=(n,x)->piecewise(x>0+2*n*Pi and x<Pi + 2*n*Pi, 1,
                      x>Pi+2*n*Pi and x<2*Pi+2*n*Pi ,-1,0):
> N:=2:
> plot([seq(f(n,x),n=-N..N)],x=-2*N*Pi..2*N*Pi,discont=true);
```



[26] Voir la proposition 18.5 p. 844 du *Cours de première année*.

L'application f étant impaire[27], pour tout $n \in \mathbb{N}$ on a $a_n = 0$ et

$$b_n = \frac{2}{\pi}\int_0^\pi f(x)\,\sin(nx)\,\mathrm{d}x = \frac{2}{\pi}\int_0^\pi \sin(nx)\,\mathrm{d}x = \frac{2}{n\pi}(1-\cos(n\pi)).$$

On a donc $b_n = 0$ si n est pair et $b_n = \frac{4}{n\pi}$ si n est impair. La série de Fourier de f est donc

$$\mathcal{S}_F(f)(x) = \sum_{n=0}^{+\infty} \frac{4}{(2n+1)\pi}\,\sin((2n+1)x).$$

Puisque f est dérivable par morceaux sur $\mathbb{R}$ et continue sur $]0,\pi[\cup]\pi,2\pi[$, d'après le théorème de Dirichlet, la série de Fourier de f converge vers $f(x)$ pour tout $x \in \mathbb{R}\setminus\pi\mathbb{Z}$, *i.e.*

$$\forall x \in \mathbb{R}\setminus\pi\mathbb{Z} \qquad f(x) = \sum_{n=0}^{+\infty} \frac{4}{(2n+1)\pi}\,\sin((2n+1)x).$$

À l'origine, l'application f est discontinue et $\lim_{x\to 0^+} f(x) + \lim_{x\to 0^-} f(x) = 0$; d'après le théorème de Dirichlet la série de Fourier de f converge en 0 vers 0. On a la même chose en π, puis par périodicité en tout $x \in \pi\mathbb{Z}$.

Comme $\sin((2n+1)\frac{\pi}{2}) = \sin(\frac{\pi}{2}+n\pi) = (-1)^n$, on obtient en choisissant $x = \pi/2$ dans la relation précédente :

$$f(\pi/2) = \frac{4}{\pi}\sum_{n=0}^{+\infty}\frac{(-1)^n}{(2n+1)},$$

et puisque $f(\pi/2) = 1$, on a établi la formule :

$$\boxed{\sum_{n=0}^{+\infty}\frac{(-1)^n}{(2n+1)} = \frac{\pi}{4}}$$

L'égalité de Parseval nous indique par ailleurs que la série $\sum_n \frac{1}{(2n+1)^2}$ converge et que

$$\frac{16}{\pi^2}\sum_{n=0}^{+\infty}\frac{1}{(2n+1)^2} = \frac{1}{\pi}\int_0^\pi 1^2\,\mathrm{d}x + \frac{1}{\pi}\int_\pi^{2\pi}(-1)^2\,\mathrm{d}x = 2.$$

On en déduit que :

$$\boxed{\sum_{n=0}^{+\infty}\frac{1}{(2n+1)^2} = \frac{\pi^2}{8}}$$

[27] Voir le corollaire 6.1 p. 315.

Soit $(S_n)_n$ la suite des sommes partielles associée à la série $\sum_{n\geqslant 1}\frac{1}{n^2}$ qui est une série de Riemann convergente et $(T_n)_n$ la suite des sommes partielles associée à la série $\sum_n \frac{1}{(2n+1)^2}$. Pour tout $n\in\mathbb{N}^*$, on a

$$S_{2n+1}=\sum_{k=1}^{2n+1}\frac{1}{k^2}=\sum_{p=1}^{n}\frac{1}{(2p)^2}+\sum_{p=0}^{n}\frac{1}{(2p+1)^2}=\frac{1}{4}S_n+T_n.$$

Comme $\sum_n \frac{1}{(2n+1)^2}$ a pour somme $\frac{1}{8}\pi^2$, la suite $(T_n)_n$ converge vers $\frac{1}{8}\pi^2$. Puisque la suite de terme général S_{2n+1} est extraite de la suite $(S_n)_n$ qui converge vers S, ces deux suites ont la même limite. On déduit de la relation précédente que :

$$S=\frac{1}{4}S+\frac{\pi^2}{8},$$

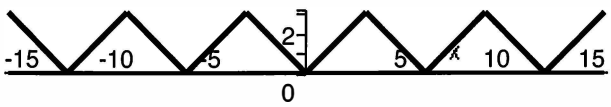
ce qui donne $\frac{3}{4}S=\frac{1}{8}\pi^2$; on a par conséquent :

$$\boxed{\sum_{n=1}^{+\infty}\frac{1}{n^2}=\frac{\pi^2}{6}}$$

Un second exemple

Soit f l'application 2π-périodique et paire définie sur $[0,\pi]$ par $f(x)=x$.

```
> f:=(n,x)->piecewise(x>-Pi+2*n*Pi and x<2*n*Pi, -x+2*n*Pi,
                      x>2*n*Pi and x<Pi+2*n*Pi ,x-2*n*Pi,0):
> N:=2:
> plot([seq(f(n,x),n=-N..N)],x=-Pi-2*N*Pi..Pi+2*N*Pi,discont=true);
```



La parité de f implique que pour tout $n\in\mathbb{N}$ on a $b_n=0$. De plus,

$$a_0=\frac{1}{\pi}\int_0^{\pi}x\,\mathrm{d}x=\frac{\pi}{2}$$

et pour $n\in\mathbb{N}^*$

$$a_n=\frac{2}{\pi}\int_0^{\pi}f(x)\cos(nx)\,\mathrm{d}x=\frac{2}{\pi}\int_0^{\pi}x\cos(nx)\,\mathrm{d}x.$$

En intégrant par parties, on obtient :

$$a_n=\frac{2}{n\pi}\Big[x\sin(nx)\Big]_0^{\pi}-\frac{2}{n\pi}\int_0^{\pi}\sin(nx)\,\mathrm{d}x=\frac{2}{\pi n^2}(\cos(n\pi)-1).$$

On a donc $a_n = 0$ si n est pair et non nul et $a_n = -\frac{4}{n^2\pi}$ si n est impair. La série de Fourier de f est donc

$$\mathcal{S}_F(f)(x) = \frac{\pi}{2} - \sum_{k\in\mathbb{N}} \frac{4}{(2k+1)^2\pi}\cos((2k+1)x).$$

Puisque f est continue et dérivable par morceaux sur $\mathbb{R}$, d'après le théorème de Dirichlet la série de Fourier de f converge vers $f(x)$ pour tout $x \in \mathbb{R}$, *i.e.*

$$\forall x \in \mathbb{R} \qquad f(x) = \frac{\pi}{2} - \sum_{n=0}^{+\infty} \frac{4}{(2n+1)\pi}\cos((2n+1)x).$$

En prenant $x = 0$, on obtient à nouveau la relation obtenue dans l'exemple précédent :

$$\sum_{n=0}^{+\infty} \frac{1}{(2n+1)^2} = \frac{\pi^2}{8}.$$

Par ailleurs, l'égalité de Parseval nous donne :

$$\frac{\pi^2}{2} + \frac{16}{\pi^2}\sum_{n=0}^{+\infty} \frac{1}{(2n+1)^4} = \frac{1}{\pi}\int_0^{2\pi} f(x)^2\,\mathrm{d}x = \frac{2}{\pi}\int_0^{\pi} x^2\,\mathrm{d}x = \frac{2}{3}\pi^2.$$

On en déduit la relation :

$$\boxed{\sum_{n=0}^{+\infty} \frac{1}{(2n+1)^4} = \frac{\pi^4}{96}}$$

Reprenant l'approche utilisée dans l'exemple précédent, désignons par $(S_n)_n$ la suite des sommes partielles associée à la série $\sum_{n\geqslant 1}\frac{1}{n^4}$ qui est une série de Riemann convergente et par $(T_n)_n$ la suite des sommes partielles associée à la série $\sum_n \frac{1}{(2n+1)^4}$. Pour tout $n \in \mathbb{N}^*$, on a

$$S_{2n+1} = \sum_{k=0}^{2n+1} \frac{1}{k^4} = \sum_{p=0}^{n} \frac{1}{(2p)^4} + \sum_{p=0}^{n} \frac{1}{(2p+1)^4} = \frac{1}{16}S_n + T_n.$$

Comme $\sum_n \frac{1}{(2n+1)^2}$ a pour somme $\frac{1}{96}\pi^4$, la suite $(T_n)_n$ converge vers $\frac{1}{96}\pi^4$. Par ailleurs, puisque la suite de terme général S_{2n+1} est extraite de la suite $(S_n)_n$ qui converge vers la somme S de la série $\sum_{n\geqslant 1}\frac{1}{n^4}$, ces deux suites ont la même limite. On en déduit que $S = \frac{1}{16}S + \frac{1}{96}\pi^4$ ce qui donne $\frac{15}{16}S = \frac{1}{96}\pi^4$. On en conclut que

$$\boxed{\sum_{n=1}^{+\infty} \frac{1}{n^4} = \frac{\pi^4}{90}}$$

On pourra consulter l'exercice 6.4 pour une autre manière d'obtenir la valeur des sommes $\sum_{n=1}^{+\infty} 1/n^p$ lorsque p est un entier pair.

Remarque On sait très peu de chose sur les valeurs de $\sum_{n=1}^{+\infty} \frac{1}{n^p}$ lorsque p est un entier impair. Ainsi, il a fallu attendre 1979 pour que soit démontré par Roger Apéry[28] le fait que $\sum_{n=1}^{+\infty} 1/n^3$ est un nombre irrationnel. ◇

6.3 Exercices de synthèse

Exercice 6.4 ***Calcul de valeurs de la fonction ζ de Riemann***

On s'intéresse au calcul de la valeur prise par la fonction ζ de Riemann $\zeta : x \mapsto \sum_{n=1}^{+\infty} n^{-x}$ en $x = 2k$ où $k \in \mathbb{N}^*$. On pourra consulter l'exercice 2.10 p. 96 pour l'étude de quelques propriétés de la fonctions ζ de Riemann.

1 - Pour $t \in \mathbb{R} \setminus \mathbb{Z}$ fixé, on considère l'application 2π-périodique f définie par $f(x) = \cos(tx)$ pour tout $x \in]-\pi, \pi]$.

a) Calculer les coefficients de Fourier de f et donner l'expression de la série de Fourier de f.

b) Indiquer sur quel ensemble la série de Fourier de f converge vers f.

c) En déduire que pour tout $t \in \mathbb{R} \setminus \mathbb{Z}$ on a[29]

$$\operatorname{cotan}(\pi t) = \frac{1}{\pi}\left(\frac{1}{t} - \sum_{n=1}^{+\infty} \frac{2t}{n^2 - t^2}\right)$$

puis que pour tout $x \in \mathbb{R} \setminus \pi\mathbb{Z}$ on a : $x \operatorname{cotan}(x) = 1 - \sum_{n=1}^{+\infty} \frac{2x^2}{n^2\pi^2 - x^2}$.

2 - a) Pour $n \in \mathbb{N}^*$ fixé, on définit l'application $g_n : x \in]-\pi, \pi[\mapsto \frac{x^2}{n^2\pi^2 - x^2}$. Montrer que le développement en série entière de g_n dans un voisinage de l'origine est $\sum_{k=1}^{+\infty} (n\pi)^{-2k} x^{2k}$ et indiquer le rayon de convergence de cette série entière.

b) Pour $N \in \mathbb{N}^*$;, on définit $S_N : x \in]-\pi, \pi[\mapsto \sum_{n=1}^{N} \frac{2x^2}{n^2\pi^2 - x^2}$. Montrer que le développement en série entière de S_N dans un voisinage de l'origine que l'on précisera est $2\sum_{k=1}^{+\infty} \frac{1}{\pi^{2k}} \zeta_N(2k)\, x^{2k}$ où $\zeta_N : x \mapsto \sum_{n=1}^{N} n^{-x}$.

3 - a) Soit $T : x \in]-\pi, \pi[\mapsto 2\sum_{n=1}^{+\infty} \zeta(2n) \frac{x^{2n}}{\pi^{2n}}$; montrer que pour tout $x \in]-\pi, \pi[$ on a $\lim_{N \to +\infty} |T(x) - S_N(x)| = 0$.

b) En déduire que : $x \operatorname{cotan}(x) = 1 - 2\sum_{n=1}^{+\infty} \zeta(2n) \frac{x^{2n}}{\pi^{2n}}$ pour tout $x \in]-\pi, \pi[$.

c) On a effectué le calcul suivant avec MAPLE :

```
> taylor(x*cot(x),x=0,9);
```

$$1 - \frac{1}{3}x^2 - \frac{1}{45}x^4 - \frac{2}{945}x^6 - \frac{1}{4725}x^8 + \mathcal{O}_0(x^{10})$$

Déduire, en le justifiant, les valeurs de $\zeta(2)$, $\zeta(4)$, $\zeta(6)$ et $\zeta(8)$.

---

[28] APÉRY, Roger (1916, Rouen - 1994, Caen). Mathématicien français.

[29] cotan désigne la fonction cotangente définie par $\operatorname{cotan}(x) = \cos(x)/\sin(x)$ pour tout $x \in \mathbb{R} \setminus \pi\mathbb{Z}$ où $\mathbb{R} \setminus \pi\mathbb{Z} = \{x \in \mathbb{R} \mid \forall k \in \mathbb{Z} \;\; x \neq k\pi\}$.

Exercice 6.5 ***Phénomène de Gibbs***

Au voisinage d'un point où une fonction f présente une discontinuité, les valeurs prises par les sommes partielles S_n de la série de Fourier de f oscillent autour de la valeur vers laquelle elles doivent converger et l'augmentation de l'ordre n de la somme partielle considérée ne diminue pas l'amplitude de l'oscillation contrairement à ce que l'intuition pourrait suggérer. Au contraire, l'écart reste toujours supérieur à une valeur seuil, aussi grand soit l'ordre de la somme partielle. Toutefois, l'oscillation s'effectue sur des intervalles de plus en plus petits. Ce phénomène, appelé « phénomène de Gibbs », fait l'objet de l'étude effectuée dans cet exercice. Il fut pour la première fois expliqué du point de vue mathématique par J.W. Gibbs[30] *en 1899. Ce phénomène reflète les difficultés qu'il y a à approcher une fonction discontinue par une suite de polynômes trigonométriques (qui sont des fonctions continues). Ce phénomène est également lié au fait que la décroissance quand n tend vers l'infini des coefficients de Fourier d'une fonction est contrôlée par la régularité de cette fonction. Les fonctions très régulières ont des coefficients de Fourier qui décroissent très rapidement quand n augmente, tandis que les fonctions discontinues ont des coefficients de Fourier qui décroissent très lentement.*

Pour répondre aux questions qui suivent, on aura recours à MAPLE *à chaque fois que cela sera possible, en particulier pour effectuer tous les calculs longs et fastidieux. Toutefois, les résultats proposés devront être en totalité justifiés et basés sur un raisonnement rigoureux.*

On considère la fonction 2π-périodique et impaire définie par $f(0) = 0$ et par $f(x) = \frac{1}{2}(\pi - x)$ pour $x \in]0, \pi[$.

1 - a) Déterminer la série de Fourier de f.

b) Représenter graphiquement sur l'intervalle $[-2\pi, 2\pi]$ les sommes partielles de la série de Fourier de f correspondant aux rangs 6, 12 et 24 ainsi que f. Qu'observe-t-on ?

On considère la suite de fonctions $(S_n)_n$ définie pour $n \in \mathbb{N}^*$ par :

$$S_n : x \in [-\pi, \pi] \longmapsto \sum_{k=1}^{n} \frac{\sin(kx)}{k}.$$

2 - a) Vers quelle fonction converge la suite de fonctions $(S_n)_n$? Quel type de convergence a-t-on ?

b) Montrer que pour tout $x \neq 0$ on a

$$S_n'(x) = \frac{\sin(nx/2)\cos((n+1)x/2)}{\sin(x/2)}.$$

---

[30] Gibbs, Josiah Willard (1839, New Haven - 1903, New Haven). Physicien américain, J.W. Gibbs se distingua par ses importants travaux en thermodynamique et en mécanique statistique. En particulier, il appliqua les principes de la thermodynamique à l'étude des phénomènes physico-chimiques, introduisant l'emploi de l'entropie dans les diagrammes d'état.

Que vaut $S_n'(0)$?

c) Résoudre l'équation $S_n'(x) = 0$ sur $[0, \pi]$.

Pour $k \in \mathbb{N}$, on note $u_k = 2k\pi/n$, $v_k = (2k+1)\pi/(n+1)$ et $t_k = (u_k + v_k)/2$, $s_k = (v_k + u_{k+1})/2$.

d) Déterminer le signe de $S_n'(t_k)$ et de $S_n'(s_k)$ pour tout $k \in \mathbb{N}$ avec $k < n/2$.

e) Donner le tableau de variation de S_n.

L'entier n étant fixé, on notera que le tableau de variation indique que les maxima relatifs de $S_n(x)$ sur $[0, \pi]$ sont les réels $m_k^{(n)} = S_n(v_k)$ où $k \in \mathbb{N}$ est tel que $k \leqslant n/2$.

3 - a) Montrer que pour tout $k \in \mathbb{N}^*$ tel que $k \leqslant n/2$ on a

$$m_k^{(n)} - m_{k-1}^{(n)} = \int_{v_{k-1}}^{v_k} \frac{\sin(\frac{1}{2}nt)\cos(\frac{1}{2}(n+1)t)}{\sin(\frac{1}{2}t)}\,\mathrm{d}t.$$

b) En déduire que pour tout $k \in \mathbb{N}^*$ tel que $k \leqslant n/2$ on a

$$m_k^{(n)} - m_{k-1}^{(n)} \leqslant \frac{1}{\sin(\frac{1}{2}v_{k-1})} \int_{v_{k-1}}^{v_k} \sin\left(\tfrac{1}{2}nt\right)\cos\left(\tfrac{1}{2}(n+1)t\right)\,\mathrm{d}t.$$

c) Montrer que

$$\int_{v_{k-1}}^{v_k} \sin\left(\tfrac{1}{2}nt\right)\cos\left(\tfrac{1}{2}(n+1)t\right)\,\mathrm{d}t = 2\left(\frac{(-1)^{2k+1}}{2n+1} - 1\right)\sin\left(\frac{k}{n+1}\pi\right)\sin\left(\frac{1}{2(n+1)}\pi\right).$$

d) En déduire que la suite $(m_k^{(n)})_{k\in\{0,\ldots,\mathrm{E}(n/2)\}}$ est strictement décroissante.

4 - On considère la suite $(\mu_n)_n$ de terme général $\mu_n = \sup_{x\in[0,\pi]} S_n(x)$.

a) Montrer que $\mu_n = m_0^{(n)}$.

b) En déduire que la suite $(\mu_n)_n$ tend vers $I = \displaystyle\int_0^\pi \frac{\sin(t)}{t}\,\mathrm{d}t$.

(Indication : on pourra interpréter μ_n comme une somme de Riemann[31].)

c) Calculer une valeur approchée de I.

5 - Comparer les valeurs de

$$\sup_{x\in[0,\pi]}\left(\lim_{n\to+\infty} S_n(x)\right) \quad \text{et} \quad \lim_{n\to+\infty}\left(\sup_{x\in[0,\pi]} S_n(x)\right).$$

et interpréter les résultats obtenus relativement au phénomène de Gibbs.

---

[31] Voir la définition 18.8 p. 868 du *Cours de première année*.

6.4 Solutions des exercices

Solution de l'exercice 6.1

En effectuant le changement de variable $y = x/2$, on obtient

$$a_0 = \frac{1}{2\pi}\int_0^{2\pi} |\sin(x/2)|\ \mathrm{d}x = \frac{1}{\pi}\int_0^{\pi} |\sin(y)|\ \mathrm{d}y = \frac{1}{\pi}\int_0^{\pi} \sin(y)\ \mathrm{d}y = \frac{2}{\pi}.$$

De la même manière, pour $n \in \mathbb{N}^*$ on a

$$a_n = \frac{1}{\pi}\int_0^{2\pi} |\sin(x/2)|\cos(nx)\ \mathrm{d}x = \frac{2}{\pi}\int_0^{\pi} \sin(y)\cos(2ny)\ \mathrm{d}y.$$

Comme $\sin(y)\cos(2ny) = \frac{1}{2}\sin(y-2ny) + \frac{1}{2}\sin(y+2ny)$, on obtient :

$$\begin{aligned} a_n &= -\frac{1}{\pi}\int_0^{\pi} \sin((2n-1)y)\ \mathrm{d}y + \frac{1}{\pi}\int_0^{\pi} \sin((2n+1)y)\ \mathrm{d}y \\ &= \frac{1}{\pi}\Big[\frac{1}{2n-1}\cos((2n-1)y)\Big]_0^{\pi} - \frac{1}{\pi}\Big[\frac{1}{2n+1}\cos((2n+1)y)\Big]_0^{\pi}. \end{aligned}$$

Compte tenu du fait que $\cos((2n-1)\pi) = \cos((2n+1)\pi) = -1$, on obtient $a_n = -\dfrac{4}{\pi(4n^2-1)}$. Par ailleurs, pour $n \in \mathbb{N}^*$, on a

$$b_n = \frac{1}{\pi}\int_0^{2\pi} |\sin(x/2)|\sin(nx)\ \mathrm{d}x = \frac{2}{\pi}\int_0^{\pi} \sin(y)\sin(2ny)\ \mathrm{d}y.$$

Comme $\sin(y)\sin(2ny) = \frac{1}{2}\cos(2ny-y) - \frac{1}{2}\sin(2ny+y)$, on obtient

$$\begin{aligned} b_n &= \frac{1}{\pi}\int_0^{\pi} \cos((2n-1)y)\ \mathrm{d}y - \frac{1}{\pi}\int_0^{\pi} \cos((2n+1)y)\ \mathrm{d}y \\ &= \frac{1}{\pi}\Big[\frac{1}{2n-1}\sin((2n-1)y)\Big]_0^{\pi} - \frac{1}{\pi}\Big[\frac{1}{2n+1}\sin((2n+1)y)\Big]_0^{\pi} = 0. \end{aligned}$$

Nous verrons ultérieurement (voir le corollaire 6.1 p. 315) que ce dernier résultat était prévisible puisque la fonction f est paire.

Solution de l'exercice 6.2

Appliquons la formule de Dirichlet à la fonction constante égale à 1 (qui est bien une application 2π-périodique et Riemann-intégrable sur $[0, 2\pi]$) ; pour tout $n \in \mathbb{N}$

$$\int_0^{\pi} \frac{\sin((2n+1)\frac{t}{2})}{\sin(\frac{t}{2})}\ \mathrm{d}t = \pi S_n(0),$$

où S_n désigne le n–ième terme de la suite des sommes partielles associée à la série de Fourier de la fonction constante égale à 1. Calculons les coefficients de Fourier de cette fonction. Pour tout $n \in \mathbb{N}^*$, on a

$$\begin{aligned} a_n &= \frac{1}{\pi}\int_0^{2\pi} \cos(nx)\ \mathrm{d}x = \Big[\frac{1}{\pi n}\sin(nx)\Big]_0^{2\pi} = 0, \\ b_n &= \frac{1}{\pi}\int_0^{2\pi} \sin(nx)\ \mathrm{d}x = \Big[-\frac{1}{\pi n}\cos(nx)\Big]_0^{2\pi} = 0. \end{aligned}$$

De plus, $a_0 = \frac{1}{2\pi}\int_0^{2\pi} 1 \ \mathrm{d}x = 1$. On en déduit que $S_n(x) = a_0 = 1$ ce qui permet d'établir la formule.

Solution de l'exercice 6.3

Pour tout $n \in \mathbb{N}^*$, on obtient en intégrant par parties

$$\begin{aligned}\pi b_n &= \int_0^{2\pi} f(x)\sin(nx)\ \mathrm{d}x \\ &= \Big[-\frac{1}{n} f(x)\ \cos(nx)\Big]_0^{2\pi} + \frac{1}{n}\int_0^{2\pi} f'(x)\cos(nx)\ \mathrm{d}x \\ &= \frac{1}{n}\int_0^{2\pi} f'(x)\cos(nx)\ \mathrm{d}x.\end{aligned}$$

Comme f est de classe $\mathcal{C}^1$ sur $\mathbb{R}$, l'application f' est continue sur l'intervalle fermé et borné $[0, 2\pi]$; par conséquent, l'application f' est bornée[32] et atteint ses bornes sur $[0, 2\pi]$. Notons $M = \sup_{x\in[0,2\pi]} |f'(x)|$; on a

$$|\pi b_n| \leqslant \frac{1}{n}\int_0^{2\pi} \underbrace{|f'(x)|}_{\leqslant M}\ \underbrace{|\cos(nx)|}_{\leqslant 1}\ \mathrm{d}x \leqslant \frac{2\pi}{n} M.$$

Par le théorème d'encadrement, on en déduit que la suite $(b_n)_n$ converge vers 0. On procède de même pour montrer que la suite $(a_n)_n$ converge vers 0.

Solution de l'exercice 6.4

1 - a) Comme f est une fonction paire, les coefficients $b_n = \frac{1}{\pi}\int_0^{2\pi} f(x)\ \sin(nx)\ \mathrm{d}x$ sont nuls. De plus,

$$a_0 = \frac{1}{\pi}\int_0^{\pi} \cos(tx)\ \mathrm{d}x = \Big[\frac{1}{\pi t}\sin(tx)\Big]_0^{\pi} = \frac{1}{\pi t}\ \sin(\pi t)$$

et pour tout $n \in \mathbb{N}^*$ on a

$$\begin{aligned}a_n &= \frac{2}{\pi}\int_0^{\pi} \cos(tx)\ \cos(nx)\ \mathrm{d}x = \frac{1}{\pi}\int_0^{\pi} \big(\cos((n+t)x) + \cos((n-t)x)\big)\ \mathrm{d}t \\ &= \Big[\frac{1}{\pi(n+t)}\ \sin((n+t)x) + \frac{1}{\pi(n-t)}\ \sin((n-t)x)\Big]_0^{\pi} \\ &= \frac{1}{\pi(n+t)}\ \sin((n+t)\pi) + \frac{1}{\pi(n-t)}\ \sin((n-t)\pi).\end{aligned}$$

Or, pour tout $\alpha \in \mathbb{R}$ on a $\sin(\alpha + n\pi) = (-1)^n \sin(\alpha)$, donc

$$a_n = \frac{(-1)^n}{\pi}\ \sin(\pi t)\ \left(\frac{1}{n+t} - \frac{1}{n-t}\right) = \frac{(-1)^n}{\pi}\ \sin(\pi t)\ \frac{2t}{n^2 - t^2}.$$

---

[32] Voir le théorème 13.4 p. 596 du *Cours de première année*.

La série de Fourier de f est donc

$$\mathcal{S}(f)(x) = \frac{1}{\pi t}\sin(\pi t) + \sum_{n=1}^{+\infty} \frac{(-1)^{n+1}}{\pi}\frac{2t}{n^2-t^2}\sin(\pi t)\cos(nx).$$

b) La fonction f est de classe $\mathcal{C}^\infty$ sur $[-\pi,\pi]$ et continue sur $\mathbb{R}$ car $\lim_{x\to-\pi^+} f(x) = \lim_{x\to\pi^-} = \cos(\pi t)$. D'après le théorème de Dirichlet, on en déduit que la série de Fourier de f converge simplement vers f sur $\mathbb{R}$.

c) La série de Fourier de f converge en $x=\pi$ vers $f(\pi)$ donc pour tout $t\in\mathbb{R}\setminus\mathbb{Z}$, on a

$$\begin{aligned}\cos(\pi t) &= \frac{1}{\pi t}\sin(\pi t) + \sum_{n=1}^{+\infty}\frac{(-1)^{n+1}}{\pi}\frac{2t}{n^2-t^2}\sin(\pi t)\underbrace{\cos(n\pi)}_{=(-1)^n}\\ &= \frac{1}{\pi}\sin(\pi t)\left(\frac{1}{t} - \sum_{n=1}^{+\infty}\frac{2t}{n^2-t^2}\right).\end{aligned}$$

Pour tout $t\in\mathbb{R}\setminus\mathbb{Z}$, on a $\sin(\pi t)\neq 0$ donc

$$\operatorname{cotan}(\pi t) = \frac{\cos(\pi t)}{\sin(\pi t)} = \frac{1}{\pi}\left(\frac{1}{t} - \sum_{n=1}^{\infty}\frac{2t}{n^2-t^2}\right)$$

On déduit de la relation précédente que

$$\pi t\operatorname{cotan}(\pi t) = 1 - \sum_{n=1}^{+\infty}\frac{2t^2}{n^2-t^2}.$$

En effectuant le changement de variable $s=\pi t$, pour $t\in\mathbb{R}\backslash\mathbb{Z}$ on a $s\in\mathbb{R}\setminus\pi\mathbb{Z}$ et

$$s\operatorname{cotan}(s) = 1 - \sum_{n=1}^{+\infty}\frac{2s^2}{(n\pi)^2-s^2}.$$

2 - a) Pour tout $x\in[-\pi,0[\cup]0,\pi]$ on a

$$g_n(x) = \frac{x^2}{n^2\pi^2}\,\frac{1}{1-\left(\frac{x}{n\pi}\right)^2}.$$

Quand x est a voisinage de 0, le réel $u=\left(\frac{x}{n\pi}\right)^2$ est au voisinage de 0. De plus, pour tout $u\in]-1,1[$, on a $\frac{1}{1-u} = \sum_{k=0}^{+\infty}u^k$ donc

$$g_n(x) = \frac{x^2}{n^2\pi^2}\sum_{k=0}^{+\infty}\left(\frac{x}{n\pi}\right)^{2k} = \sum_{k=0}^{+\infty}\left(\frac{x}{n\pi}\right)^{2k+2} = \sum_{k=1}^{+\infty}\left(\frac{x}{n\pi}\right)^{2k}.$$

Le développement en série entière de g_n au voisinage de 0 est donc $\sum_{k=1}^{+\infty}\left(\frac{x}{n\pi}\right)^{2k}$ et on a $g_n(x) = \sum_{k=1}^{+\infty}\left(\frac{x}{n\pi}\right)^{2k}$ pour tout $u=\frac{x}{n\pi}$ dans $]-1,1[$ *i.e.* pour tout $x\in]-n\pi,n\pi[$. Le rayon de convergence de la série est donc $R=n\pi$.

b) Pour tout $N \in \mathbb{N}^*$ et pour tout $x \in]-\pi, \pi[$, on a

$$\begin{aligned} S_N(x) &= 2\sum_{n=1}^{N} g_n(x) = 2\sum_{n=1}^{N}\left(\sum_{k=1}^{+\infty}\left(\frac{x}{n\pi}\right)^{2k}\right) \\ &= 2\sum_{k=1}^{+\infty}\left(\sum_{n=1}^{N}\left(\frac{x}{n\pi}\right)^{2k}\right) = 2\sum_{k=1}^{+\infty}\frac{1}{\pi^{2k}}\left(\sum_{n=1}^{N}\frac{1}{n^{2k}}\right)x^{2k} \\ &= 2\sum_{k=1}^{+\infty}\frac{1}{\pi^{2k}}\,\zeta_N(2k)\,x^{2k} \end{aligned}$$

où $\zeta_N : t \mapsto \sum_{n=1}^{N}\frac{1}{n^t}$. Le rayon de convergence de la série entière $\sum_{k=1}^{+\infty}\frac{1}{\pi^{2k}}\zeta_N(2k)\,x^{2k}$ vaut π car[33] on a effectué la somme de N séries entières de rayon de convergence $\pi, 2\pi, \ldots, N\pi$ tous distincts dont le minimum est π.

3 - a) On obtient, en utilisant le développement en série entière de S_N établi à la question précédente, que pour $x \in]-\pi, \pi[$:

$$\begin{aligned} S_N(x) - T(x) &= 2\sum_{n=1}^{+\infty}\zeta_N(2n)\frac{x^{2n}}{\pi^{2n}} - 2\sum_{n=1}^{+\infty}\zeta(2n)\frac{x^{2n}}{\pi^{2n}} \\ &= 2\left(\sum_{n=1}^{+\infty}\left(\zeta_N(2n)-\zeta(2n)\right)\frac{x^{2n}}{\pi^{2n}}\right). \end{aligned}$$

On a $\zeta_N(2n) - \zeta(2n) = \sum_{k=1}^{N}\frac{1}{k^{2n}} - \sum_{k=1}^{+\infty}\frac{1}{k^{2n}} = \sum_{k=N+1}^{+\infty}\frac{1}{k^{2n}} = \lim_{p\to+\infty}\sum_{k=N+1}^{p}\frac{1}{k^{2n}}$.

De plus, la fonction $\psi : t \mapsto 1/t^{2n}$ étant décroissante sur $]0, +\infty[$, on a[34] pour tout $k \in \mathbb{N}^*$:

$$\frac{1}{(k+1)^{2n}} \leqslant \int_k^{k+1}\frac{1}{t^{2n}}\,\mathrm{d}t \leqslant \frac{1}{k^{2n}}.$$

On en déduit que :

$$\underbrace{\sum_{k=N+1}^{p}\int_k^{k+1}\frac{1}{t^{2n}}\,\mathrm{d}t}_{=\int_{N+1}^{P+1}\frac{1}{t^{2n}}\,\mathrm{d}t} \leqslant \sum_{k=N+1}^{P}\frac{1}{k^{2n}} \leqslant \underbrace{\sum_{k=N+1}^{p}\int_{k-1}^{k}\frac{1}{t^{2n}}\,\mathrm{d}t}_{=\int_N^P\frac{1}{t^{2n}}\,\mathrm{d}t}$$

---

[33] Voir la proposition 5.6 p. 244.

[34] La décroissance de ψ sur $[k, k+1]$ implique que pour tout $t \in [k, k+1]$ on a $\psi(k+1) \leqslant \psi(t) \leqslant \psi(k)$ ce qui entraîne que

$$\psi(k+1) = \int_k^{k+1}\psi(k+1)\,\mathrm{d}t \leqslant \int_k^{k+1}\psi(t)\,\mathrm{d}t \leqslant \int_k^{k+1}\psi(k)\,\mathrm{d}t = \psi(k).$$

En considérant la limite quand P tend vers $+\infty$, on obtient pour tout $n \in \mathbb{N}^*$,

$$\underbrace{\int_N^{+\infty} \frac{1}{t^{2n}}\,\mathrm{d}t}_{= \frac{1}{2n-1}\frac{1}{(N+1)^{2n-1}}} \leqslant \zeta_N(2n) - \zeta(2n) \leqslant \underbrace{\int_{N+1}^{+\infty} \frac{1}{t^{2n}}\,\mathrm{d}t}_{= \frac{1}{2n-1}\frac{1}{N^{2n-1}}}$$

Pour tout $n \in \mathbb{N}^*$ on a donc $0 \leqslant \zeta_N(2n) - \zeta(2n) \leqslant 1/N$ car la suite de terme général $v_n = 1/(2n-1)N^{2n-1}$ est une suite décroissante qui est majorée par $v_1 = 1/N$. On en déduit que pour tout $x \in]-\pi, \pi[$ et tout $n \in \mathbb{N}^*$, on a

$$0 \leqslant (\zeta_N(2n) - \zeta(2n))\,\frac{x^{2n}}{\pi^{2n}} \leqslant \frac{1}{N}\,\frac{x^{2n}}{\pi^{2n}}$$

ce qui implique que

$$0 \leqslant T(x) - S_N(x) \leqslant \frac{1}{N}\sum_{n=1}^{+\infty}\left(\frac{x}{\pi}\right)^{2n}.$$

Le théorème d'encadrement indique alors que $\lim_{N\to+\infty} S_N(x) = T(x)$ pour tout $x \in]-\pi, \pi[$. Puisque $\lim_{N\to+\infty} S_N(x) = \sum_{n=1}^{+\infty} \frac{2x^2}{n^2\pi^2 - x^2}$ on en conclut que pour tout $x \in]-\pi, \pi[$ on a

$$\sum_{n=1}^{+\infty} \frac{2x^2}{n^2\pi^2 - x^2} = 2\sum_{n=1}^{+\infty} \zeta(2n)\,\frac{x^{2n}}{\pi^{2n}}.$$

b) Pour $x \in]-\pi, \pi[$, on a : $x\,\mathrm{cotan}(x) = 1 - \sum_{n=1}^{+\infty} \frac{2x^2}{n^2\pi^2 - x^2}$ et d'après ce qui précède

$$\sum_{n=1}^{+\infty} \frac{2x^2}{n^2\pi^2 - x^2} = 2\sum_{n=1}^{+\infty} \zeta(2n)\,\frac{x^{2n}}{\pi^{2n}}.$$

On en déduit que pour tout $x \in]-\pi, \pi[$, on a

$$x\,\mathrm{cotan}(x) = 1 - 2\sum_{n=1}^{+\infty} \zeta(2n)\,\frac{x^{2n}}{\pi^{2n}}.$$

c) Calculons le développement limité en 0 de $\phi : x \mapsto x\,\mathrm{cotan}(x)$. Le développement limité en 0 à l'ordre 7 de la fonction sinus est :

$$\sin(x) = x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \frac{1}{7!}x^7 + o_0(x^7).$$

On en déduit que :

$$\frac{\sin(x)}{x} = \underbrace{1 - \frac{1}{3!}x^2 + \frac{1}{5!}x^4 - \frac{1}{7!}x^6}_{= Q(x)} + o_0(x^6).$$

Le développement limité en 0 à l'ordre 6 de la fonction cosinus est :

$$\cos(x) = \underbrace{1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 - \frac{1}{6!}x^6}_{= P(x)} + o_0(x^6).$$

Pour tout $x \in]-\pi, \pi[$, on a $x\operatorname{cotan}(x) = \cos(x)/(\sin(x)/x)$. La partie régulière du développement limité à l'ordre 6 au voisinage de 0 de la fonction ϕ s'obtient donc en effectuant la division selon les puissances croissantes de P par Q à l'ordre 6. On obtient :

$$x\operatorname{cotan}(x) = 1 - \frac{1}{3}x^2 - \frac{1}{45}x^4 - \frac{2}{945}x^6 + o_0(x^6).$$

Or, on a établi à la question précédente que pour tout $x \in]-\pi, \pi[$,

$$x\operatorname{cotan}(x) = 1 - 2\sum_{n=1}^{+\infty} \zeta(2n)\frac{x^{2n}}{\pi^{2n}} = 1 - \frac{2\zeta(2)}{\pi^2}x^2 - \frac{2\zeta(4)}{\pi^4}x^4 - \frac{2\zeta(6)}{\pi^6}x^6 + R(x)$$

où $R(x) = -2\displaystyle\sum_{n=4}^{+\infty} \zeta(2n)\frac{x^{2n}}{\pi^{2n}} = -2x^8 \sum_{n=0}^{+\infty} \zeta(2n+8)\frac{x^{2n}}{\pi^{2n+8}} = o_0(x^6)$.

Puisque

$$x\operatorname{cotan}(x) = 1 - \frac{1}{3}x^2 - \frac{1}{45}x^4 - \frac{2}{945}x^6 + o_0(x^6)$$

et $$x\operatorname{cotan}(x) = 1 - \frac{2\zeta(2)}{\pi^2}x^2 - \frac{2\zeta(4)}{\pi^4}x^4 - \frac{2\zeta(6)}{\pi^6}x^6 + o_0(x^6)$$

et puisque le développement limité à un ordre fixé et en un point donné d'une fonction est unique, on en déduit que :

$$\frac{2\zeta(2)}{\pi^2} = \frac{1}{3} \quad \text{soit} \quad \zeta(2) = \frac{\pi^2}{6},$$
$$\frac{2\zeta(4)}{\pi^4} = \frac{1}{45} \quad \text{soit} \quad \zeta(4) = \frac{\pi^4}{90},$$
$$\frac{2\zeta(6)}{\pi^6} = \frac{2}{945} \quad \text{soit} \quad \zeta(6) = \frac{\pi^6}{945}.$$

On peut bien entendu poursuivre le développement limité à un ordre plus élevé et obtenir ainsi les valeurs de $\zeta(2k)$ pour $k \geqslant 4$.

Solution de l'exercice 6.5

1 - La fonction f étant impaire, les coefficients $(a_n)_n$ sont nuls et pour tout $n \in \mathbb{N}^*$ on a

$$b_n = \frac{1}{\pi}\int_0^{\pi} (\pi - x)\sin(nx)\,\mathrm{d}x = \int_0^{\pi} \sin(nx)\,\mathrm{d}x - \frac{1}{\pi}\int_0^{\pi} x\sin(nx)\,\mathrm{d}x.$$

La première intégrale vaut

$$\int_0^{\pi} \sin(nx)\ \mathrm{d}x = \left[-\frac{1}{n}\cos(nx)\right]_0^{\pi} = \frac{1}{n} - \frac{(-1)^n}{n}.$$

La seconde intégrale se calcule en effectuant une intégration par parties :

$$\int_0^{\pi} x\sin(nx)\ \mathrm{d}x = \left[-\frac{x}{n}\cos(nx)\right]_0^{\pi} + \frac{1}{n}\int_0^{\pi}\cos(nx)\ \mathrm{d}x = \frac{(-1)^n\,\pi}{n}.$$

On obtient $b_n = 1/n$ pour tout $n \in \mathbb{N}^*$. La série de Fourier de f est donc

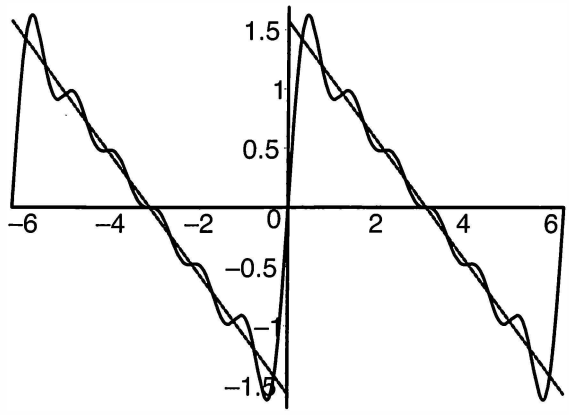
$$\mathcal{S}_F(f)(x) = \sum_{n=1}^{+\infty} \frac{1}{n}\sin(nx).$$

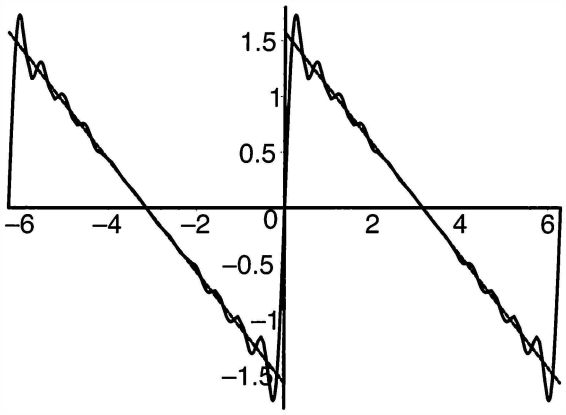
Pour calculer les sommes partielles de la série de Fourier $\mathcal{S}_F(f)$ avec MAPLE, on peut utiliser la procédure suivante qui pour une fonction `f` donnée retourne la somme partielle `SF` d'ordre `N`.

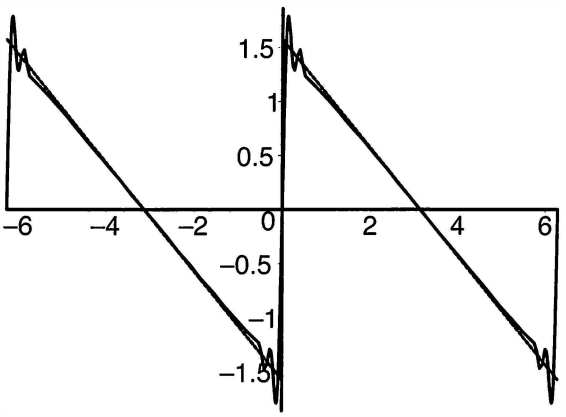
```
FScoef:=proc(f,N,SF)
  local n,a0,a,b;
  assume(n::integer): additionaly(n>0):
  a0:= (1/(2*Pi))*int(f(t),t=0..2*Pi):
   a:= n-> simplify((1/Pi)*int(f(t)*cos(n*t),t=0..2*Pi)):
   b:= n-> simplify((1/Pi)*int(f(t)*sin(n*t),t=0..2*Pi)):
  SF:=t->a0+ sum(a(n)*cos(n*t)+b(n)*sin(n*t),n=1..N):
end proc:
```

Pour visualiser les sommes partielles, on peut procéder de la manière suivante.

```
> f:=x->piecewise(x>-2*Pi and x<0, (1/2)*(Pi+x),
                  x>0 and x<2*Pi, (1/2)*(x-Pi));
> for i in [6,12,24] do FScoef(f,i,SF[i]): od:
> plot([f,SF[6]],-2*Pi..2*Pi,linestyle=[DOT,SOLID]);
  plot([f,SF[12]],-2*Pi..2*Pi,linestyle=[DOT,SOLID]);
  plot([f,SF[24]],-2*Pi..2*Pi,linestyle=[DOT,SOLID]);
```


1.5
1
0.5
−6
−4
−2
0
2
4
6
−0.5
−1
−1.5


1.5
1
0.5
−6
−4
−2
0
2
4
6
−0.5
−1
−1.5


1.5
1
0.5
−6
−4
−2
0
2
4
6
−0.5
−1
−1.5

On peut observer le phénomène de Gibbs : au voisinage des points de discontinuité $-\pi, 0$ et π, les valeurs prises par les sommes partielles de la série de Fourier oscillent autour de la valeur $\frac{\pi}{2}$ ou $-\frac{\pi}{2}$. Par ailleurs, on peut observer qu'une augmentation de l'ordre de la somme partielle ne réduit pas l'amplitude de l'oscillation ; toutefois, l'oscillation s'effectue sur des intervalles de plus en plus petits.

2 - a) La suite de fonctions $(S_n)_n$ correspond à la suite des sommes partielles associée à la série de Fourier de f. La fonction f est dérivable par morceaux sur $[-\pi, \pi]$ puisque sur chaque intervalle $]-\pi, 0[$ et $]0, \pi[$ elle est la restriction des fonctions affines $x \mapsto -\frac{1}{2}(\pi + x)$ et $x \mapsto \frac{1}{2}(\pi - x)$. De plus, comme f est continue sur $]-\pi, 0[$ et sur $]0, \pi[$, d'après le théorème de Dirichlet (voir p. 322), pour tout $x \in]-\pi, 0[\cup]0, \pi[$ la série $\mathcal{S}_F(f)(x)$ converge et a pour somme $f(x)$. Pour $x = 0$, la série $\mathcal{S}_F(f)(0)$ a pour somme

$$\tfrac{1}{2} \lim_{x \to 0^+} f(x) + \tfrac{1}{2} \lim_{x \to 0^-} f(x) = 0 = f(0)$$

et pour $x = \pi$, la série $\mathcal{S}_F(f)(0)$ a pour somme

$$\tfrac{1}{2} \lim_{x \to \pi^+} f(x) + \tfrac{1}{2} \lim_{x \to \pi^-} f(x) = 0 = f(\pi).$$

On en déduit que la série de Fourier de f converge simplement sur $[-\pi, \pi]$ et a pour somme f. La suite $(S_n)_n$ converge donc simplement sur $[-\pi, \pi]$ vers f.

La convergence de la série de Fourier de f n'est pas uniforme sur $[-\pi, \pi]$ car f n'est pas continue sur $[-\pi, \pi]$. Comme pour tout $n \in \mathbb{N}^*$, les fonctions $f_n : x \mapsto -\frac{1}{n}\sin(nx)$ sont continues sur $[-\pi, \pi]$, si la convergence était uniforme, la somme de la série serait continue, voir le théorème 3.7 p. 135.

b) L'application S_n est dérivable sur $[-\pi, \pi]$ en tant que somme finie d'applications dérivables sur $[-\pi, \pi]$. Pour tout $x \in [-\pi, \pi]$ on a $S_n'(x) = \sum_{k=1}^n \cos(kx)$. On en déduit que $S_n'(0) = \sum_{k=1}^n \cos(0) = n$. Pour $x \in [-\pi, 0[\cup]0, \pi]$, en faisant intervenir les propriétés des sommes de progressions géométriques, on obtient :

$$S_n'(x) = \mathcal{R}\mathrm{e}\left(\sum_{k=1}^{n} \mathrm{e}^{\mathrm{i}kx}\right) = \mathcal{R}\mathrm{e}\left(\mathrm{e}^{\mathrm{i}x}\,\frac{1-\mathrm{e}^{\mathrm{i}nx}}{1-\mathrm{e}^{\mathrm{i}x}}\right) = \mathcal{R}\mathrm{e}\left(\frac{1-\mathrm{e}^{\mathrm{i}nx}}{\mathrm{e}^{-\mathrm{i}x}-1}\right).$$

On a

$$\frac{1-\mathrm{e}^{\mathrm{i}nx}}{\mathrm{e}^{-\mathrm{i}x}-1} = \frac{1-\mathrm{e}^{\mathrm{i}nx}}{\cos(x)-1-\mathrm{i}\sin(x)}.$$

En multipliant numérateur et dénominateur par $\cos(x) - 1 + \mathrm{i}\sin(x) = \mathrm{e}^{\mathrm{i}x} - 1$, on obtient :

$$\frac{1-\mathrm{e}^{\mathrm{i}nx}}{\mathrm{e}^{-\mathrm{i}x}-1} = \frac{(1-\mathrm{e}^{\mathrm{i}nx})(\mathrm{e}^{\mathrm{i}x}-1)}{(\cos(x)-1)^2+\sin^2(x)} = \frac{\mathrm{e}^{\mathrm{i}x}-1+\mathrm{e}^{\mathrm{i}nx}-\mathrm{e}^{\mathrm{i}(n+1)x}}{2(1-\cos(x))}.$$

La partie réelle de cette expression vaut

$$S_n'(x) = \frac{\cos(x)-1+\cos(nx)-\cos((n+1)x)}{2(1-\cos(x))}.$$

En utilisant la formule[35] $\cos(a)-\cos(b) = -2\sin(\frac{a+b}{2})\sin(\frac{a-b}{2})$, on obtient

$$\begin{aligned}\cos(x)-1 &= -2\sin^2(\tfrac{1}{2}x)\\ \cos(nx)-\cos((n+1)x) &= 2\sin((n+\tfrac{1}{2})x)\sin(\tfrac{1}{2}x)\end{aligned}$$

ce qui nous donne

$$S_n'(x) = \frac{-\sin^2(\frac{1}{2}x)+\sin((n+\frac{1}{2})x)\sin(\frac{1}{2}x)}{2\sin^2(\frac{1}{2}x)} = \frac{-\sin(\frac{1}{2}x)+\sin((n+\frac{1}{2})x)}{2\sin(\frac{1}{2}x)}.$$

En utilisant la formule $\sin(a)-\sin(b) = 2\sin(\frac{a-b}{2})\cos(\frac{a+b}{2})$ au numérateur avec $a=(n+\frac{1}{2})x$ et $b=\frac{1}{2}x$, on obtient pour tout $x\in[-\pi,0[\cup]0,\pi]$

$$S_n'(x) = \frac{\sin(\frac{1}{2}nx)\cos(\frac{1}{2}(n+1)x)}{\sin(\frac{1}{2}x)}.$$

c) Comme $S_n'(0)=n$, pour $x\in[0,\pi]$, on a $S_n'(x)=0$ si $x\in]0,\pi]$ et si $\sin(\frac{1}{2}nx)\cos(\frac{1}{2}(n+1)x)=0$, i.e. $\sin(\frac{1}{2}nx)=0$ ou $\cos(\frac{1}{2}(n+1)x)=0$. Or,

$$\Big(\sin(\tfrac{1}{2}nx)=0 \text{ et } x\in]0,\pi]\Big) \iff x\in\left\{\frac{2k\pi}{n}\ ;\ k\in\{1,\dots \mathrm{E}(\tfrac{n}{2})\}\right\}$$

et

$$\Big(\cos(\tfrac{1}{2}(n+1)x)=0 \text{ et } x\in]0,\pi]\Big) \iff x\in\left\{\frac{(2k+1)\pi}{n+1}\ ;\ k\in\{0,\dots \mathrm{E}(\tfrac{n}{2})\}\right\}.$$

On en conclut que sur $[0,\pi]$, l'ensemble des solutions de l'équation $S_n'(x)=0$ est

$$\left\{\frac{2k\pi}{n}, \frac{(2k+1)\pi}{n+1}\ ;\ k\in\{1,\dots \mathrm{E}(\tfrac{n}{2})\}\right\}\cup\left\{\frac{\pi}{n+1}\right\}.$$

On ne peut pas dire que MAPLE soit très indiqué pour déterminer les zéros de $S_n'(x)$:

```
> assume(n,integer);solve(sin(n*x/2)*cos((n+1)*x/2)=0,x);
```

$$0, \frac{\pi}{n+1}$$

d) Remarquons que pour $k\in\{0,\dots,\mathrm{E}(n/2)\}$, les réels s_k et t_k appartiennent à $[0,\pi]$. Pour $x\in[0,\pi]$, on a $\sin(x/2)\geqslant 0$ et le signe de $S_n'(x)$ est celui du numérateur $T_n(x)=\sin(\frac{1}{2}nx)\cos(\frac{1}{2}(n+1)x)$. Pour déterminer le signe de $S_n(t_k)$ et celui de $T_n(t_k)$, on peut avoir recours à MAPLE.

```
> u:= k->2*k*Pi/n: v:=k->(2*k+1)*Pi/(n+1):
> t:= k->(u(k)+v(k))/2: s:=k->(u(k+1)+v(k))/2:
> Tn:=(n,x)-> sin(n*x/2)*cos((n+1)*x/2):
```

[35] Voir le *Cours de première année* p. 156.

```
> simplify(Tn(n,t(k)));
```

$$\sin\left(\pi\frac{4nk+n+2k}{4(n+1)}\right)\ \cos\left(\pi\frac{4nk+n+2k}{4n}\right)$$

```
> simplify(Tn(n,s(k)));
```

$$\sin\left(\pi\frac{4nk+2k+3n+2}{4(n+1)}\right)\ \cos\left(\pi\frac{4nk+2k+3n+2}{4n}\right)$$

Posons $a_{nk}=\dfrac{\pi}{4}\,\dfrac{2k(2n+1)+n}{n+1}$ et $b_{nk}=\dfrac{\pi}{4}\,\dfrac{2k(2n+1)+n}{n}$; on a

$$T_n(t_k)=\sin(a_{nk})\cos(b_{nk}),$$
$$T_n(s_k)=\sin\left(a_{nk}+\frac{\pi}{2}\right)\cos\left(b_{nk}+\frac{\pi}{2n}+\frac{\pi}{2}\right)=-\cos(a_{nk})\sin\left(b_{nk}+\frac{\pi}{2n}\right).$$

On peut remarquer que l'on a aussi $a_{nk}=k\pi+\alpha_{nk}$ et $b_{nk}=k\pi+\beta_{nk}$ où

$$\alpha_{nk}=\frac{\pi}{4}\,\frac{n-2k}{n+1}\quad\text{et}\quad\beta_{nk}=\frac{\pi}{4}\,\frac{n+2k}{n}.$$

Comme pour tout réel θ, on a $\sin(\theta+k\pi)=(-1)^k\sin(\theta)$ et $\cos(\theta+k\pi)=(-1)^k\cos(\theta)$, on en déduit que :

$$T_n(t_k)=\sin(\alpha_{nk})\cos(\beta_{nk})\quad\text{et}\quad T_n(s_k)=-\cos(\alpha_{nk})\sin\left(\beta_{nk}+\frac{\pi}{2n}\right).$$

Comme $k<\frac{n}{2}$, on a $\alpha_{nk}=\frac{\pi}{4}\,\frac{n-2k}{n+1}\in]0,\frac{\pi}{4}]$ et $\beta_{nk}=\frac{\pi}{4}\,\frac{n+2k}{n}\in[\frac{\pi}{4},\frac{\pi}{2}[$. De plus $\beta_{nk}+\frac{\pi}{2n}\in[\frac{\pi}{4},\pi[$. On en déduit d'une part que $\sin(\alpha_{nk})>0$, $\cos(\beta_{nk})>0$ et d'autre part que $\cos(\alpha_{nk})>0$, $\sin\left(\beta_{nk}+\frac{\pi}{2n}\right)>0$. On en conclut que $T_n(t_k)>0$ et $T_n(s_k)<0$ pour tout $k\in\mathbb{N}$ avec $k<\frac{n}{2}$.

e) On déduit de l'étude précédente le tableau de variation suivant pour S_n sur $[0,\pi]$:

x	0	t_0	v_0	s_0	u_1	t_1	v_1	...	u_m	t_m	v_m	π
$S_n'(x)$	n	+	0	−	0	+	0	...	0	+	0	−
$S_n(x)$	0	↗		↘		↗				↗		↘ 0

où on a posé $m=\mathrm{E}(n/2)$. Le tableau de variation indique que les maxima relatifs de $S_n(x)$ sont les réels $m_k^{(n)}=S_n(v_k)$ où $k\in\mathbb{N}$ est tel que $k\leqslant n/2$.

3 - a) Soit $k\in\mathbb{N}^*$ tel que $k\leqslant n/2$; on a

$$m_k^{(n)}-m_{k-1}^{(n)}=S_n(v_k)-S_n(v_{k-1})=\int_{v_{k_1}}^{v_k}S_n'(t)\,\mathrm{d}t$$
$$=\int_{v_{k-1}}^{v_k}\frac{\sin(\frac{1}{2}nt)\cos(\frac{1}{2}(n+1)t)}{\sin(\frac{1}{2}t)}\,\mathrm{d}t.$$

b) Pour tout $k \in \mathbb{N}^*$ tel que $k \leqslant n/2$ on a $[v_{k_1}, v_k] \subset [0, \pi]$ donc si $t \in [v_{k_1}, v_k]$ alors $\frac{1}{2}t \in [0, \pi/2]$. Comme la fonction sinus est croissante sur $[0, \pi/2]$, on en déduit que pour tout $t \in [v_{k_1}, v_k]$ on a $\sin(\frac{1}{2}t) \geqslant \sin(\frac{1}{2}v_{k-1}) \geqslant 0$. On déduit alors de l'expression obtenue à la question précédente que

$$m_k^{(n)} - m_{k-1}^{(n)} \leqslant \frac{1}{\sin(\frac{1}{2}v_{k-1})} \int_{v_{k-1}}^{v_k} \sin(\tfrac{1}{2}nt)\cos(\tfrac{1}{2}(n+1)t)\ \mathrm{d}t.$$

c) On a

$$\begin{aligned} J &= \int_{v_{k-1}}^{v_k} \sin(\tfrac{1}{2}nt)\cos(\tfrac{1}{2}(n+1)t)\ \mathrm{d}t = \frac{1}{2}\int_{v_{k-1}}^{v_k} \left(\sin(\tfrac{2n+1}{2}t) - \sin(\tfrac{1}{2}t)\right)\ \mathrm{d}t \\ &= \left[-\tfrac{1}{2n+1}\cos\left(\tfrac{2n+1}{2}t\right) + \cos(\tfrac{1}{2}t)\right]_{v_{k-1}}^{v_k} \\ &= \tfrac{1}{2n+1}\left(\cos\left(\tfrac{2n+1}{2}v_{k-1}\right) - \cos\left(\tfrac{2n+1}{2}v_k\right)\right) + \cos(\tfrac{1}{2}v_k) - \cos(\tfrac{1}{2}v_{k-1}). \end{aligned}$$

Comme

$$\frac{2n+1}{2}v_{k-1} = \frac{2n+1}{2}\frac{2k-1}{n+1}\pi = (2k-1)\pi - \frac{2k-1}{n+1}\frac{\pi}{2}$$

on en déduit que $\cos\left(\frac{2n+1}{2}v_{k-1}\right) = (-1)^{2k-1}\cos\left(\frac{2k-1}{n+1}\frac{\pi}{2}\right)$. De même, $\cos\left(\frac{2n+1}{2}v_k\right) = (-1)^{2k+1}\cos\left(\frac{2k+1}{n+1}\frac{\pi}{2}\right)$. On a donc

$$\begin{aligned} J &= \tfrac{(-1)^{2k+1}}{2n+1}\left(\cos\left(\tfrac{2k-1}{n+1}\tfrac{\pi}{2}\right) - \cos\left(\tfrac{2k+1}{n+1}\tfrac{\pi}{2}\right)\right) + \cos\left(\tfrac{2k+1}{n+1}\tfrac{\pi}{2}\right) - \cos\left(\tfrac{2k-1}{n+1}\tfrac{\pi}{2}\right) \\ &= \left(1 - \tfrac{(-1)^{2k+1}}{2n+1}\right)\left(\cos\left(\tfrac{2k+1}{n+1}\tfrac{\pi}{2}\right) - \cos\left(\tfrac{2k-1}{n+1}\tfrac{\pi}{2}\right)\right) \\ &= 2\left(\tfrac{(-1)^{2k+1}}{2n+1} - 1\right)\sin(\tfrac{k}{n+1}\pi)\sin(\tfrac{1}{2(n+1)}\pi). \end{aligned}$$

On notera que MAPLE ne parvient pas à obtenir une expression aussi simple que celle-ci.

d) Pour tout $k \in \{1, \ldots, \mathrm{E}(\frac{n}{2})\}$ on a $\frac{k}{n+1}\pi \in]0, \pi[$ donc $\sin(\frac{k}{n+1}\pi) > 0$ et par conséquent $J < 0$. Comme $\frac{1}{2}v_{k-1} = \frac{2k-1}{2n+2}\pi$ appartient à $]0, \pi[$ pour tout $k \in \{1, \ldots, \mathrm{E}(\frac{n}{2})\}$, on a $\sin(\frac{1}{2}v_{k-1}) > 0$ et on en déduit que $m_k^{(n)} - m_{k-1}^{(n)} < 0$. La suite $(m_k^{(n)})_{k \in \{0,\ldots,\mathrm{E}(\frac{n}{2})\}}$ est donc strictement décroissante.

4 - a) Pour tout $n \in \mathbb{N}$, on a

$$\mu_n = \sup_{x \in [0,\pi]} S_n(x) = \max\{m_k^{(n)}\ ;\ k \in \{0, \ldots, \mathrm{E}(\tfrac{n}{2})\} = m_0^{(n)}$$

car la suite $(m_k^{(n)})_{k \in \{0,\ldots,\mathrm{E}(\frac{n}{2})\}}$ est strictement décroissante.

b) Pour tout $n \in \mathbb{N}$, on a

$$\begin{aligned} \mu_n &= m_0^{(n)} = S_n(v_0) = S_n(\tfrac{1}{n+1}\pi) = \sum_{k=1}^{n} \frac{1}{k}\sin\left(\frac{k\pi}{n+1}\right) \\ &= \sum_{k=0}^{n} \frac{1}{k}\sin\left(\frac{k\pi}{n+1}\right) = \frac{\pi}{n+1}\sum_{k=0}^{n} \frac{n+1}{k\pi}\sin\left(\frac{k\pi}{n+1}\right). \end{aligned}$$

La quantité μ_n s'interprète comme la somme de Riemann associée à la fonction $x \mapsto \sin(x)/x$ sur l'intervalle $[0, \pi]$ pour une subdivision uniforme de $[0, \pi]$ de pas $h = \frac{\pi}{n+1}$. Cette somme de Riemann μ_n tend vers $\int_0^\pi \sin(t)/t \, \mathrm{d}t$ lorsque n tend vers $+\infty$, voir le théorème 18.3 p. 868 du *Cours de première année*.

c) Pour obtenir une valeur approchée de I, on peut avoir recours à MAPLE :

```
> evalf(Int(sin(t)/t,t=0..Pi));
```

$$1.851937052$$

5 - On a $\lim_{n\to+\infty} S_n(x) = \sum_{k=1}^{+\infty} \frac{\sin(kx)}{k}$. Cette série de fonctions est la série de Fourier associée à f. Nous avons établi à la question 1 que sa somme était égale à f. On a donc

$$\sup_{x\in[0,\pi]} \left(\lim_{n\to+\infty} S_n(x) \right) = \sup_{x\in[0,\pi]} f(x) = \frac{\pi}{2} \approx 1,570796327.$$

Par ailleurs,

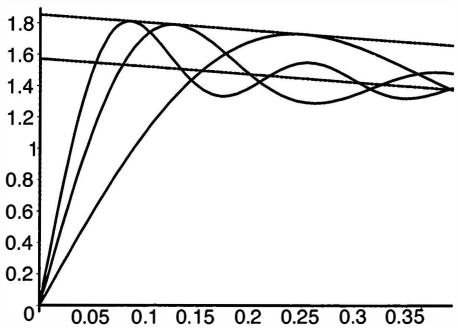
$$\lim_{n\to+\infty} \left(\sup_{x\in[0,\pi]} S_n(x) \right) = \lim_{n\to+\infty} \mu_n = I \approx 1,851937052.$$

Les sommes partielles S_n associées à la série de Fourier de f prennent pour valeur maximale μ_n en $v_0 = \frac{\pi}{n+1}$. Cette valeur maximale tend vers $\mu \approx 1,851937052$ lorsque n tend vers $+\infty$ et la valeur pour laquelle ce maximum est réalisé tend vers 0. De plus, l'écart entre cette valeur maximale μ_n et la valeur prise par f en v_0 ne tend pas vers 0 mais vers $\mu - \frac{\pi}{2} \approx 0,281140725$. L'augmentation de l'ordre n de la somme partielle considérée ne diminue donc pas l'écart entre f et S_n contrairement à ce que l'intuition pourrait suggérer. Au contraire, l'écart tend vers une valeur seuil non nulle lorsque l'ordre n de la somme partielle augmente. Toutefois, l'oscillation s'effectue sur des intervalles de plus en plus petits puisque $v_0 = \frac{\pi}{n+1}$ tend vers 0 quand n tend vers $+\infty$.

```
> I1:=evalf(Int(sin(t)/t,t=0..Pi));
```

$$I1 := 1.851937052$$

```
> for i in [12,24,36] do FScoef(f,0,2*Pi,i,a,b,SF[i]): od:
> g:=x-> I1-x/2:
> plot([f,SF[12],SF[24],SF[36],g],0..Pi/8);
```



CHAPITRE 7

Espaces métriques et espaces vectoriels normés

7.1 Distance et espace métrique

Nous avons étudié dans le cours de première année la notion de limite et de continuité pour les fonctions de la variable réelle ainsi que la notion de convergence pour les suites numériques réelles. On a pu constater qu'intervenait de façon prépondérante dans leur définition et dans l'étude de leurs propriétés la notion de distance entre deux nombres réels. Cette distance entre deux nombres réels x et y définie à l'aide de la valeur absolue par $|x-y|$ possède les propriétés suivantes : c'est une quantité positive qui est nulle si et seulement si les deux réels sont égaux, elle est symétrique (la distance $|x-y|$ entre x et y est la même que celle $|y-x|$ entre y et x) et la distance entre x et y est toujours inférieure à celle séparant x et z d'une part et z et y d'autre part (c'est l'inégalité triangulaire $|x-y| \leqslant |x-z| + |z-y|$). Nous avons par ailleurs pu constater que la notion de limite et de continuité pour les fonctions de la variable complexe ainsi que la notion de convergence pour les suites numériques complexes étaient également basées sur la notion de distance entre deux nombres complexes. Cette distance définie à l'aide du module possédait des propriétés analogues à celles énumérées ici pour la distance entre réels.

Ces constatations motivent l'introduction d'une notion générale de distance sur un ensemble quelconque afin de pouvoir définir la notion de limite pour une fonction définie sur un ensemble autre que $\mathbb{R}$ ou $\mathbb{C}$ (on pense en particulier à $\mathbb{R}^2$ ou $\mathbb{R}^3$ pour la modélisation de phénomènes physiques) ou celle de convergence d'une suite sur un tel ensemble.

DÉFINITION 7.1 *On appelle* distance *sur un ensemble E non vide toute application d de E^2 dans $\mathbb{R}$ qui vérifie les conditions suivantes*[1] *:*

1. $\forall (x,y) \in E^2 \quad d(x,y) \geqslant 0$;
2. $\forall (x,y) \in E^2 \quad (d(x,y) = 0 \iff x = y)$;
3. $\forall (x,y) \in E^2 \quad d(x,y) = d(y,x)$;
4. $\forall (x,y,z) \in E^3 \quad d(x,y) \leqslant d(x,z) + d(z,y)$.

Un ensemble E muni d'une distance d est appelé un espace métrique *et est noté (E,d).*

Remarque La première condition de la définition 7.1 est appelée *condition de positivité*, la deuxième *condition de séparation* et la troisième *condition de symétrie* ; la quatrième condition est appelée *première inégalité triangulaire.* ◇

Exemples

1. L'application $d : (x, y) \in \mathbb{R}^2 \times \mathbb{R}^2 \mapsto \sqrt{(y_1 - x_1)^2 + (y_2 - x_2)^2}$ où (x_1, x_2) et (y_1, y_2) désignent respectivement les coordonnées de x et de y est une distance sur $\mathbb{R}^2$ appelée *distance euclidienne* dans le plan. Il est clair que l'application d est à valeurs positives et que la condition de symétrie est satisfaire par d. De plus, on a $d(x, y) = 0$ si et seulement si $(y_1 - x_1)^2 + (y_2 - x_2)^2 = 0$; une somme de termes positifs étant nulle si et seulement si chacun de ses termes est nulle, l'égalité précédente a lieu uniquement lorsque $x_1 = y_1$ et $x_2 = y_2$. On a donc bien $d(x, y) = 0$ si et seulement si $x = y$. Il reste à établir l'inégalité triangulaire pour d. Pour $x = (x_1, x_2)$,$y = (y_1, y_2)$ et $z = (z_1, z_2)$ dans $\mathbb{R}^2$ on a

$$\begin{aligned} d(x,y)^2 &= (y_1 - x_1)^2 + (y_2 - x_2)^2 \\ &= \big(\overbrace{(y_1 - z_1) + (z_1}^{=0} - x_1)\big)^2 + \big(\overbrace{(y_2 - z_2) + (z_2}^{=0} - x_2)\big)^2 \\ &= \underbrace{(x_1 - z_1)^2 + (x_2 - z_2)^2}_{=d(x,z)^2} + \underbrace{(y_1 - z_1)^2 + (y_2 - z_2)^2}_{=d(z,y)^2} \\ &\quad + 2\underbrace{\big((y_1 - z_1)(z_1 - x_1) + (y_2 - z_2)(z_2 - x_2)\big)}_{=\alpha} \end{aligned}$$

et d'après l'inégalité de Cauchy-Schwarz[(2)], on a

$$\begin{aligned} \alpha &\leqslant \big((y_1 - z_1)^2 + (y_2 - z_2)^2\big)^{\frac{1}{2}} \times \big((z_1 - x_1)^2 + (z_2 - x_2)^2\big)^{\frac{1}{2}} \\ &= d(y,z)\, d(z,x). \end{aligned}$$

On en déduit que

$$d(x,y)^2 \leqslant d(x,z)^2 + d(z,y)^2 + 2d(y,z)\, d(z,x) = \big(d(x,z) + d(z,y)\big)^2;$$

comme toutes les quantités sont positives, on a $d(x, y) \leqslant d(x, z) + d(z, y)$. L'inégalité triangulaire pour d est établie.

2. On vérifie aisément à partir des propriétés de la valeur absolue que l'application $d : (x, y) \in \mathbb{R}^2 \times \mathbb{R}^2 \mapsto |y_1 - x_1| + |y_2 - x_2|$ où (x_1, x_2) et (y_1, y_2) désignent les coordonnées de x et y est également une distance sur $\mathbb{R}^2$.

3. Si E est un ensemble non vide, l'application d de E^2 dans $\mathbb{R}^+$ définie par $d(x, y) = 1$ si $x \neq y$ et 0 si $x = y$ est une distance sur E appelée *distance discrète.* Cette distance est souvent utilisée pour construire des contre-exemples.

---

(1) On rappelle que l'on note E^2 le produit cartésien de E par E ($E^2 = E \times E$). Plus généralement si n désigne un entier supérieur à 2, on note E^n le produit cartésien $E \times E \times \cdots \times E$ (n fois).

(2) Voir la proposition 3.6 p. 101 du *Cours de première année.*

Nous avons évoqué en introduction le fait que l'application $(x,y) \in \mathbb{R}^2 \mapsto |x-y|$ définit une distance sur $\mathbb{R}$. Cette distance est qualifiée de distance « usuelle » (ou « canonique ») sur $\mathbb{R}$ mais l'ensemble des réels peut être muni d'autres distances comme par exemple celle définie dans l'exercice suivant.

Exercice 7.1 Montrer que l'application

$$d : (x,y) \in \mathbb{R}^2 \longmapsto \ln\left(1 + \frac{|x-y|}{1+|x-y|}\right)$$

définit une distance dans $\mathbb{R}$.

D'une façon générale, si f est une application injective définie sur $\mathbb{R}$ à valeurs dans $\mathbb{R}$ alors l'application $d : (x,y) \in \mathbb{R}^2 \mapsto |f(x)-f(y)|$ est une distance sur $\mathbb{R}$. En effet, il résulte des propriétés de la valeur absolue que pour tout $(x,y) \in \mathbb{R}^2$ on a $d(x,y) \geqslant 0$ et $d(x,y) = d(y,x)$. Par ailleurs, $d(x,y) = 0$ si et seulement si $f(x) = f(y)$; sous l'hypothèse que f est une application injective[3] cette dernière relation équivaut à $x = y$. Enfin, pour tout $(x,y,z) \in \mathbb{R}^3$ on obtient en utilisant les propriétés de la valeur absolue[4],

$$\begin{aligned} d(x,y) &= |f(x)-f(y)| = |f(x)-f(z)+f(z)-f(y)| \\ &\leqslant |f(x)-f(z)| + |f(z)-f(y)| = d(x,z)+d(z,y). \end{aligned}$$

Par exemple, l'application arc-tangente étant injective de $\mathbb{R}$ dans $]-\frac{\pi}{2},\frac{\pi}{2}[$, l'application $d : (x,y) \in \mathbb{R}^2 \mapsto |\arctan(x) - \arctan(y)|$ est une distance sur $\mathbb{R}$.

Sur le même principe, l'application f de $\overline{\mathbb{R}}$ dans $[-1,1]$ définie par[5]

$$f(x) = \frac{x}{1+|x|} \text{ si } x \in \mathbb{R} \quad \text{et } f(+\infty) = 1,\ f(-\infty) = -1$$

étant injective, l'application $d : (x,y) \in \overline{\mathbb{R}} \times \overline{\mathbb{R}} \mapsto |f(x)-f(y)|$ est une distance sur $\overline{\mathbb{R}}$.

Exercice 7.2
1 - Montrer que si ϕ est une application de $[0,+\infty[$ dans $\mathbb{R}$, strictement croissante, vérifiant $\phi(0) = 0$ et $\phi(u+v) \leqslant \phi(u) + \phi(v)$ pour tout $(u,v) \in [0,+\infty[^2$ alors l'application $d : (x,y) \in \mathbb{R}^2 \mapsto \phi(|x-y|)$ est une distance sur $\mathbb{R}$.
2 - En déduire que $(x,y) \in \mathbb{R}^2 \mapsto \frac{|x-y|}{1+|x-y|}$ est une distance sur $\mathbb{R}$.
3 - Plus généralement, montrer que si δ est une distance sur $\mathbb{R}$ alors $d : (x,y) \in \mathbb{R}^2 \mapsto \phi(\delta(x,y))$ est aussi une distance sur $\mathbb{R}$.

Bien entendu, il existe de nombreux ensembles autres de $\mathbb{R}$ ou $\mathbb{C}$ qui peuvent

[3] Voir la définition 2.19 p. 41 du *Cours de première année*.
[4] Voir la proposition 3.7 p. 102 du *Cours de première année*.
[5] On rappelle que $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty, -\infty\}$ désigne la droite numérique achevée, voir le *Cours de première année* p. 112.

être munis de distances. On peut en particulier introduire des distances sur des ensembles de fonctions comme l'illustrent les exemples suivants.

Exemples On désigne par a et b deux réels tels que $a < b$.

1. L'application

$$\boxed{d_1 : (f,g) \in \mathcal{C}^0([a,b],\mathbb{R}) \times \mathcal{C}^0([a,b],\mathbb{R}) \longmapsto \int_a^b |f(x) - g(x)| \, \mathrm{d}x}$$

est une distance sur l'ensemble $\mathcal{C}^0([a,b],\mathbb{R})$ des applications continues sur $[a,b]$, à valeurs réelles. Remarquons que pour tous f et g appartenant à $\mathcal{C}^0([a,b],\mathbb{R})$, l'application $f : x \in [a,b] \mapsto |f(x) - g(x)|$ est continue et donc Riemann-intégrable[6] sur $[a,b]$. Par conséquent $d_1(f,g)$ est bien définie, voir la figure 1. Vérifions que d_1 satisfait aux quatre conditions de la définition 7.1.

- Si f et g sont deux applications continues sur $[a,b]$ alors l'application $x \in [a,b] \mapsto |f(x) - g(x)|$ étant continue et positive sur l'intervalle $[a,b]$, on a[7] d'une part $\int_a^b |f(x) - g(x)| \, \mathrm{d}x \geqslant 0$ et d'autre part

$$\int_a^b |f(x) - g(x)| \, \mathrm{d}x = 0 \iff \forall x \in [a,b] \quad |f(x) - g(x)| = 0 \iff f = g,$$

  autrement dit : $d_1(f,g) \geqslant 0$ et $(d_1(f,g) = 0 \Leftrightarrow f = g)$.
- Soit $(f,g) \in \mathcal{C}^0([a,b],\mathbb{R})^2$; pour tout réel $x \in [a,b]$ on a $|f(x) - g(x)| = |g(x) - f(x)|$ ce qui implique que $d_1(f,g) = d_1(g,f)$.
- Pour $(f,g,h) \in \mathcal{C}^0([a,b],\mathbb{R})^3$ on obtient en utilisant la première inégalité triangulaire pour la valeur absolue :

$$\forall x \in [a,b] \qquad |f(x) - h(x)| \leqslant |f(x) - g(x)| + |g(x) - h(x)|.$$

  On en déduit[8] que :

$$\int_a^b |f(x) - h(x)| \, \mathrm{d}x \leqslant \int_a^b |f(x) - g(x)| \, \mathrm{d}x + \int_a^b |g(x) - h(x)| \, \mathrm{d}x,$$

  autrement dit que : $d_1(f,h) \leqslant d_1(f,g) + d_1(g,h)$.

2. L'application

$$\boxed{d_\infty : (f,g) \in \mathcal{C}^0([a,b],\mathbb{R}) \times \mathcal{C}^0([a,b],\mathbb{R}) \longmapsto \sup_{x \in [a,b]} |f(x) - g(x)|}$$

est également une distance sur $\mathcal{C}^0([a,b],\mathbb{R})$ appelée *distance de la convergence uniforme.*

---

[6] Voir la proposition 18.2 p. 842 du *Cours de première année.*

[7] Voir la proposition 18.5 p. 844 du *Cours de première année.*

[8] Voir la proposition 18.4 p. 844 du *Cours de première année.*

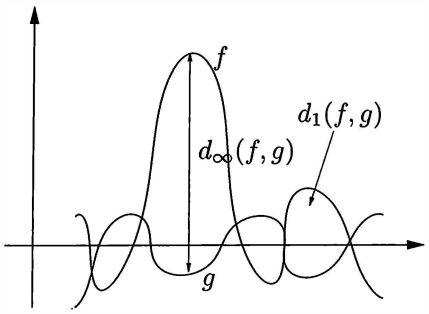


Fig. 1 Illustration graphique des distances $d_\infty(f,g)$ et $d_1(f,g)$ (valeur de l'aire grisée) entre les applications f et g.

Si f et g appartiennent à $\mathcal{C}^0([a,b],\mathbb{R})$ alors la fonction $f-g$ est continue sur l'intervalle fermé borné $[a,b]$. Elle est donc bornée[(9)] et atteint ses bornes. La quantité

$$d_\infty(f,g) = \sup_{x\in[a,b]} |f(x)-g(x)| = \max_{x\in[a,b]} |f(x)-g(x)|$$

est bien définie, voir la figure 1. Vérifions que d_∞ satisfait aux quatre conditions de la définition 7.1.

- On a clairement $d_\infty(f,g) \geqslant 0$ pour tout $(f,g) \in \mathcal{C}^0([a,b],\mathbb{R})^2$.
- Supposons que $d_\infty(f,g) = 0$; on a alors $\max_{x\in[a,b]} |f(x)-g(x)| = 0$, ce qui implique que pour tout $x \in [a,b]$ on a $|f(x)-g(x)| = 0$, autrement dit que $f = g$. La réciproque est quant à elle évidente : si $f = g$ alors $\max_{x\in[a,b]} |f(x)-g(x)| = 0$.
- Soit $(f,g) \in \mathcal{C}^0([a,b],\mathbb{R})^2$; pour tout réel $x \in [a,b]$ on a $|f(x)-g(x)| = |g(x)-f(x)|$ donc $\max_{x\in[a,b]} |f(x)-g(x)| = \max_{x\in[a,b]} |g(x)-f(x)|$. On en déduit que $d_\infty(f,g) = d_\infty(g,f)$.
- Pour tout $(f,g,h) \in \mathcal{C}^0([a,b],\mathbb{R})^3$ on a :

$$\forall x \in [a,b] \qquad |f(x)-h(x)| \leqslant |f(x)-g(x)| + |g(x)-h(x)|.$$

Or, pour tout $x \in [a,b]$

$$|f(x)-g(x)| \leqslant \max_{y\in[a,b]} |f(y)-g(y)| = d_\infty(f,g)$$

et

$$|g(x)-h(x)| \leqslant \max_{y\in[a,b]} |g(y)-h(y)| = d_\infty(g,h).$$

(9) Voir le théorème 13.4 p. 596 du *Cours de première année*. Rappelons que sur un intervalle ouvert, une application continue n'est pas nécessairement bornée (considérer par exemple la fonction logarithme sur l'intervalle $]0,1[$). La fonction $d : (f,g) \in \mathcal{C}^0(]a,b[)^2 \mapsto \sup_{x\in]a,b[} |f(x)-g(x)|$ n'est donc pas une application et ne permet pas de définir une distance sur $\mathcal{C}^0(]a,b[)$. Par contre, sur le sous-ensemble de $\mathcal{C}^0(]a,b[)$ constitué des applications bornées sur $]a,b[$, on vérifie par un raisonnement analogue à celui qui est fait ici pour $\mathcal{C}^0([a,b],\mathbb{R})$ que cette fonction d définit bien une distance.

On en déduit donc que $|f(x) - g(x)| \leqslant d_\infty(f,g) + d_\infty(g,h)$ pour tout $x \in [a,b]$. Cela implique que $\max_{x\in[a,b]} |f(x) - g(x)| \leqslant d_\infty(f,g) + d_\infty(g,h)$, autrement dit que $d_\infty(f,h) \leqslant d_\infty(f,g) + d_\infty(g,h)$.

Exercice 7.3 Montrer que l'application

$$d_2 : (f,g) \in \mathcal{C}^0([a,b],\mathbb{R}) \times \mathcal{C}^0([a,b],\mathbb{R}) \longmapsto \left(\int_a^b (f(x) - g(x))^2 \, \mathrm{d}x \right)^{\frac{1}{2}}$$

est une distance sur $\mathcal{C}^0([a,b],\mathbb{R})$.

Remarques

1. On vérifie sans difficulté que si d_1 et d_2 sont deux distances sur un ensemble E alors $d_1 + d_2$ est aussi une distance sur E. De plus, si λ est un réel strictement positif alors $\lambda\, d_1$ est une distance sur E.

2. Si (E,d) est un espace métrique et A est un sous-ensemble de E, la restriction de l'application d à l'ensemble $A \times A$ qui est définie par $d_A(x,y) = d(x,y)$ pour tout $(x,y) \in A^2$ est une distance sur A. On qualifie (A,d) de sous-espace métrique de (E,d). Par exemple l'ensemble $\mathbb{Q}$ des nombres rationnels muni de la distance $d_{\mathbb{Q}} : (x,y) \in \mathbb{Q}^2 \mapsto |x-y|$ est un sous-espace métrique de $\mathbb{R}$ muni de la distance usuelle $d : (x,y) \in \mathbb{R}^2 \mapsto |x-y|$. ◇

La distance de Hamming qui est utilisée en informatique et dans les télécommunications joue un rôle très important en théorie des codes correcteurs. Elle permet de quantifier la différence entre deux séquences de symboles. Si on pose $a = a_0 ... a_n$ et $b = b_0 ... b_n$, où les a_i et b_i sont des symboles appartenant à un ensemble A ($A = \{0,1\}$ dans le cas de bits en informatique), la distance de Hamming est définie par : $d(a,b) = \mathrm{Card}\{i \ ; \ a_i \neq b_i\}$. Si on considère les séquences binaires suivantes : $a = 0\ 0\ 0\ 1\ 1\ 1\ 1$ et $b = 1\ 1\ 0\ 1\ 0\ 1\ 1$ on a $d(a,b) = 3$; il y a 3 bits qui diffèrent.

En infographie, elle est utilisée pour comparer deux images et quantifier la façon dont ces deux images diffèrent. Calculer la distance de Hamming entre deux images revient à comptabiliser le nombre de pixels dont la couleur diffère. Par exemple, la distance de Hamming entre les deux images de gauche sur la figure 2 vaut 4078 : il y a 4078 pixels qui diffèrent entre les deux images qui sont chacune constituées de 512^2 pixels. L'image de droite indique les pixels qui diffèrent entre les deux images. Bien entendu, en pratique on essaie d'avoir une valeur relative qui ne dépende pas du nombre de pixel des deux images. Dans notre cas la distance de Hamming relative est $4078/512^2 = 0.0156$.

Ainsi définie, la distance de Hamming relative est une application à valeurs dans $[0,1]$. La valeur 0 est atteinte lorsque les deux images sont identiques alors que la valeur 1 est atteinte dans le cas où aucun des pixels entre les deux images est identique.

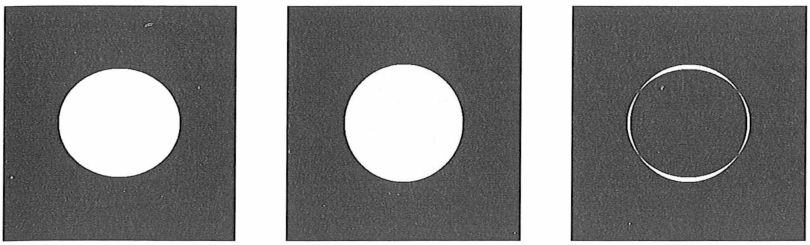


Fig. 2 Illustration de l'utilisation de la distance de Hamming en infographie.

Exercice 7.4
1 - Montrer que quelle que soit la suite numérique $(u_n)_n$, la série $\sum_n \frac{1}{2^n}\frac{|u_n|}{1+|u_n|}$ converge.
2 - Montrer que sur l'ensemble des suites numériques $\mathcal{S}(\mathbb{K})$, l'application $d : ((u_n)_n, (v_n)_n) \in \mathcal{S}(\mathbb{K})^2 \mapsto \sum_{n=0}^{+\infty} \frac{1}{2^n}\frac{|u_n - v_n|}{1+|u_n - v_n|}$ est une distance[10].

PROPOSITION 7.1 (Seconde inégalité triangulaire)
Soit (E, d) un espace métrique ; pour tout $(x, y, z) \in E^3$ on a

$$|d(x, z) - d(y, z)| \leqslant d(x, y).$$

Démonstration Remarquons pour commencer que, sans restreindre la portée de la démonstration, on peut supposer que $d(x, z) - d(y, z)$ est positif car si l'on se donne trois éléments quelconques de l'ensemble E, on peut appeler z l'un des trois et nommer x celui des deux autres éléments dont la distance à z est la plus grande. En utilisant la première inégalité triangulaire (quatrième condition de la définition 7.1), on obtient

$$d(x, z) \leqslant d(x, y) + d(y, z)$$

et par conséquent $d(x, z) - d(y, z) \leqslant d(x, y)$. Comme par hypothèse le réel $d(x, z) - d(y, z)$ est positif, on a $d(x, z) - d(y, z) = |d(x, z) - d(y, z)|$ et le résultat est démontré. □

Remarque Si E est l'ensemble $\mathbb{R}$ muni de la distance définie par $d : (x, y) \in \mathbb{R}^2 \mapsto |x - y|$ alors en prenant $z = 0$, la relation de la proposition 7.1 indique que pour tout $(x, y) \in \mathbb{R}^2$ on a

$$||x| - |y|| \leqslant |x - y|.$$

[10] On pourra établir que pour $(x, y) \in \mathbb{R}^+ \times \mathbb{R}^+$ on a : $x \leqslant y \Rightarrow \frac{x}{1+x} \leqslant \frac{y}{1+y}$.

On retrouve la seconde inégalité triangulaire pour les réels[11]. ◇

Il arrive parfois que deux distances différentes d_1 et d_2 sur un même ensemble E soient assez « ressemblantes » pour que les espaces métriques (E, d_1) et (E, d_2) possèdent certaines propriétés communes. On définit ainsi la notion de distances équivalentes.

DÉFINITION 7.2 (Distances équivalentes)
On dit que deux distances d_1 et d_2 sur un ensemble E sont équivalentes *s'il existe deux réels strictement positifs C_1 et C_2 tels que*

$$\forall (x,y) \in E^2 \qquad C_1\, d_1(x,y) \leqslant d_2(x,y) \leqslant C_2\, d_1(x,y).$$

Remarques

1. Sur l'ensemble des distances sur un ensemble E donné, la relation « être équivalentes à » est une relation réflexive (une distance est équivalente à elle même), symétrique (si une distance d_1 est équivalente à une distance d_2, cette dernière est également équivalente à la distance d_1) et transitive (si d_1 est une distance équivalente à d_2 et que d_2 est équivalente à d_3 alors d_1 est équivalente à d_3). Il s'agit donc d'une *relation d'équivalence*[12]. La propriété de symétrie indique en particulier que pour montrer que deux distances d_1 et d_2 sont équivalentes, il faut et il suffit d'établir que l'une des deux assertions suivantes est vraie :

$$\exists (C_1, C_2) \in {\mathbb{R}_+^*}^2 \quad \forall (x,y) \in E^2 \quad C_1\, d_1(x,y) \leqslant d_2(x,y) \leqslant C_2\, d_1(x,y)$$
$$\text{ou} \quad \exists (C_3, C_4) \in {\mathbb{R}_+^*}^2 \quad \forall (x,y) \in E^2 \quad C_3\, d_2(x,y) \leqslant d_1(x,y) \leqslant C_4\, d_2(x,y).$$

2. Pour montrer que deux distances ne sont pas équivalentes, il faut et il suffit de montrer que l'une des deux assertions suivantes est vraie :

$$\forall C_1 \in \mathbb{R}_+^* \quad \exists (x,y) \in E^2 \qquad C_1\, d_1(x,y) > d_2(x,y)$$
$$\text{ou} \quad \forall C_2 \in \mathbb{R}_+^* \quad \exists (x,y) \in E^2 \qquad d_1(x,y) > C_2\, d_2(x,y).$$

◇

Exemples

1. Les deux applications $d_1 : (x,y) \in \mathbb{R}^2 \times \mathbb{R}^2 \mapsto |y_1 - x_1| + |y_2 - x_2|$ et $d_2 : (x,y) \in \mathbb{R}^2 \times \mathbb{R}^2 \mapsto \sqrt{(y_1 - x_1)^2 + (y_2 - x_2)^2}$ où (x_1, x_2) et (y_1, y_2) désignent respectivement les composantes de x et y, sont des distances équivalentes sur $\mathbb{R}^2$. En effet, la relation $\sqrt{a^2 + b^2} \leqslant |a| + |b| \leqslant \sqrt{2}\sqrt{a^2 + b^2}$ valable[13] pour tous réels a, b implique que pour tous $x, y \in \mathbb{R}^2$, on a

$$\sqrt{(y_1 - x_1)^2 + (y_2 - x_2)^2} \leqslant |y_1 - x_1| + |y_2 - x_2| \leqslant \sqrt{2}\sqrt{(y_1 - x_1)^2 + (y_2 - x_2)^2}.$$

[11] Voir la proposition 3.7 p. 102 du *Cours de première année*.
[12] Voir la définition 2.28 p. 55 du *Cours de première année*.
[13] La vérification est immédiate en élevant au carré chaque terme.

2. Sur $\mathbb{R}$, la distance $d_1 : (x,y) \in \mathbb{R}^2 \mapsto |\arctan(x) - \arctan(y)|$ n'est pas équivalente à la distance usuelle $d_2 : (x,y) \in \mathbb{R}^2 \mapsto |x-y|$. Remarquons tout d'abord que pour tout $(x,y) \in \mathbb{R}^2$ on a $d_1(x,y) \leqslant d_2(x,y)$ car d'après le théorème des accroissements finis appliqué à la fonction arc-tangente entre x et y (avec $x < y$), il existe un réel $c \in]x,y[$ tel que :

$$\arctan(x) - \arctan(y) = \frac{1}{1+c^2}\,(x-y).$$

On en déduit que[14] $|\arctan(x) - \arctan(y)| = \frac{1}{1+c^2}|x-y| \leqslant |x-y|$ pour tout $(x,y) \in \mathbb{R}^2$.

Pour montrer que ces deux distances ne sont pas équivalentes il faut donc montrer qu'il n'existe pas de réel strictement positif C tel que $d_2(x,y) \leqslant C\,d_1(x,y)$, autrement dit, il faut montrer que

$$\forall C \in \mathbb{R}_+^* \quad \exists (x,y) \in \mathbb{R}^2 \quad C_3|x-y| > |\arctan(x) - \arctan(y)|.$$

Pour $C \in \mathbb{R}_+^*$ fixé, considérons les réels $x = \pi/C$ et $y = 0$. On a $C|x-y| = \pi$ et $|\arctan(x) - \arctan(y)| = \arctan(\pi/C)$. Comme la fonction arc-tangente est bornée par $\pi/2$ on a $C|x-y| > |\arctan(x) - \arctan(y)|$.
Cet exemple illustre le fait que pour deux distances non équivalentes, l'une des inégalités de la définition 7.2 peut toutefois être vraie.

3. Sur $\mathcal{C}^0([a,b],\mathbb{R})$, les distances d_1 et d_∞ définies p. 348 ne sont pas équivalentes. Pour tout $(f,g) \in \mathcal{C}^0([a,b],\mathbb{R})^2$ on a bien :

$$d_1(f,g) = \int_a^b |f(x)-g(x)|\,\mathrm{d}x \leqslant (b-a) \sup_{x\in[a,b]} |f(x)-g(x)| = (b-a)\,d_\infty(f,g),$$

mais il n'existe pas de réel strictement positif C tel que pour tout $(f,g) \in \mathcal{C}^0([a,b],\mathbb{R})^2$ on ait $d_\infty(f,g) \leqslant C\,d_1(f,g)$. Pour le vérifier, montrons que :

$$\forall C \in \mathbb{R}_+^* \quad \exists (f,g) \in \mathcal{C}^0([a,b],\mathbb{R})^2 \quad \sup_{x\in[a,b]} |f(x)-g(x)| > C \int_a^b |f(x)-g(x)|\,\mathrm{d}x.$$

Pour $C \in \mathbb{R}_+^*$ fixé, considérons l'application $f : x \in [a,b] \mapsto (b-x)^{n-1}$ où l'entier naturel n vaut[15] $2 + \mathrm{E}(C\,(b-a))$ et prenons pour g l'application nulle. On a,

$$\sup_{x\in[a,b]} |f(x)-g(x)| = (b-a)^{n-1} \quad \text{et} \quad \int_a^b |f(x)-g(x)|\,\mathrm{d}x = \frac{1}{n}(b-a)^n.$$

Puisque[15] $n > C(b-a)$, on en déduit que :

$$C\,d_1(f,g) = \frac{C}{n}(b-a)^n = \frac{1}{n}C(b-a)\,(b-a)^{n-1} < (b-a)^{n-1} = d_\infty(f,g).$$

---

[14] Les cas où $x = y$ est trivial et pour $y < x$ on peut procéder de la même manière en inversant les rôles de x et y.

[15] ici E désigne la fonction partie entière ; voir le *Cours de première année* p. 104.

Pour tout $C \in \mathbb{R}_+^*$, on dispose donc de deux fonctions f et g telles que $d_\infty(f,g) > C\, d_1(f,g)$. Les distances ne sont par conséquent pas équivalentes.

Dans la suite du cours nous serons amenés à considérer les propriétés des fonctions de plusieurs variables réelles et nous aurons besoin de distances sur $\mathbb{R}^n$. On désigne par x un élément de $\mathbb{R}^n$ et par $(x_1, \ldots, x_n)$ le n-uplet de $\mathbb{R}^n$ correspondant à x.

PROPOSITION 7.2 (Distances usuelles sur $\mathbb{R}^n$)
Soit $n \in \mathbb{N}$ tel que $n \geqslant 2$; les applications suivantes définissent trois distances sur $\mathbb{R}^n$:

$$d_\infty : (x,y) \in \mathbb{R}^n \times \mathbb{R}^n \longmapsto \max_{i \in \{1,\ldots,n\}} |x_i - y_i|,$$

$$d_1 : (x,y) \in \mathbb{R}^n \times \mathbb{R}^n \longmapsto \sum_{i=1}^{n} |x_i - y_i|,$$

$$d_2 : (x,y) \in \mathbb{R}^n \times \mathbb{R}^n \longmapsto \sqrt{\sum_{i=1}^{n} |x_i - y_i|^2}.$$

Ces distances sont équivalentes deux à deux et on a pour tout $(x,y) \in E^2$:

$$d_\infty(x,y) \leqslant d_1(x,y) \leqslant \sqrt{n}\, d_2(x,y) \leqslant n\, d_\infty(x,y).$$

Démonstration ▷ La vérification que d_∞ et d_1 sont des distances est laissée en exercice. Montrons que d_2 définit une distance sur $\mathbb{R}^n$ en établissant les conditions de la définition 7.1.

1. La condition de positivité est évidente.
2. Soit $(x,y) \in \mathbb{R}^n \times \mathbb{R}^n$ tel que $d_2(x,y) = 0$. La quantité $d_2(x,y)^2$ est alors une somme de carrés qui est nulle. On en déduit que pour tout entier i avec $1 \leqslant i \leqslant n$ on a $|x_i - y_i| = 0$. Ceci implique que $x_i = y_i$ pour tout $i \in \{1, \ldots, n\}$ et donc que $x = y$. Il est immédiat que si $x = y$ alors $d_2(x,y) = 0$.
3. Soit $(x,y) \in \mathbb{R}^n \times \mathbb{R}^n$; l'égalité $d_2(x,y) = d_2(y,x)$ résulte du fait que pour tout $i \in \{1, \ldots, n\}$ on a $(x_i - y_i)^2 = (y_i - x_i)^2$.
4. Soient x, y et z des éléments de $\mathbb{R}^n$. Montrer que $d_2(x,y) \leqslant d_2(x,z) + d_2(z,y)$ revient à montrer que $d_2^2(x,y) \leqslant d_2^2(x,z) + d_2^2(z,y) + 2d_2(x,z)d_2(z,y)$. On a

$$\begin{aligned} d_2(x,y)^2 &= \sum_{i=1}^{n} (x_i - y_i)^2 = \sum_{i=1}^{n} \big((x_i - z_i) + (z_i - y_i)\big)^2 \\ &= \sum_{i=1}^{n} (x_i - z_i)^2 + \sum_{i=1}^{n} (z_i - y_i)^2 + 2\sum_{i=1}^{n} (x_i - y_i)(z_i - y_i). \end{aligned}$$

D'après l'inégalité de Cauchy-Schwarz[16] on a

$$\begin{aligned}\sum_{i=1}^{n}(x_i-y_i)(z_i-y_i) &\leqslant \left(\sum_{i=1}^{n}(x_i-y_i)^2\right)^{\frac{1}{2}} \times \left(\sum_{i=1}^{n}(z_i-y_i)^2\right)^{\frac{1}{2}} \\ &= d_2(x,z)\times d_2(z,y).\end{aligned}$$

On a donc bien $d_2^2(x,y) \leqslant d_2^2(x,z) + d_2^2(z,y) + 2d_2(x,z)d_2(z,y)$.

$\trianglerighteq$ Montrons que les trois distances sont équivalentes. Tout d'abord, la relation $d_\infty(x,y) \leqslant d_1(x,y)$ est évidente car $d_1(x,y)$ est égal à la somme de $d_\infty(x,y)$ (le plus grand des réels positifs $|x_i - y_i|$ lorsque i varie dans $\{1,\ldots,n\}$) et de réels positifs. Par ailleurs, en utilisant l'inégalité de Cauchy-Schwarz[16] on obtient :

$$d_1^2(x,y) = \left(\sum_{i=1}^{n} 1\times|x_i-y_i|\right)^2 \leqslant \sum_{i=1}^{n}1^2 \times \sum_{i=1}^{n}|x_i-y_i|^2 = n\times d_2^2(x,y).$$

On a donc $d_1(x,y) \leqslant \sqrt{n}\, d_2(x,y)$.

Enfin, $d_\infty(x,y)$ est le plus grand des réels positifs $|x_i - y_i|$ lorsque i varie dans $\{1,\ldots,n\}$ donc la somme $\sum_{i=1}^{n}|x_i-y_i|^2$ est constituée de n termes tous inférieurs ou égaux à $d_\infty^2(x,y)$. On en déduit que :

$$d_2(x,y) = \sqrt{\sum_{i=1}^{n}(x_i-y_i)^2} \leqslant \sqrt{\sum_{i=1}^{n} d_\infty(x,y)^2} = \sqrt{n}\; d_\infty(x,y).$$

On a donc montré que : $d_\infty(x,y) \leqslant d_1(x,y) \leqslant \sqrt{n}\, d_2(x,y) \leqslant n\, d_\infty(x,y)$; les distances d_∞, d_1 et d_2 sont donc équivalentes deux à deux. □

La distance d_2 dans $\mathbb{R}^2$ mesure la distance « à vol d'oiseau » entre deux points. On peut également donner une interprétation à la distance d_1 : elle permet de calculer la distance entre deux points en ville où le déplacement ne peut se faire que selon les rues qui sont toutes orientées dans deux directions perpendiculaires[17].

Exercice 7.5 Montrer que dans $\mathbb{R}^3$ l'application

$$d : (x,y) \in \mathbb{R}^3\times\mathbb{R}^3 \mapsto \sqrt{(|x_1-y_1|+|x_2-y_2|)^2+(x_3-y_3)^2}$$

est une distance[18] équivalente à la distance d_2 définie à la proposition 7.2.

[16] Voir la proposition 3.6 p. 101 du *Cours de première année*.

[17] Le quartier de Siam à Brest en est un exemple.

[18] On peut l'assimiler à la distance associée au déplacement d'un oiseau dans une ville possédant de très hauts immeubles.

Remarques

1. La distance d_2 est appelée *distance euclidienne* sur $\mathbb{R}^n$. On notera que pour $n = 1$ les trois distances usuelles sur $\mathbb{R}^n$ se confondent et on retrouve la distance usuelle sur $\mathbb{R}$: $(x, y) \in \mathbb{R}^2 \mapsto |x - y|$.

2. De façon plus générale, pour tout $p \in \mathbb{N}^*$ l'application

$$d_p : (x, y) \in \mathbb{R}^n \times \mathbb{R}^n \longmapsto \left(\sum_{i=1}^{n} (x_i - y_i)^p \right)^{1/p}$$

définit une distance sur $\mathbb{R}^n$ qui est équivalentes aux distances d_∞, d_1 et d_2.

3. On montre également que sur $\mathbb{C}^n$ ($n \in \mathbb{N}^*$) les applications suivantes (où $|\cdot|$ désigne le module) définissent des distances équivalentes :

$$d_\infty : (x, y) \in \mathbb{C}^n \times \mathbb{C}^n \longmapsto \max_{i \in \{1, \ldots, n\}} |x_i - y_i|,$$

$$d_1 : (x, y) \in \mathbb{C}^n \times \mathbb{C}^n \longmapsto \sum_{i=1}^{n} |x_i - y_i|,$$

$$d_2 : (x, y) \in \mathbb{C}^n \times \mathbb{C}^n \longmapsto \sqrt{\sum_{i=1}^{n} |x_i - y_i|^2},$$

avec les notations : $x = (x_1, \ldots, x_n)$ et $y = (y_1, \ldots, y_n)$. Plus généralement, si $(E_1, \delta_1), \ldots, (E_n, \delta_n)$ sont des espaces métriques alors sur le produit cartésien $E = E_1 \times \ldots \times E_n$ les applications :

$$\begin{aligned} d_\infty : (x, y) \in E^2 &\longmapsto \max(\delta_1(x_1, y_1), \ldots, \delta_n(x_n, y_n)), \\ d_1 : (x, y) \in E^2 &\longmapsto \delta_1(x_1, y_1) + \ldots + \delta_n(x_n, y_n), \\ d_2 : (x, y) \in E^2 &\longmapsto \sqrt{\delta_1(x_1, y_1)^2 + \ldots + \delta_n(x_n, y_n)^2}, \end{aligned}$$

définissent des distances équivalentes. ◇

7.2 Normes et espaces vectoriels normés

Un cas particulier d'espace métrique (E, d) est constitué par les espaces vectoriels dits « normés » : l'ensemble E possède une structure d'espace vectoriel et la distance d est définie par une « norme ».

Rappelons[19] qu'un espace vectoriel sur un corps $(\mathbb{K}, +_\mathbb{K}, \times_\mathbb{K})$ (on parle de $\mathbb{K}$-espace vectoriel et on note $\mathbb{K}$-e.v.) est un ensemble E muni d'une *loi de composition interne* $+$:

$$(\mathbf{x}, \mathbf{y}) \in E \times E \longmapsto \mathbf{x} + \mathbf{y} \in E$$

[19] Voir le chap. 8 du *Cours de première année* consacré aux espaces vectoriels.

et muni d'une *loi de composition externe* $\cdot$ sur le corps $\mathbb{K}$:

$$(\alpha, \mathbf{x}) \in \mathbb{K} \times E \longmapsto \alpha \cdot \mathbf{x} \in E$$

telles que

1. $(E, +)$ est un groupe commutatif, *i.e.* :
   - la loi $+$ est associative,
   - il existe un élément neutre dans $(E, +)$,
   - tout élément de $(E, +)$ admet un symétrique,
   - la loi $+$ est commutative ;
2. la loi externe $\cdot$ possède les propriétés suivantes :

$$\begin{array}{llll} \forall \alpha \in \mathbb{K} & \forall (\mathbf{x}, \mathbf{y}) \in E^2 & \alpha \cdot (\mathbf{x} + \mathbf{y}) = \alpha \cdot \mathbf{x} + \alpha \cdot \mathbf{y}, \\ \forall (\alpha, \beta) \in \mathbb{K}^2 & \forall \mathbf{x} \in E & (\alpha +_{\mathbb{K}} \beta) \cdot \mathbf{x} = \alpha \cdot \mathbf{x} + \beta \cdot \mathbf{x}, \\ \forall (\alpha, \beta) \in \mathbb{K}^2 & \forall \mathbf{x} \in E & \alpha \cdot (\beta \cdot \mathbf{x}) = (\alpha \times_{\mathbb{K}} \beta) \cdot \mathbf{x}, \\ & \forall \mathbf{x} \in E & 1_{\mathbb{K}} \cdot \mathbf{x} = \mathbf{x}. \end{array}$$

Les éléments de E sont appelés *vecteurs* et les éléments de $\mathbb{K}$ sont appelés *scalaires.*

On considérera par la suite uniquement des espaces vectoriels sur le corps $\mathbb{R}$ ou $\mathbb{C}$. Si aucune ambiguïté n'est à craindre, on se contentera de parler d'espace vectoriel pour désigner un espace vectoriel sur le corps $\mathbb{R}$ ou sur le corps $\mathbb{C}$. Rappelons[20] les principaux exemples d'espaces vectoriels réels et complexes ($\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$)

- $\mathbb{K}[X]$ le $\mathbb{K}$-espace vectoriel des polynômes[21] à coefficients dans $\mathbb{K}$;
- $\mathcal{A}(X, \mathbb{K})$ le $\mathbb{K}$-espace vectoriel des applications[22] de $X \subset \mathbb{K}$ dans $\mathbb{K}$;
- $\mathcal{S}(\mathbb{K})$ le $\mathbb{K}$-espace vectoriel des suites numériques[23] ;
- $\Sigma(\mathbb{K})$ le $\mathbb{K}$-espace vectoriel des séries numériques[24] ;
- $\mathcal{M}_n(\mathbb{K})$ le $\mathbb{K}$-espace vectoriel[25] des matrices carrées $n \times n$ dont les éléments appartiennent à $\mathbb{K}$.

Rappelons[26] également qu'un sous-ensemble non vide F d'un $\mathbb{K}$-espace vectoriel E est qualifié de sous-espace vectoriel de E si

$$\forall (\alpha, \beta) \in \mathbb{K}^2 \quad \forall (x, y) \in F^2 \quad \alpha \cdot x + \beta \cdot y \in F.$$

---

[20] On pourra consulter le *Cours de première année* p. 298 et suivantes pour le détail de la vérification.

[21] Voir le *Cours de première année* p. 217 et p. 299.

[22] Voir le *Cours de première année* p. 299 et p. 559.

[23] Voir le *Cours de première année* p. 167 et p. 300.

[24] Voir le chap. 1 p. 2.

[25] Voir le *Cours de première année* p. 397.

[26] Voir la définition 8.5 p. 306 du *Cours de première année.*

En particulier un sous-espace vectoriel d'un $\mathbb{K}$-espace vectoriel est lui-même un $\mathbb{K}$-espace vectoriel.

Exemples Soit X un ensemble de réels.

1. L'ensemble $\mathcal{C}^0(X,\mathbb{R})$ des applications continues sur X à valeurs réelles est un sous-espace vectoriel de $\mathcal{A}(X,\mathbb{R})$.

2. L'ensemble $\mathcal{B}(X,\mathbb{R})$ des applications bornées sur X à valeurs réelles est un sous-espace vectoriel de $\mathcal{A}(X,\mathbb{R})$.

3. L'ensemble $\mathbb{K}_n[X]$ des polynômes de $\mathbb{K}[X]$ de degré au plus n est un sous-espace vectoriel de $\mathbb{K}[X]$.

4. L'ensemble $\ell^1(\mathbb{K})$ des séries numériques à coefficients dans $\mathbb{K}$, absolument convergentes, est un sous-espace vectoriel de $\Sigma(\mathbb{K})$.

DÉFINITION 7.3 (Norme sur un espace vectoriel)
On appelle norme *sur l'espace vectoriel E toute application N de E dans $\mathbb{R}$ vérifiant les quatre conditions suivantes :*

1. Pour tout $\mathbf{x} \in E$, $N(\mathbf{x}) \geqslant 0$;

2. $N(\mathbf{x}) = 0$ si et seulement si $\mathbf{x} = 0$;

3. pour tout $\mathbf{x} \in E$ et pour tout $\lambda \in \mathbb{K}$ on a $N(\lambda\mathbf{x}) = |\lambda| N(\mathbf{x})$;

4. pour tout $(\mathbf{x},\mathbf{y}) \in E^2$ on a $N(\mathbf{x}+\mathbf{y}) \leqslant N(\mathbf{x}) + N(\mathbf{y})$.

Un espace vectoriel normé *est défini par la donnée simultanée d'un espace vectoriel E et d'une norme N sur cet espace vectoriel. Il est noté (E,N).*

Remarques

1. Il est courant de noter $\|\cdot\|_E$ (ou plus simplement $\|\cdot\|$ lorsqu'il n'y a pas d'ambiguïté possible) une norme sur l'espace vectoriel E. Par ailleurs, le terme *espace vectoriel normé* est souvent écrit de manière abrégée *e.v.n.*.

2. On appelle *semi-norme* sur l'espace vectoriel E toute application N de E dans $\mathbb{R}^+$ vérifiant les conditions 1, 3 et 4 de la définition 7.3 mais pas nécessairement la condition 2.

3. Il est aisé de vérifier, voir l'exercice 7.6, que si N_1 et N_2 sont deux normes (resp. semi-normes) sur un espace vectoriel E, leur somme $N_1 + N_2$ définit elle aussi une norme (resp. semi-norme) sur E. On vérifiera également que si a désigne un réel strictement positif, alors $a\,N_1$ est également une norme (resp. semi-norme) sur E.

4. Si (E,N) est un espace vectoriel normé et si A est un sous-espace vectoriel de E alors la restriction à A de l'application N définit une norme sur A. On qualifie alors (A,N) de sous-espace vectoriel normé de (E,N).

5. On appelle *vecteur unitaire* tout vecteur $\mathbf{x}$ de l'espace vectoriel normé (E,N) tel que $N(\mathbf{x}) = 1$. ◇

Exercice 7.6
1 - Soit (E, N) un e.v.n. et a un réel strictement positif. Montrer que $\widetilde{N} : \mathbf{x} \in E \mapsto aN(\mathbf{x})$ est aussi une norme sur E.
2 - Soient E un espace vectoriel, N_1, N_2 deux normes sur E et a, b deux réels strictement positifs. Montrer que $N : \mathbf{x} \in E \mapsto a\,N_1(\mathbf{x}) + b\,N_2(\mathbf{x})$ est aussi une norme sur E.

Exemples

1. Il résulte des propriétés de la valeur absolue (resp. du module) que sur $\mathbb{R}$ (resp. sur $\mathbb{C}$) l'application $x \in \mathbb{K} \mapsto |x|$ est une norme.

2. Sur $\mathbb{K}[X]$, l'application N_1 qui au polynôme $P = \sum_{k=0}^{n} a_k X^k$ associe le réel positif $\sum_{k=0}^{n} |a_k|$ est une norme.

- Si $N_1(P) = 0$ alors pour tout $k \in \{0, \ldots, n\}$ on a $|a_k| = 0$ car $N_1(P)$ est une somme de termes positifs. Tous les coefficients du polynôme P étant nuls, P est le polynôme nul. Il est par ailleurs clair que si P est le polynôme nul alors $N_1(P) = 0$.
- Les deux autres conditions de la définition 7.3 résultent de la linéarité de la somme et des propriétés de la valeur absolue (ou du module) : $|\lambda a_k| = |\lambda||a_k|$ et $|a_k + b_k| \leqslant |a_k| + |b_k|$.

On vérifiera sans peine que l'application N_∞ qui au polynôme $P = \sum_{k=0}^{n} a_k X^k$ associe $\sup_{k\in\{0,\ldots,n\}} |a_k|$ est aussi une norme sur $\mathbb{K}[X]$.

3. Montrons que sur l'espace vectoriel $\ell^1(\mathbb{K})$ des séries numériques absolument convergentes, l'application N_1 qui à la série $u = \sum_n u_n$ associe le réel positif $\sum_{n=0}^{+\infty} |u_n|$ est une norme.

- Si $N_1(u) = 0$ alors $\sum_{n=0}^{+\infty} |u_n| = 0$; or la seule série à termes positifs dont la somme est nulle est la série nulle donc $|u_n| = 0$ pour tout $n \in \mathbb{N}$. Ceci implique que $u_n = 0$ pour tout $n \in \mathbb{N}$ et par conséquent que u est la série nulle. Il est par ailleurs clair que si u est la série nulle alors on a : $N_1(u) = 0$.
- La troisième condition de la définition 7.3 résulte des propriétés générales des séries (voir la proposition 1.3 p. 12).
- Enfin, si $u = \sum_n u_n$ et $v = \sum_n v_n$ sont deux séries absolument convergentes, pour tout $N \in \mathbb{N}$ on a :

$$\sum_{n=0}^{N} |u_n + v_n| \leqslant \sum_{n=0}^{N} \big(|u_n| + |v_n|\big) = \sum_{n=0}^{N} |u_n| + \sum_{n=0}^{N} |v_n|.$$

En passant à la limite quand N tend vers l'infini[27], on obtient : $N_1(u + v) \leqslant N_1(u) + N_1(v)$.

[27] Ces limites existent puisque les séries sont absolument convergentes.

On pourra vérifier que l'application $N_\infty : \sum_n u_n \in \ell^1(\mathbb{K}) \mapsto \sup_{n\in\mathbb{N}} |u_n|$ est aussi une norme sur $\ell^1(\mathbb{K})$.

DÉFINITION 7.4 (Normes équivalentes)
Deux normes N_1 et N_2 sur un même espace vectoriel E sont dites équivalentes *s'il existe deux réels strictement positifs C_1 et C_2 tels que :*

$$\forall \mathbf{x} \in E \qquad C_1\, N_1(\mathbf{x}) \leqslant N_2(\mathbf{x}) \leqslant C_2\, N_1(\mathbf{x}).$$

Remarque Pour montrer que deux normes N_1 et N_2 sur un espace vectoriel E ne sont pas équivalentes, il faut montrer que (c'est la négation de l'assertion quantifiée de la définition 7.4)

$$\begin{aligned} &\forall C_1 \in \mathbb{R}_+^* \quad \exists \mathbf{x} \in E \quad N_2(\mathbf{x}) < C_1 N_1(\mathbf{x}) \\ \text{ou} \quad &\forall C_2 \in \mathbb{R}_+^* \quad \exists \mathbf{x} \in E \quad N_2(\mathbf{x}) > C_2 N_1(\mathbf{x}). \end{aligned}$$

Signalons que dans le cas où deux normes ne sont pas équivalentes, il arrive souvent que l'une des deux inégalités intervenant dans la définition 7.4 soit tout de même vraie. Dans une telle situation, pour montrer que les deux normes ne sont pas équivalentes, il conviendra donc d'identifier au préalable l'inégalité qui n'est pas vraie. D'un point de vue pratique, pour un réel strictement positif C_2 fixé, il s'agit de déterminer, ou construire, un élément $\mathbf{x}$ de E (qui dépendra de C_2) pour lequel $N_2(\mathbf{x}) > C_2 N_1(\mathbf{x})$. Le second exemple détaillé ci-dessous illustre la méthode. ◇

Exemples

1. Sur l'espace vectoriel $\mathbb{K}_n[X]$ des polynômes $P = \sum_{k=0}^n a_k X^k$ à coefficients dans $\mathbb{K}$ qui sont de degré au plus n, les applications :

$$N_1 : P \in \mathbb{K}_n[X] \longmapsto \sum_{k=0}^{n} |a_k| \quad \text{et} \quad N_\infty : P \in \mathbb{K}_n[X] \longmapsto \max_{k\in\{0,\dots,n\}} |a_k|$$

définissent des normes équivalentes. Pour tout $P \in \mathbb{K}_n[X]$, on a d'une part $N_1(P) \leqslant (n+1) N_\infty(P)$ car $|a_k| \leqslant \sup_{j\in\{0,\dots,n\}} |a_j|$ pour tout $k \in \{0,\dots,n\}$ et d'autre part $N_\infty(P) \leqslant N_1(P)$ car $\sup_{j\in\{0,\dots,n\}} |a_j|$ est l'un des termes, tous positifs, de la somme définissant $N_1(P)$. On en déduit que

$$\forall P \in \mathbb{K}_n[X] \qquad \frac{1}{n+1} N_1(P) \leqslant N_\infty(P) \leqslant N_1(P).$$

2. Sur l'espace vectoriel $\mathbb{K}[X]$ des polynômes à coefficients dans $\mathbb{K}$, les normes N_1 et N_∞ de l'exemple précédent ne sont pas équivalentes. On a toujours $N_\infty(P) \leqslant N_1(P)$ pour tout $P \in \mathbb{K}[X]$ mais nous allons vérifier qu'il n'existe pas de réel positif C tel que pour tout $P \in \mathbb{K}[X]$ on ait[28] $N_1(P) \leqslant C\, N_\infty(P)$.

[28] On remarquera à ce propos que la constante obtenue dans l'exemple précédent avec $\mathbb{K}_n[X]$ dépend de n.

Considérons pour $n \in \mathbb{N}$, le polynôme $P_n = \sum_{k=0}^{n} X^k$; on a $N_\infty(P_n) = 1$ et $N_1(P_n) = n+1$. Quel que soit le réel strictement positif C, en prenant pour n la partie entière de C, on obtient $N_1(P_n) = n+1 = \mathrm{E}(C)+1 > C = C\, N_\infty(P_n)$. On en déduit que l'assertion « $\forall C \in \mathbb{R}_+^* \quad \exists P \in \mathbb{K}[X] \quad N_1(P) > C\, N_\infty(P)$ » est vraie, ce qui établit la non-équivalence des normes.

Remarque Nous établirons ultérieurement que sur un espace vectoriel de *dimension finie* toutes les normes sont équivalentes. ◇

Proposition 7.3 (Distance subordonnée à une norme)
Si (E, N) est un espace vectoriel normé, l'application

$$d : (\mathbf{x}, \mathbf{y}) \in E^2 \longmapsto N(\mathbf{x} - \mathbf{y}) \in \mathbb{R}^+$$

définit une distance sur E appelée distance subordonnée à la norme N.

Démonstration Il s'agit de montrer que l'application d vérifie les quatre conditions de la définition 7.1. Les deux premières conditions de la définition 7.3 impliquent que l'application d vérifie pour tout $(\mathbf{x}, \mathbf{y}) \in E^2$,

1. $d(\mathbf{x}, \mathbf{y}) = N(\mathbf{x} - \mathbf{y}) \geqslant 0$;
2. $d(\mathbf{x}, \mathbf{y}) = 0 \iff N(\mathbf{x} - \mathbf{y}) = 0 \iff \mathbf{x} - \mathbf{y} = 0 \iff \mathbf{x} = \mathbf{y}$.

En utilisant la troisième condition avec $\lambda = -1$, on obtient :

$$d(\mathbf{y}, \mathbf{x}) = N(\mathbf{y} - \mathbf{x}) = N(-(\mathbf{x} - \mathbf{y})) = |-1|\, N(\mathbf{x} - \mathbf{y}) = N(\mathbf{x} - \mathbf{y}) = d(\mathbf{x}, \mathbf{y}).$$

Enfin, en utilisant la quatrième condition de la définition 7.3, on établit que pour tout $(\mathbf{x}, \mathbf{y}, \mathbf{z}) \in E^3$ on a :

$$d(\mathbf{x}, \mathbf{y}) = N(\mathbf{x}-\mathbf{y}) = N(\mathbf{x}-\mathbf{z}+\mathbf{z}-\mathbf{y}) \leqslant N(\mathbf{x}-\mathbf{z}) + N(\mathbf{z}-\mathbf{y}) = d(\mathbf{x}, \mathbf{z}) + d(\mathbf{z}, \mathbf{y}).$$

L'application d satisfaisant aux conditions de la définition 7.1, il s'agit bien d'une distance. □

Remarque Sur un espace vectoriel E, deux normes N_1 et N_2 sont équivalentes si et seulement si les deux distances qui leur sont subordonnées sont équivalentes. Ceci résulte du fait que si d est la distance subordonnée à la norme N alors pour tout $(\mathbf{x}, \mathbf{y}) \in E^2$ on a $d(\mathbf{x}, \mathbf{y}) = N(\mathbf{x} - \mathbf{y})$ et $N(\mathbf{x}) = d(\mathbf{x}, \mathbf{0}_E)$. ◇

Sur un espace vectoriel, il peut exister des distances qui ne sont pas subordonnées à une norme. C'est le cas, par exemple, sur $\mathbb{R}$ de la distance $d : (x, y) \in \mathbb{R}^2 \mapsto |\arctan(x) - \arctan(y)|$. En effet, s'il existait une norme N sur $\mathbb{R}$ telle que d lui soit subordonnée, on aurait pour tous réels x, y la relation $N(x-y) = |\arctan(x) - \arctan(y)|$. En particulier, pour $x = 2$ et $y = 0$ on obtiendrait $N(2) = \arctan(2)$. Si N est une norme, la seconde relation de la définition 7.3 implique que $N(2) = 2N(1) = 2\arctan(1) = 2 \times \frac{\pi}{4} = \frac{\pi}{2}$. On aurait alors $\arctan(2) = \frac{\pi}{2}$, ce qui est absurde puisque la fonction arctangente est à valeurs dans $]-\frac{\pi}{2}, \frac{\pi}{2}[$.

Dans la suite du cours nous serons amenés à considérer les propriétés des fonc-

tions de plusieurs variables réelles et nous aurons recours aux normes sur $\mathbb{R}^n$. La proposition suivante indique les normes les plus utilisées sur $\mathbb{R}^n$ et $\mathbb{C}^n$; $|\cdot|$ désigne la valeur absolue ou le module selon que $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$. Étant donné un élément $\mathbf{x}$ de l'espace vectoriel $\mathbb{K}^n$, on désigne par $(x_1, \ldots, x_n)$ ses composantes.

Proposition 7.4 (Normes usuelles sur $\mathbb{K}^n$)
Pour tout $n \in \mathbb{N}^*$, les applications suivantes définissent des normes sur $\mathbb{K}^n$:

$$N_\infty : \mathbf{x} \in \mathbb{K}^n \longmapsto \max_{i \in \{1,\ldots,n\}} |x_i|,$$

$$N_1 : \mathbf{x} \in \mathbb{K}^n \longmapsto \sum_{i=1}^{n} |x_i|,$$

$$N_2 : \mathbf{x} \in \mathbb{K}^n \longmapsto \sqrt{\sum_{i=1}^{n} |x_i|^2}.$$

Ces normes sont équivalentes deux à deux et pour tout $\mathbf{x} \in \mathbb{K}^n$ on a :

$$N_\infty(\mathbf{x}) \leqslant N_1(\mathbf{x}) \leqslant \sqrt{n}\, N_2(\mathbf{x}) \leqslant n\, N_\infty(\mathbf{x}).$$

Démonstration ▷ Montrons que N_2 définit une norme sur $\mathbb{K}^n$. Elle est appelée la *norme euclidienne.*

1. Soit $\mathbf{x} \in \mathbb{K}^n$ tel que $N_2(\mathbf{x}) = 0$. Le réel $N_2(\mathbf{x})^2 = \sum_{i=1}^n |x_i|^2$ est une somme de carrés qui vaut 0. Chacun des termes de la somme est donc nul, autrement dit, pour $i \in \{1, \ldots, n\}$ on a $x_i = 0$. On en déduit que $\mathbf{x} = 0$. Inversement, si $\mathbf{x} = 0$ alors pour tout $i \in \{1, \ldots, n\}$ on a $x_i = 0$ et $N_2(\mathbf{x}) = \left(\sum_{i=1}^n |x_i|^2\right)^{\frac{1}{2}} = 0$.
2. Soient $\mathbf{x} \in \mathbb{K}^n$ et $\lambda \in \mathbb{K}$; on a

$$N_2(\lambda \mathbf{x})^2 = \sum_{i=1}^{n} |\lambda\, x_i|^2 = \sum_{i=1}^{n} |\lambda|^2\, |x_i|^2 = |\lambda|^2 \sum_{i=1}^{n} |x_i|^2 = |\lambda|^2\, N_2(\mathbf{x})^2.$$

   On en déduit que $N_2(\lambda \mathbf{x}) = |\lambda|\, N_2(\mathbf{x})$.
3. Soient $\mathbf{x}$ et $\mathbf{y}$ deux éléments de $\mathbb{K}^n$; montrer que $N_2(\mathbf{x}+\mathbf{y}) \leqslant N_2(\mathbf{x}) + N_2(\mathbf{y})$ revient à montrer que $N_2(\mathbf{x}+\mathbf{y})^2 \leqslant N_2(\mathbf{x})^2 + N_2(\mathbf{y})^2 + 2N_2(\mathbf{x})N_2(\mathbf{y})$. On a

$$N_2(\mathbf{x}+\mathbf{y})^2 = \sum_{i=1}^{n} |x_i + y_i|^2 = \sum_{i=1}^{n} |x_i|^2 + \sum_{i=1}^{n} |y_i|^2 + 2\sum_{i=1}^{n} (|x_i| \times |y_i|),$$

   et d'après l'inégalité de Cauchy-Schwarz[29]

$$\sum_{i=1}^{n} (|x_i| \times |y_i|) \leqslant \left(\sum_{i=1}^{n} |x_i|^2\right)^{\frac{1}{2}} \times \left(\sum_{i=1}^{n} |y_i|^2\right)^{\frac{1}{2}} = N_2(\mathbf{x}) \times N_2(\mathbf{y}).$$

[29] Voir la proposition 3.6 p. 101 du *Cours de première année.*

On a donc bien $N_2(\mathbf{x}+\mathbf{y})^2 \leqslant N_2(\mathbf{x})^2 + N_2(\mathbf{y})^2 + 2N_2(\mathbf{x})N_2(\mathbf{y})$.

$\trianglerighteq$ La vérification du fait que N_1 et N_∞ sont des normes sur $\mathbb{K}^n$ est laissée en exercice.

$\trianglerighteq$ Montrons que les trois normes N_1, N_2 et N_∞ sont équivalentes. La relation $N_\infty(\mathbf{x}) \leqslant N_1(\mathbf{x})$ est évidente car $N_1(\mathbf{x})$ est égal à la somme de $N_\infty(\mathbf{x})$ (le plus grand des réels positifs $|x_i|$ lorsque i varie dans $\{1,\dots,n\}$) et de réels positifs (les autres $|x_i|$). Par ailleurs, en utilisant l'inégalité de Cauchy-Schwarz[29] on obtient :

$$N_1(\mathbf{x})^2 = \left(\sum_{i=1}^{n} 1 \times |x_i|\right)^2 \leqslant \sum_{i=1}^{n} 1^2 \times \sum_{i=1}^{n} |x_i|^2 = n \times N_2(\mathbf{x})^2.$$

On a donc $N_1(\mathbf{x}) \leqslant \sqrt{n}\, N_2(\mathbf{x})$.

Enfin, $N_\infty(\mathbf{x})$ est le plus grand des réels positifs $|x_i|$ lorsque i varie dans $\{1,\dots,n\}$ donc la somme $\sum_{i=1}^{n} |x_i|^2$ est constituée de n termes tous inférieurs ou égaux à $N_\infty(\mathbf{x})^2$. On en déduit que

$$N_2(\mathbf{x}) = \sqrt{\sum_{i=1}^{n} |x_i|^2} \leqslant \sqrt{\sum_{i=1}^{n} N_\infty(\mathbf{x})^2} = \sqrt{n\, N_\infty(\mathbf{x})^2} = \sqrt{n}\; N_\infty(\mathbf{x}).$$

On en conclut que : $N_\infty(\mathbf{x}) \leqslant N_1(\mathbf{x}) \leqslant \sqrt{n}\, N_2(\mathbf{x}) \leqslant n\, N_\infty(\mathbf{x})$. □

Remarques

1. Les distances usuelles sur $\mathbb{R}^n$ définies à la proposition 7.2 sont les distances subordonnées aux normes définies à la proposition 7.4.

2. On peut montrer que pour tout réel $p \in [1, +\infty[$ l'application

$$N_p : \mathbf{x} \in \mathbb{K}^n \longmapsto \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}}$$

est une norme sur $\mathbb{K}^n$. ◇

Exercice 7.7 Soit $\|\cdot\|$ une norme sur $\mathbb{K}^n$ et A une matrice non nulle de $\mathcal{M}_n(\mathbb{K})$.
1 - Montrer que l'application $N : \mathbf{x} \in \mathbb{K}^n \mapsto \|\mathrm{A}\,\mathbf{x}\|$ est une semi-norme sur $\mathbb{K}^n$.
2 - Quelle propriété doit vérifier A pour que N soit une norme ?

Les normes sur $\mathbb{K}^n$ peuvent être utilisées pour définir des normes sur n'importe quel $\mathbb{K}$-espace vectoriel E de dimension n. Considérons une base $\mathcal{B} = (\mathbf{e_1},\dots,\mathbf{e_n})$ de E ; l'application qui à tout élément $\mathbf{x}$ de E associe ses

composantes $(x_1, \dots, x_n)$ dans la base $\mathcal{B}$ est un isomorphisme[30] de E dans $\mathbb{K}^n$. Cela implique que si N désigne une norme sur $\mathbb{K}^n$, l'application

$$N_E : \mathbf{x} \in E \longmapsto N(x_1, \dots, x_n)$$

est une norme sur E.

Exemple L'espace vectoriel $\mathbb{K}_n[X]$ des polynômes à coefficients dans $\mathbb{K}$ de degré inférieur ou égal à n est un espace vectoriel de dimension $n+1$. Dans la base canonique $\{0_{\mathbb{K}[X]}, X, \dots, X^n\}$ de $\mathbb{K}_n[X]$, un polynôme P de $\mathbb{K}_n[X]$ s'écrit $P = \sum_{k=0}^{n} a_k X^k$ où $a_0, \dots, a_n$ désignent les coefficients de P. Les applications suivantes sont donc des normes sur $\mathbb{K}_n[X]$:

$$N_\infty : P \in \mathbb{K}_n[X] \longmapsto \max_{k \in \{0,\dots,n\}} |a_k|,$$

$$N_1 : P \in \mathbb{K}_n[X] \longmapsto \sum_{k=0}^{n} |a_k|,$$

$$N_2 : P \in \mathbb{K}_n[X] \longmapsto \sqrt{\sum_{k=0}^{n} |a_k|^2}.$$

Ces normes sont équivalentes puisque d'après la proposition 7.4 on a pour tout $P \in \mathbb{K}_n[X]$,

$$N_\infty(P) \leqslant N_1(P) \leqslant \sqrt{n+1}\, N_2(P) \leqslant (n+1)\, N_\infty(P).$$

Ces trois applications définissent également des normes sur l'espace vectoriel $\mathbb{K}[X]$ des polynômes à coefficients dans $\mathbb{K}$ mais qu'elles ne sont pas équivalentes sur $\mathbb{K}[X]$, voir p. 360.

Par ailleurs, on déduit également de la proposition 7.4 que si $(E_1, N_1), \dots, (E_p, N_p)$ sont des espaces vectoriels normés alors les applications suivantes sont des normes sur le produit cartésien $E = E_1 \times \cdots \times E_p$:

$$\nu_0 : (x_1, \dots, x_p) \in E \longmapsto \max(N_1(x_1), \dots, N_p(x_p)),$$

$$\nu_1 : (x_1, \dots, x_p) \in E \longmapsto N_1(x_1) + \cdots + N_p(x_p),$$

$$\nu_2 : (x_1, \dots, x_p) \in E \longmapsto \sqrt{N_1(x_1)^2 + \cdots + N_p(x_p)^2}.$$

De plus, ces trois normes sont équivalentes sur E : pour tout $(x_1, \dots, x_p) \in E$ on a :

$$\nu_0(x_1, \dots, x_p) \leqslant \nu_1(x_1, \dots, x_p) \leqslant \sqrt{p}\, \nu_2(x_1, \dots, x_p) \leqslant p\, \nu_0(x_1, \dots, x_p).$$

[30] Autrement dit, une application linéaire bijective ; voir la proposition 8.9 p. 324 du *Cours de première année* pour une justification.

Un autre espace vectoriel de dimension finie qui est très souvent utilisé est celui des matrices. La proposition suivantes précise les normes usuelles sur $\mathcal{M}_{m,n}(\mathbb{K})$.

PROPOSITION 7.5 (Normes usuelles sur $\mathcal{M}_{m,n}(\mathbb{K})$)
Pour tous entiers naturels m, n non nuls, les applications suivantes définissent des normes sur $\mathcal{M}_{m,n}(\mathbb{K})$:

$$N_1 : \mathrm{A} \in \mathcal{M}_{m,n}(\mathbb{K}) \longmapsto \max_{1 \leqslant j \leqslant m} \sum_{i=1}^{n} |a_{ij}|,$$

$$N_s : \mathrm{A} \in \mathcal{M}_{m,n}(\mathbb{K}) \longmapsto \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{m} |a_{ij}|^2},$$

$$N_\infty : \mathrm{A} \in \mathcal{M}_{m,n}(\mathbb{K}) \longmapsto \max_{1 \leqslant i \leqslant n} \sum_{j=1}^{m} |a_{ij}|.$$

Démonstration La vérification que N_1 et N_∞ sont des normes fait l'objet de l'exercice 7.8. Montrons que N_s est une norme sur $\mathcal{M}_{m,n}(\mathbb{K})$. Elle est appelée *norme de Schur*[31]. Plutôt que de vérifier les 4 conditions de la définition 7.3, on peut avoir à nouveau recours à la proposition 7.4. La base canonique[32] de l'espace vectoriel $\mathcal{M}_{m,n}(\mathbb{K})$ des matrices à m lignes et n colonnes est formée par l'ensemble des matrices E_{ij} pour $i \in \{1, \ldots, m\}$ et $j \in \{1, \ldots, n\}$ où E_{ij} désigne la matrice dont tous les termes sont nuls à l'exception du terme placé à la i–ième ligne et à la j–ième colonne. Pour toute matrice $\mathrm{A} = (a_{ij})_{1 \leqslant i \leqslant m, 1 \leqslant j \leqslant n}$ de $\mathcal{M}_{m,n}(\mathbb{K})$ on a

$$A = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} \cdot E_{ij}.$$

On peut réécrire l'application N_s sous la forme

$$N_s : \mathrm{A} \in \mathcal{M}_{m,n}(\mathbb{K}) \longmapsto N_2\left((a_{ij})_{1 \leqslant i \leqslant m, 1 \leqslant j \leqslant n}\right),$$

où N_2 désigne la norme euclidienne de $\mathbb{K}^{mn}$. Comme l'application qui à une matrice de $\mathcal{M}_{m,n}(\mathbb{K})$ associe ses composantes dans la base canonique de $\mathcal{M}_{m,n}(\mathbb{K})$ est un isomorphisme, on en conclut que N_s est une norme sur $\mathcal{M}_{m,n}(\mathbb{K})$. □

[31] Issai SCHUR (1875-1941)
[32] Voir le *Cours de première année* p. 398.

Exercice 7.8

1 - Montrer que les applications N_1 et N_∞ définies à la proposition 7.5 sont des normes sur $\mathcal{M}_{m,n}(\mathbb{K})$.

2 - Montrer, en utilisant l'inégalité de Cauchy-Schwarz[33], que si $A = (a_{ij})_{i,j}$ et $B = (b_{ij})_{i,j}$ sont deux matrices de $\mathcal{M}_n(\mathbb{K})$ alors

$$N_s(A \times B) \leqslant N_s(A) \times N_s(B).$$

On sera attentif au fait que cette relation portant sur la norme du produit n'est pas une propriété que possèdent toutes les normes.

Nous avons utilisé au cours de la démonstration de la proposition 7.5 l'isomorphisme entre $\mathcal{M}_{m,n}(\mathbb{K})$ et $\mathbb{K}^{mn}$ pour relier la norme de Schur sur $\mathcal{M}_{m,n}(\mathbb{K})$ à la norme euclidienne sur $\mathbb{K}^{mn}$. Indiquons quelques liens supplémentaires entre normes sur l'ensemble des matrices et normes sur l'ensemble des vecteurs.

✗ Si on identifie le vecteur $\mathbf{x}$ de $\mathbb{K}^n$ au vecteur colonne $X \in \mathcal{M}_{n,1}(\mathbb{K})$ alors $\|\mathbf{x}\|_\infty = N_\infty(X)$ où N_∞ désigne la norme infinie dans $\mathcal{M}_{n,1}(\mathbb{K})$ et $\|\mathbf{x}\|_\infty = \max_{i\in\{1,\dots,n\}} |x_i|$ désigne la norme infinie de $\mathbf{x}$.

✗ Par ailleurs, pour tout $A \in \mathcal{M}_n(\mathbb{K})$ et pour tout $\mathbf{x} \in \mathbb{K}^n$ on a

$$\begin{aligned}\|A\,\mathbf{x}\|_\infty &= \max_{i\in\{1,\dots,m\}} \Big|\sum_{j=1}^{n} a_{ij}x_j\Big| \leqslant \max_{i\in\{1,\dots,m\}} \sum_{j=1}^{n} |a_{ij}|\ |x_j| \\ &\leqslant \|\mathbf{x}\|_\infty \max_{i\in\{1,\dots,m\}} \sum_{j=1}^{n} |a_{ij}| = \|\mathbf{x}\|_\infty\, N_\infty(A),\end{aligned}$$

où $N_\infty(A)$ désigne ici la norme infinie de A dans $\mathcal{M}_n(\mathbb{K})$.

✗ On appelle *rayon spectral* de la matrice $A \in \mathcal{M}_n(\mathbb{C})$ et on note $\rho(A)$ le réel positif égal au module des valeurs propres de A le plus grand, *i.e.*

$$\rho(A) = \max\big\{|\lambda| \ ;\ \exists \mathbf{x} \in \mathbb{C}^n \quad \mathbf{x} \neq \mathbf{0} \quad \text{et} \quad A\mathbf{x} = \lambda\mathbf{x}\big\}.$$

En notant λ_1 une valeur propre de A telle que $\rho(A) = |\lambda_1|$ et $\mathbf{x}_1$ un vecteur propre associé, on obtient

$$|\lambda_1|\,\|\mathbf{x}_1\|_\infty = \|\lambda_1\,\mathbf{x}_1\|_\infty = \|A\,\mathbf{x}_1\|_\infty \leqslant N_\infty(A)\,\|\mathbf{x}_1\|_\infty.$$

On en déduit que $\rho(A) = |\lambda_1| \leqslant N_\infty(A)$. Le rayon spectral d'une matrice est donc toujours inférieur à la norme infinie de cette matrice.

Donnons à présent un résultat indiquant que sur $\mathbb{R}^n$ et sur $\mathbb{C}^n$ ($n \in \mathbb{N}^*$), toutes le normes sont équivalentes deux à deux. Sa démonstration, qui utilise la notion de suite de Cauchy, sera donnée ultérieurement (voir p. 406).

[33] Voir la proposition 3.6 p. 101 du *Cours de première année*.

THÉORÈME 7.1 *Toutes les normes sur $\mathbb{K}^n$ $(n \in \mathbb{N}^*)$ sont équivalentes.*

Sur $\mathbb{K}^n$, toutes les normes sont équivalentes mais toutes les distances ne le sont pas. Par exemple, sur $\mathbb{R}^2$ la distance euclidienne d_2 n'est pas équivalente à la distance[34] d définie par :

$$d(\mathbf{x},\mathbf{y}) = d_2(\mathbf{x},\mathbf{y}) \quad \text{si } \mathbf{x},\mathbf{y} \text{ et l'origine sont alignés}$$
$$d(\mathbf{x},\mathbf{y}) = d_2(\mathbf{x},\mathbf{0}) + d(\mathbf{0},\mathbf{y}) \quad \text{sinon.}$$

On a $d_2 \leqslant d$ mais pour tout réel strictement positif C les éléments $\mathbf{x} = (C,0)$ et $\mathbf{y} = (C,1)$ vérifient : $d_2(\mathbf{x},\mathbf{y}) = 1$ et $d(\mathbf{x},\mathbf{y}) = C + \sqrt{1+C^2}$. On a donc $d(\mathbf{x},\mathbf{y}) > C\, d_2(\mathbf{x},\mathbf{y})$.

On peut toutefois déduire du théorème 7.1 que sur $\mathbb{K}^n$, toutes les distances subordonnées à des normes sont équivalentes.

Plus généralement, on a le résultat suivant qui est un corollaire du théorème 7.1.

THÉORÈME 7.2 *Sur un $\mathbb{K}$-espace vectoriel de dimension finie, toutes les normes sont équivalentes deux à deux.*

Démonstration Soit E un $\mathbb{K}$-espace vectoriel de dimension n dont la base canonique est notée $(\mathbf{e}_1,\ldots,\mathbf{e}_n)$. Considérons l'isomorphisme ϕ qui à un élément $\mathbf{x} = \sum_{k=1}^{n} x_i \mathbf{e}_i$ de E associe ses composantes $(x_1,\ldots,x_n)$ dans la base canonique :

$$\phi : \mathbf{x} \in E \longmapsto (x_1,\ldots,x_n) \in \mathbb{K}^n.$$

À deux normes N_1 et N_2 sur E, on peut associer les applications

$$\nu_1 : (x_1,\ldots,x_n) \in \mathbb{K}^n \longmapsto N_1(\phi^{-1}(x_1,\ldots,x_n)) \in \mathbb{R}^+$$
$$\text{et} \quad \nu_2 : (x_1,\ldots,x_n) \in \mathbb{K}^n \longmapsto N_2(\phi^{-1}(x_1,\ldots,x_n)) \in \mathbb{R}^+$$

qui sont des normes sur $\mathbb{K}^n$. D'après le théorème 7.1, elles sont équivalentes : il existe donc $C_1, C_2 \in \mathbb{R}_+^*$ tels que pour tout $(x_1,\ldots,x_n) \in \mathbb{K}^n$

$$C_1\, \nu_1(x_1,\ldots,x_n) \leqslant \nu_2(x_1,\ldots,x_n) \leqslant C_2\, \nu_1(x_1,\ldots,x_n),$$

autrement dit tels que,

$$C_1\, N_1(\phi^{-1}(x_1,\ldots,x_n)) \leqslant N_2(\phi^{-1}(x_1,\ldots,x_n)) \leqslant C_2\, N_1(\phi^{-1}(x_1,\ldots,x_n)).$$

On en conclut que pour tout $\mathbf{x} \in E$, on a $C_1\, N_1(\mathbf{x}) \leqslant N_2(\mathbf{x}) \leqslant C_2\, N_1(\mathbf{x})$; les deux normes N_1 et N_2 sont donc équivalentes. □

[34] Cette distance est assimilable à la « distance SNCF » entre gares, le terme « aligné » signifiant dans ce cas « sur la même ligne ferroviaire ».

Sur un espace vectoriel qui n'est pas de dimension finie, cette propriété n'est pas vraie. Certaines normes peuvent être équivalentes, mais toutes les normes ne sont pas équivalentes deux à deux. Nous avons par exemple vu p. 360 que sur l'espace vectoriel $\mathbb{K}[X]$ les normes N_1 et N_∞ n'étaient pas équivalentes.

Corollaire 7.1 *Soient E un espace vectoriel sur $\mathbb{K}$ de dimension n et $\mathcal{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ une base de E. Toute norme sur E est équivalente à la norme infinie définie pour tout vecteur $\mathbf{x}$ de E de composantes $(x_1, \ldots, x_n)$ dans la base $\mathcal{B}$ par $\|\mathbf{x}\|_\infty = \max_{1 \leqslant i \leqslant n} |x_i|$*

Terminons cette partie consacrée aux espaces vectoriels normés en donnant les normes usuelles sur l'espace vectoriel $\mathcal{C}^0([a,b], \mathbb{R})$ des applications continues de $[a,b]$ dans $\mathbb{R}$.

Proposition 7.6 (Normes usuelles sur $\mathcal{C}^0([a,b], \mathbb{R})$)
Les applications suivantes sont des normes sur l'espace vectoriel $\mathcal{C}^0([a,b], \mathbb{R})$ des applications continues de $[a,b]$ dans $\mathbb{R}$:

$$N_1 : f \in \mathcal{C}^0([a,b], \mathbb{R}) \longmapsto \int_a^b |f(x)| \, \mathrm{d}x,$$

$$N_2 : f \in \mathcal{C}^0([a,b], \mathbb{R}) \longmapsto \sqrt{\int_a^b |f(x)|^2 \, \mathrm{d}x}$$

$$N_\infty : f \in \mathcal{C}^0([a,b], \mathbb{R}) \longmapsto \sup_{x \in [a,b]} |f(x)|.$$

L'application N_1 est appelée norme de la convergence en moyenne, *N_2 norme de la convergence* quadratique *et N_∞ norme de la convergence* uniforme.

Démonstration Il résulte des propriétés de la borne supérieure (voir la proposition 2.1 p. 68) que l'application N_∞ est une norme. La vérification du fait que N_1 et N_2 sont des normes est laissée en exercice ; on pourra s'inspirer des exemples de la p. 348. □

Remarque L'application $N_\infty : f \in \mathcal{B}([a,b], \mathbb{R}) \mapsto \sup_{x \in [a,b]} |f(x)|$ définit aussi une norme sur l'ensemble $\mathcal{B}([a,b], \mathbb{R})$ des applications bornées de $[a,b]$ dans $\mathbb{R}$. Par contre, une application bornée n'étant pas nécessairement intégrable, les fonctions N_1 et N_2 ne sont pas des normes sur $\mathcal{B}([a,b], \mathbb{R})$. ◇

7.3 Topologie dans un espace métrique

Nous allons étendre au cas d'un espace métrique (E, d) quelconque un certain nombre de notions de topologie[35] qui ont été introduites dans le *Cours de première année* dans le cas de la droite réelle[36] puis dans le chap. 4 de cet ouvrage dans le cas du plan complexe. Toutes les notions topologiques introduites dans cette partie restent bien entendu valables dans le cas d'un espace vectoriel normé (E, N) en considérant sur E la distance d subordonnée à la norme N :

$$d : (\mathbf{x}, \mathbf{y}) \in E \longmapsto N(\mathbf{y} - \mathbf{x}).$$

7.3.1 Ensembles ouverts et ensembles fermés

DÉFINITION 7.5 *Soient (E, d) un espace métrique, x_0 un élément de E et r un réel strictement positif.*

✗ *On appelle* boule ouverte *de centre x_0 et de rayon r l'ensemble*

$$B_d(x_0, r) = \{x \in E \ ; \ d(x, x_0) < r\}.$$

✗ *On appelle* boule fermée *de centre x_0 et de rayon r l'ensemble*

$$\overline{B}_d(x_0, r) = \{x \in E \ ; \ d(x, x_0) \leqslant r\}.$$

Remarques

1. La notion de boule est étroitement liée à la distance d dont est muni l'ensemble E. Si δ est une autre distance sur E, les ensembles $B_d(x_0, r)$ et $B_\delta(x_0, r)$ sont distincts (même si les distances sont équivalentes). Lorsqu'aucune confusion n'est à craindre concernant la distance utilisée, nous noterons plus simplement $B(x_0, r)$ la boule ouverte de centre x_0 et de rayon r.

2. Dans le cas où l'espace métrique (E, d) est l'ensemble des complexes muni de la distance subordonnée au module, *i.e.* $d : (z, z') \in \mathbb{C}^2 \mapsto |z - z'|$, la définition 7.5 correspond à la définition 4.1 p. 164. Plus généralement, on pourra relier les définitions données dans cette partie aux notions concernant la topologie du plan complexe étudiées au chap. 4. ◇

Exemples

1. La figure 3 donne les représentations graphiques des boules unités (*i.e.* les boules de centre l'origine du repère et de rayon $r = 1$) dans $\mathbb{R}^2$ et dans $\mathbb{R}^3$ associées aux distances d_1, d_2 et d_∞ définies à la proposition 7.2. Par exemple,

[35] Le terme « topologie » a pour signification en grec : discours du lieu. L'origine de la topologie est liée aux problèmes qu'ont posés la généralisation de l'étude des propriétés de continuité et de dérivabilité pour les fonctions de plusieurs variables.

[36] Voir la section 3.3 p. 109 du *Cours de première année*.

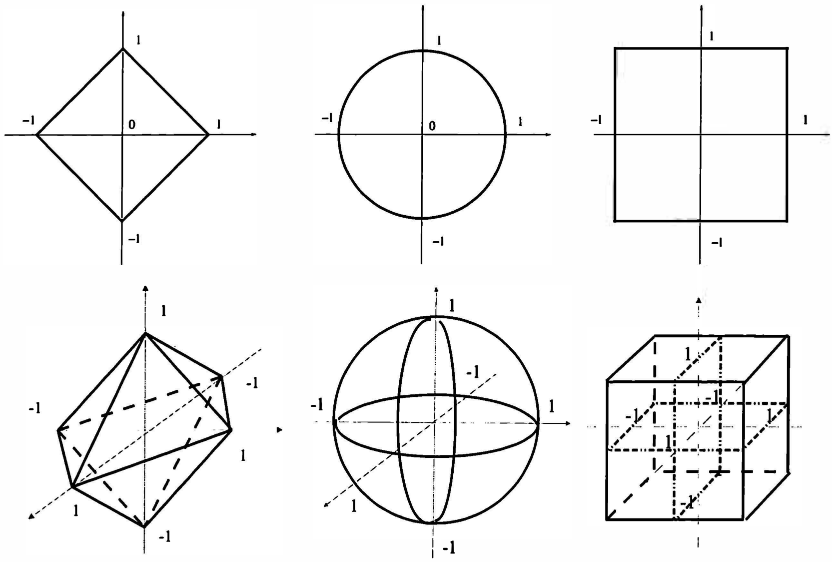


Fig. 3 Boules unités dans $\mathbb{R}^2$ et dans $\mathbb{R}^3$ pour les distances d_1, d_2 et d_∞.

dans $\mathbb{R}^2$ on a

$$B_{d_2}(0,1) = \big\{(x_1,x_2) \in \mathbb{R}^2 \ ; \ (x_1^2+x_2^2)^{\frac{1}{2}} < 1\big\} = \big\{(x_1,x_2) \in \mathbb{R}^2 \ ; \ x_1^2+x_2^2 < 1\big\}.$$

La boule unité pour la distance euclidienne dans $\mathbb{R}^2$ est donc le disque de centre l'origine et de rayon 1. Dans $\mathbb{R}^3$, la boule unité pour la distance euclidienne est la « boule » de centre l'origine et de rayon 1. On a par ailleurs,

$$\begin{aligned} B_{d_\infty}(0,1) &= \big\{(x_1,x_2) \in \mathbb{R}^2 \ ; \ \max(|x_1|,|x_2|) < 1\big\} \\ &= \big\{(x_1,x_2) \in \mathbb{R}^2 \ ; \ -1 < x_1 < 1 \text{ et } -1 < x_2 < 1\big\}. \end{aligned}$$

La boule unité pour la distance d_∞ dans $\mathbb{R}^2$ est donc le carré $]-1,1[\times]-1,1[$. Dans $\mathbb{R}^3$, la boule unité pour la distance euclidienne est le cube $]-1,1[^3$. Enfin, on pourra vérifier que la boule unité pour la distance d_1 dans $\mathbb{R}^2$ est un losange et que dans $\mathbb{R}^3$ il s'agit d'un octaèdre.

Signalons que dans $\mathbb{R}$ la boule ouverte (resp. fermée) de centre x_0 et de rayon r pour ces trois distances est l'intervalle ouvert $]x_0-r,x_0+r[$ (resp. fermé $[x_0-r,x_0+r]$).

2. Dans $\mathcal{C}^0([a,b],\mathbb{R})$ muni de la distance

$$d_\infty : (f,g) \in \mathcal{C}^0([a,b],\mathbb{R}) \times \mathcal{C}^0([a,b],\mathbb{R}) \longmapsto \sup_{x\in[a,b]} |f(x)-g(x)|,$$

la boule ouverte de centre $f \in \mathcal{C}^0([a,b],\mathbb{R})$ et de rayon $r \in \mathbb{R}_+^*$ correspond à l'ensemble des fonctions continues sur $[a,b]$ dont la représentation graphique se trouve dans le domaine du plan limité par les représentations des applications $f_- : x \in [a,b] \mapsto f(x)-r$ et $f_+ : x \in [a,b] \mapsto f(x)+r$, voir la figure 4.

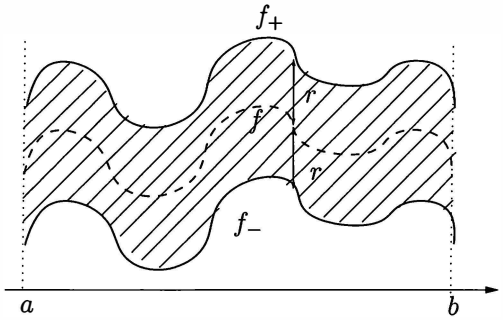


Fig. 4 Dans $C^0([a,b],\mathbb{R})$, représentation de la boule de centre f et de rayon r.

On sera attentif à ne pas confondre au niveau des notations la boule fermée de centre x_0 et de rayon r notée $\overline{B}(x_0,r)$ et le complémentaire dans E de la boule ouverte de centre x_0 et de rayon r que nous noterons $\complement_E B(x_0,r)$ ou $\complement B(x_0,r)$ quand il n'y aura pas de confusion possible sur l'ensemble E.

DÉFINITION 7.6 (Voisinage d'un point)
Soient (E,d) un espace métrique et x_0 un élément de E. On dit que le sous-ensemble A de E est un voisinage de x_0 dans E si A contient une boule ouverte de centre x_0, i.e. si

$$\exists r \in \mathbb{R}_+^* \quad B(x_0,r) \subset A.$$

On note $\mathcal{V}(x_0)$ l'ensemble des voisinages de x_0 dans E.

Remarques

1. L'intersection et la réunion de deux voisinages de l'élément x_0 sont aussi des voisinages de x_0.
2. Toute boule (ouverte ou fermée) de centre x_0 est un voisinage de x_0.
3. L'ensemble E est voisinage de tous ses éléments. ◇

DÉFINITION 7.7 (Ensemble ouvert et ensemble fermé)
✗ *Soit (E,d) un espace métrique. Un sous-ensemble non vide A de E est qualifié d'*ensemble ouvert *si pour tout élément x de A il existe une boule ouverte de centre x contenue dans A, i.e. si*

$$\forall x \in A \quad \exists r \in \mathbb{R}_+^* \quad B(x,r) \subset A.$$

✗ *Un sous-ensemble A de E est appelé* ensemble fermé *si son complémentaire dans E est un ensemble ouvert.*

Remarques

1. Il résulte de la définition 7.6 qu'un ensemble est ouvert s'il est un voisinage de chacun des ses points.

2. Par convention, l'ensemble vide est un ouvert de E. Son complémentaire qui est l'ensemble E est donc un fermé. Nous avons mentionné que l'ensemble E est voisinage de chacun de ses points ; c'est donc aussi un ouvert. Le complémentaire de l'ensemble E qui est l'ensemble vide est donc un fermé. On retiendra que l'ensemble vide et l'ensemble E sont à la fois des ouverts et des fermés.

3. Si (A, d) est un sous-espace métrique de (E, d) alors les ensembles ouverts dans A sont les intersections de A avec les ouverts de E. ◇

Exemples

1. Soit (E, d_E) est un espace métrique et x_0 un élément de E. Le singleton $\{x_0\}$ est un sous-ensemble fermé de E. En effet, son complémentaire qui est l'ensemble $A = \{x \in E \ ; \ x \neq x_0\}$ est ouvert puisque pour tout $x \in A$ la boule de centre x et de rayon $r = \frac{1}{2} d_E(x, x_0) > 0$ ne contient pas x_0 donc est incluse dans A.

2. Munissons l'ensemble E de la distance $d : (x, y) \in E \mapsto \begin{cases} 1 & \text{si } x = y \\ 0 & \text{sinon} \end{cases}$ appelée *distance discrète*. Pour tout $x_0 \in E$ on a $B(x_0, \frac{1}{2}) = \{x_0\}$ donc pour cette distance le singleton $\{x_0\}$ est un sous-ensemble ouvert de E (il est donc à la fois ouvert et fermé dans E compte tenu de ce qui a été établi à l'exemple précédent).

3. Soit (E, d_E) est un espace métrique et $D = \{(y_1, y_2) \in E^2 \ ; \ y_1 = y_2\}$. Dans E^2 muni de la distance produit[37]

$$d_1 : ((x_1, x_2), (y_1, y_2)) \in E^2 \mapsto d_E(x_1, y_1) + d_E(x_2, y_2),$$

l'ensemble D est fermé. Pour le montrer, établissons que le complémentaire de D est un ensemble ouvert. Il s'agit de montrer que si $y = (y_1, y_2)$ désigne un élément quelconque de $\complement D$ alors il existe un réel $\alpha > 0$ tel que la boule $B(y, \alpha)$ soit incluse dans $\complement D$. Notons $\beta = d_E(y_1, y_2)$ (qui est strictement positif car $y_1 \neq y_2$ puisque $y \in \complement D$) et montrons que pour $\alpha = \frac{1}{2}\beta$ on a $B(y, \alpha) \subset \complement D$. Considérons un élément quelconque $z = (z_1, z_2)$ de $B(y, \alpha)$ et montrons que $z \in \complement D$, *i.e.* que $z_1 \neq z_2$. On a $d_1(y, z) = d_E(y_1, z_1) + d_E(y_2, z_2) < \frac{1}{2}\beta$ et par l'inégalité triangulaire

$$d_E(y_1, y_2) \leqslant d_E(y_1, z_1) + d_E(z_1, z_2) + d_E(z_2, y_2).$$

On en déduit que

$$d_E(z_1, z_2) \geqslant d_E(y_1, y_2) - \underbrace{(d_E(y_1, z_1) + d_E(z_2, y_2))}_{= d_1(y, z)} > \beta - \tfrac{1}{2}\beta = \tfrac{1}{2}\beta.$$

[37] Les différentes distances produits données p. 356 sont équivalentes ; le choix fait ici facilite grandement le raisonnement.

On en conclut que $d_E(z_1, z_2)$ est strictement positif ce qui implique que z_1 est nécessairement distinct de z_2.

Un ensemble peut être ni ouvert ni fermé. Contrairement à son sens dans le langage courant, « fermé » n'est pas le contraire d' « ouvert ». Par ailleurs, un ensemble peut être à la fois ouvert et fermé. Par exemple, dans $\mathbb{R}^n$ muni de la distance discrète (voir p. 346), la boule ouverte $B(x_0, 1)$ est réduite au singleton $\{x_0\}$ qui est un ouvert et un fermé pour cette distance.

PROPOSITION 7.7 *Dans un espace métrique (E, d), toute boule ouverte est un ensemble ouvert et toute boule fermée est un ensemble fermé*[38].

Démonstration $\trianglerighteq$ Considérons la boule ouverte $B(x_0, r_0)$ où $x_0 \in E$ et $r_0 \in \mathbb{R}_+^*$ et un élément x_1 quelconque de $B(x_0, r_0)$; il s'agit de montrer que l'on peut trouver une boule ouverte de centre x_1 incluse dans $B(x_0, r_0)$, voir le dessin de gauche de la figure 5. Si $r_1 = r_0 - d(x_1, x_0)$ alors pour tout $x \in B(x_1, r_1)$, l'inégalité triangulaire[39] nous donne :

$$d(x, x_0) \leqslant d(x, x_1) + d(x_1, x_0) < r_1 + d(x_1, x_0) = r_0.$$

Cela signifie que tous les éléments x de $B(x_1, r_1)$ appartiennent à $B(x_0, r_0)$ et ceci quel que soit le choix de x_1 dans $B(x_0, r_0)$. D'après la définition 7.7, l'ensemble $B(x_0, r_0)$ est un ensemble ouvert.

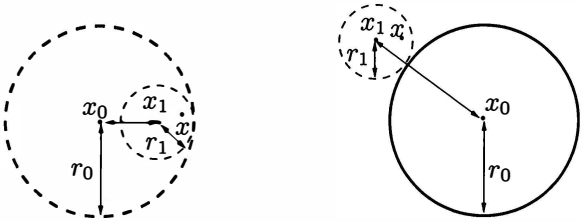


Fig. 5 Les deux situations considérées.

$\trianglerighteq$ Considérons la boule fermée $\overline{B}(x_0, r_0)$ où $x_0 \in E$ et $r_0 \in \mathbb{R}_+^*$. Pour montrer que $\overline{B}(x_0, r_0)$ est un ensemble fermé, il faut montrer que son complémentaire $\complement\overline{B}(x_0, r_0)$ est un ensemble ouvert de E. Considérons un élément x_1 quelconque de $\complement\overline{B}(x_0, r_0)$; il s'agit de montrer que l'on peut trouver une boule ouverte de centre x_1 incluse dans $\complement\overline{B}(x_0, r_0)$, voir le dessin de droite de la figure 5. Soit $r_1 = d(x_1, x_0) - r_0$; pour tout $x \in B(x_1, r_1)$, l'inégalité triangulaire nous donne :

$$d(x_1, x_0) \leqslant d(x_1, x) + d(x, x_0) < r_1 + d(x, x_0) = d(x_1, x_0) - r_0 + d(x, x_0).$$

[38] Ce qui justifie la terminologie employée.

[39] Voir la définition 7.1 p. 345.

On en déduit que $d(x, x_0) > r_0$ ce qui signifie que tous les éléments x de $B(x_1, r_1)$ appartiennent à $\complement\overline{B}(x_0, r_0)$ et ceci quel que soit le choix de x_1 dans $\complement\overline{B}(x_0, r_0)$. L'ensemble $\complement\overline{B}(x_0, r_0)$ est donc bien un ensemble ouvert ce qui implique que la boule fermée $\overline{B}(x_0, r_0)$ est un ensemble fermé. □

La notion d'ensemble ouvert dans un espace métrique (E, d) dépend de la distance d. En général, un même ensemble E possède des sous-ensembles ouverts et fermés différents si on le munit successivement de deux distances différentes.

Les ouverts définis sur un ensemble E par une distance d, forment ce qui est appelé la « topologie » induite par la distance d. Deux distances d et δ définissent la même topologie sur E si les sous-ensembles qui sont ouverts pour d le sont aussi pour δ, et inversement[(40)]. Dans ce cas on dit que les deux distances sont « topologiquement équivalentes ».

La proposition suivante indique que deux distances d et δ, équivalentes au sens de la définition 7.2, sont topologiquement équivalentes et induisent la même topologie sur E (les ensembles ouverts dans (E, d) et (E, δ) sont les mêmes).

Proposition 7.8 *Soient (E, d) un espace métrique et A un sous-ensemble de E, ouvert dans (E, d). Si δ est une distance équivalente à d alors A est aussi un ouvert dans (E, δ).*

Démonstration Pour montrer que A est un ouvert dans (E, δ), il faut montrer qu'étant donné un élément x de A, on peut trouver un réel ρ strictement positif tel que la boule $B_\delta(x, \rho)$ soit incluse dans A. D'après la définition 7.7, puisque A est ouvert dans (E, d),

$$\exists r \in \mathbb{R}_+^* \quad \forall y \in E \quad \big(d(x, y) < r \implies y \in A\big).$$

Comme d et δ sont deux distances équivalentes, il existe $C \in \mathbb{R}_+^*$ tel que $d(x, y) \leqslant C\,\delta(x, y)$. On en déduit que si $\delta(x, y) < r/C$ alors $d(x, y) < r$ et $y \in A$. On en conclut qu'en prenant $\rho = r/C$, on a

$$\forall y \in E \quad \big(\delta(x, y) < \rho \implies y \in A\big),$$

autrement dit que la boule $B_\delta(x, \rho)$ est incluse dans A. □

Un corollaire de la proposition 7.8 est que si un ensemble est fermé dans (E, d) alors il est fermé dans E muni d'une distance équivalente à d.

La proposition 7.8 n'admet pas de réciproque : deux distances peuvent être topologiquement équivalentes sans être équivalentes au sens de la définition 7.2. Par exemple, sur $\mathbb{R}$ la distance $d : (x, y) \in \mathbb{R}^2 \mapsto |\arctan(x) - \arctan(y)|$

(40) Dans ce cas les ensembles fermés sont aussi les mêmes puisqu'ils sont définis comme complémentaires d'ensembles ouverts.

n'est pas équivalente à la distance usuelle $\delta : (x,y) \in \mathbb{R}^2 \mapsto |x-y|$, voir p. 352. Toutefois, ces deux distances sont topologiquement équivalentes : on peut montrer que les ensembles ouverts dans $\mathbb{R}$ muni de la distance d sont également les ensembles ouverts dans $\mathbb{R}$ muni de la distance δ.

Remarques

1. Le cas des espaces vectoriels normés est un peu particulier. Les distances associées à deux normes qui ne sont pas équivalentes ne peuvent pas engendrer une même topologie.

2. Pour montrer que deux topologies induites sur un ensemble E par deux distances d et δ sont distinctes, il suffit de montrer qu'un sous-ensemble de E est ouvert (ou fermé, ou ni ouvert ni fermé) avec l'une des distances mais pas avec l'autre. ◇

LEMME 7.1 *Tout ouvert d'un espace métrique (E,d) est réunion de boules ouvertes.*

Démonstration Si A est un ensemble non vide et ouvert dans l'espace métrique (E,d), on peut associer, par définition, à tout élément a de A un réel $r_a > 0$ tel que $B(a,r_a) \subset A$. On a donc $\bigcup_{a\in A} B(a,r_a) \subset A$.
Par ailleurs, puisque $a \in B(a,r_a)$ on a aussi $A = \bigcup_{a\in A}\{a\} \subset \bigcup_{a\in A} B(a,r_a)$.
On en déduit que $A = \bigcup_{a\in A} B(a,r_a)$. □

D'un point de vue pratique, comme tout ensemble ouvert est réunion de boules ouvertes, pour étudier les propriétés topologiques d'un espace métrique, on peut se contenter de considérer les boules ouvertes de cet espace.

La proposition suivante indique que le lemme 7.1 admet une réciproque : la réunion de boules ouvertes est un ouvert.

PROPOSITION 7.9 (Union et intersection d'ouverts ou de fermés)

✗ *L'union d'un nombre quelconque*[41] *d'ensembles ouverts est un ensemble ouvert. L'intersection d'un nombre* fini *d'ensembles ouverts est un ensemble ouvert.*

✗ *L'intersection d'un nombre quelconque d'ensembles fermés est un ensemble fermé. L'union d'un nombre* fini *d'ensembles fermés est un ensemble fermé.*

Démonstration ≥ Soit $(O_i)_{i\in I}$ ($I \subset \mathbb{R}$) une famille (non nécessairement dénombrable) d'ensembles ouverts et soit O la réunion de ces ensembles. Un élément x quelconque de O appartient nécessairement à au moins un des ensembles O_i. Supposons que x appartienne à l'ensemble O_{i_0}, voir le dessin de gauche de la figure 6. Comme O_{i_0} est un ensemble ouvert, on peut trouver

[41] Fini, dénombrable ou infini non dénombrable.

pour chacun de ses éléments une boule ouverte centrée en l'élément considéré qui soit incluse dans O_{i_0}. C'est en particulier le cas pour x : il existe un réel $r > 0$ tel que $B(x,r) \subset O_{i_0}$. Mais comme O_{i_0} est inclus dans O la boule $B(x,r)$ est incluse dans O. On a ainsi montré que pour tout élément x de O, il existe un réel $r > 0$ tel que $B(x,r) \subset O$. L'ensemble O est donc ouvert.

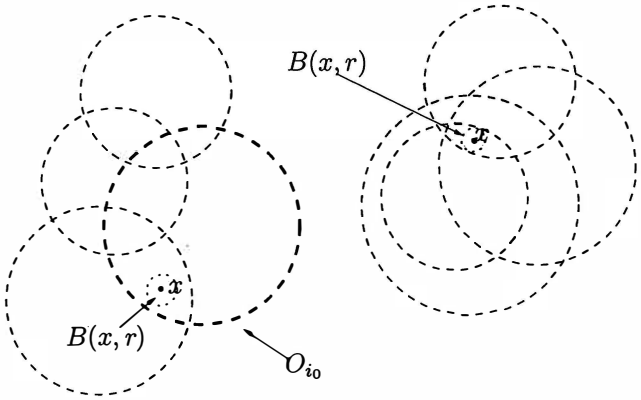


Fig. 6 Les deux situations considérées.

$\trianglerighteq$ Soit $(O_i)_{i\in\{1,\dots,n\}}$ une famille finie d'ensembles ouverts ; désignons par O l'intersection de ces ensembles. Si l'intersection est l'ensemble vide, il s'agit bien d'un ouvert. Supposons donc cette intersection non vide et considérons un élément x quelconque de O, voir le dessin de droite de la figure 6. Il appartient nécessairement à tous les ensembles O_i et comme ces ensembles sont ouverts, pour tout $i \in \{1,\dots,n\}$ il existe un réel $r_i > 0$ tel que $B(x,r_i) \subset O_i$. Si $r = \min_{i\in\{1,\dots,n\}} r_i$ alors $r > 0$ et pour tout $i \in \{1,\dots,n\}$ on a $B(x,r) \subset B(x,r_i) \subset O_i$. On a par conséquent $B(x,r) \subset O$ et O est donc un ensemble ouvert.

$\trianglerighteq$ Les propriétés concernant les ensembles fermés s'obtiennent à partir des propriétés montrées pour les ensembles ouverts en utilisant les lois de Morgan[(42)]. □

Remarque L'intersection d'un nombre infini d'ouverts n'est pas nécessairement un ensemble ouvert. Ce qui ne « marche pas » avec une intersection infinie dans la démonstration qui a été faite de la proposition 7.9 est que si l'ensemble I est infini, on doit remplacer $r = \min_{i\in\{1,\dots,n\}} r_i$ par $r = \inf_{i\in I} r_i$ et r peut alors valoir 0. Pour illustrer cette possibilité, on peut de considérer l'intersection sur $\mathbb{N}^*$ (qui est un ensemble infini) de toutes les boules ouvertes de centre 0 et de rayon $1/n$. On a $r = \inf_{n\in\mathbb{N}^*} \frac{1}{n} = 0$ et $\bigcap_{n\in\mathbb{N}^*} B(0,\frac{1}{n}) = \{0\}$ qui est un ensemble fermé. ◇

(42) Voir la proposition 2.3 p. 30 du *Cours de première année*.

7.3.2 Intérieur et adhérence d'un ensemble

DÉFINITION 7.8 (Intérieur)
Soient (E,d) un espace métrique, A un sous-ensemble non vide de E et x_0 un élément de E.

✗ *On dit que x_0 est un* point intérieur *à A si A est un voisinage de x_0. Autrement dit, si*

$$\exists r \in \mathbb{R}_+^* \quad B(x_0,r) \subset A.$$

✗ *L'ensemble des points intérieurs à A est noté* $\mathrm{Int}(A)$ *ou* $\mathring{A}$ *et est appelé* intérieur *de A.*

Remarques

1. L'intérieur d'un sous-ensemble A de E est un ensemble ouvert. Pour le montrer, il faut vérifier que pour tout $x_0 \in \mathrm{Int}\,(A)$, il existe $r \in \mathbb{R}_+^*$ tel que $B(x_0,r) \subset \mathrm{Int}\,(A)$. D'après la définition 7.8, pour tout $x_0 \in \mathrm{Int}\,(A)$, il existe $r \in \mathbb{R}_+^*$ tel que $B(x_0,r) \subset A$. Or $B(x_0,r)$ est un ouvert, donc pour tout $x \in B(x_0,r)$ il existe $\rho \in \mathbb{R}_+^*$ tel que $B(x,\rho) \subset B(x_0,r)$. Ceci implique en particulier que $B(x,\rho) \subset A$ et par conséquent que x est un point intérieur à A. On a donc bien $B(x_0,r) \subset \mathrm{Int}\,(A)$.

2. L'intérieur d'un sous-ensemble A de E est le plus grand ouvert contenu dans A. Considérons en effet un ouvert O contenu dans A ; pour tout $x \in O$, il existe $r \in \mathbb{R}_+^*$ tel que $B(x,r) \subset O$ et par conséquent tel que $B(x,r) \subset A$. Cela signifie que $x \in \mathrm{Int}\,(A)$ et donc que $O \subset \mathrm{Int}\,(A)$. Ainsi, $\mathrm{Int}\,(A)$ est un ouvert et tout ouvert inclus dans A est inclus dans $\mathrm{Int}\,(A)$; c'est donc que $\mathrm{Int}\,(A)$ est le plus grand ouvert contenu dans A. On en déduit que $\mathrm{Int}\,(A) \subset A$ et $\mathrm{Int}\,(\mathrm{Int}\,(A)) = \mathrm{Int}\,(A)$. De plus, A est ouvert si et seulement si $\mathrm{Int}\,(A) = A$.

3. L'intérieur d'un sous-ensemble A de E est la réunion de tous les ouverts contenus dans A. En effet, la réunion X de tous les ouverts contenus dans A contient l'ouvert $\mathrm{Int}\,(A)$ et est un ouvert inclus dans A ; on a donc $\mathrm{Int}\,(A) \subset X$. Par ailleurs, comme $\mathrm{Int}\,(A)$ est le plus grand ouvert contenu dans A on a également $X \subset \mathrm{Int}\,(A)$. Ainsi, $X = \mathrm{Int}\,(A)$. ◇

On a également les propriétés suivantes.

PROPOSITION 7.10 *Soient (E,d) un espace métrique et A et B deux sous-ensembles non vides de E.*

1. $A \subset B \Longrightarrow \mathrm{Int}\,(A) \subset \mathrm{Int}\,(B)$;

2. $\mathrm{Int}\,(A \cap B) = \mathrm{Int}\,(A) \cap \mathrm{Int}\,(B)$;

3. $\mathrm{Int}\,(A) \cup \mathrm{Int}\,(B) \subset \mathrm{Int}\,(A \cup B)$.

Démonstration $\trianglerighteq$ Si A est inclus dans B alors $\mathrm{Int}(A)$ qui est inclus dans A est inclus dans B. Or $\mathrm{Int}(A)$ est un ouvert et $\mathrm{Int}(B)$ est le plus grand ouvert inclus dans B. On a donc nécessairement $\mathrm{Int}(A) \subset \mathrm{Int}(B)$.

$\trianglerighteq$ D'une part, $\mathrm{Int}(A \cap B)$ est un ouvert qui est inclus dans $A \cap B$. Il est donc inclus dans A (resp. dans B). Comme $\mathrm{Int}(A)$ (resp. $\mathrm{Int}(B)$) est le plus grand ouvert inclus dans A (resp. B) on a $\mathrm{Int}(A \cap B) \subset \mathrm{Int}(A)$ (resp. $\mathrm{Int}(A \cap B) \subset \mathrm{Int}(B)$). On en déduit que $\mathrm{Int}(A \cap B) \subset \mathrm{Int}(A) \cap \mathrm{Int}(B)$. D'autre part, $\mathrm{Int}(A) \cap \mathrm{Int}(B)$ est un ouvert (c'est l'intersection de deux ouverts) qui est inclus dans $A \cap B$ car $\mathrm{Int}(A) \subset A$ et $\mathrm{Int}(B) \subset B$. Comme $\mathrm{Int}(A \cap B)$ est le plus grand ouvert inclus dans $A \cap B$ on en déduit que $\mathrm{Int}(A) \cap \mathrm{Int}(B) \subset \mathrm{Int}(A \cap B)$.

$\trianglerighteq$ $\mathrm{Int}(A) \cup \mathrm{Int}(B)$ est un ouvert (c'est la réunion de deux ouverts) qui est inclus dans $A \cup B$. Comme $\mathrm{Int}(A \cup B)$ est le plus grand ouvert inclus dans $A \cup B$, on a nécessairement $\mathrm{Int}(A) \cup \mathrm{Int}(B) \subset \mathrm{Int}(A \cup B)$. □

On n'a pas nécessairement $\mathrm{Int}(A) \cup \mathrm{Int}(B) = \mathrm{Int}(A \cup B)$. Prenons par exemple $A = [0,1[$ et $B = [1,2]$; on a $\mathrm{Int}(A) =]0,1[$ et $\mathrm{Int}(B) =]1,2[$ et par conséquent $\mathrm{Int}(A) \cup \mathrm{Int}(B) =]0,2[\backslash \{1\}$ alors que $A \cup B = [0,2]$ et $\mathrm{Int}(A \cup B) =]0,2[$.

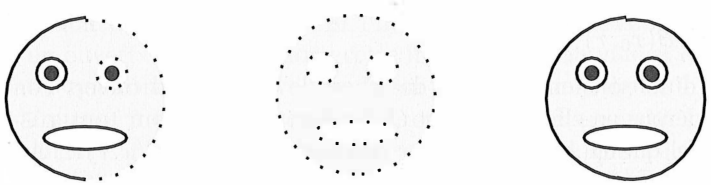


Fig. 7 Exemple d'un ensemble A de $\mathbb{R}^2$ (à gauche en grisé), de son intérieur (au centre) et de son adhérence (à droite). On matérialise par des traits pleins l'appartenance à l'ensemble et par des pointillés la non-appartenance.

DÉFINITION 7.9 (Adhérence)
Soient (E,d) un espace métrique, A un sous-ensemble non vide de E et x_0 un élément de E.

✗ *On dit que x_0 est un* point adhérent *à A si toute boule ouverte de centre x_0 contient au moins un élément de A, i.e. si*

$$\forall r \in \mathbb{R}_+^* \quad \exists y \in A \quad d(y,x_0) < r.$$

L'ensemble des points adhérents à A est noté $\overline{A}$ et est appelé adhérence *de A.*

✗ *On dit que x_0 est un* point d'accumulation *de A si toute boule ouverte de centre x_0 contient au moins un élément de A autre que x_0.*

Remarques

1. D'après la définition 7.6, x_0 est un point adhérent à A si et seulement si tout voisinage de x_0 contient au moins un élément de A et x_0 est un point d'accumulation de A si et seulement si tout voisinage de x_0 contient au moins un élément de A autre que x_0.

2. Un point d'accumulation de A est un point adhérent à A mais la réciproque est fausse. Un point adhérent à l'ensemble A qui n'est pas un point d'accumulation est appelé un *point isolé.*

3. L'adhérence d'un sous-ensemble non vide A de E est l'intersection de tous les fermés qui contiennent A (c'est donc un fermé). Il s'agit du plus petit ensemble fermé contenant A; on a donc $A \subset \overline{A}$ et $\overline{\overline{A}} = \overline{A}$. De plus, voir l'exercice 7.9,

$$\boxed{A \text{ est fermé si et seulement si } A = \overline{A}.}$$

4. On appelle frontière d'un sous-ensemble A de E et on note $\mathrm{Fr}(A)$ l'ensemble des éléments de E qui appartiennent à l'adhérence de A mais pas à l'intérieur de A, *i.e.* $\mathrm{Fr}(A) = \overline{A} \setminus \mathrm{Int}(A)$. On montre que $\mathrm{Fr}(A) = \overline{A} \cap \overline{\complement_E A}$ autrement dit que les éléments de la frontière de A sont les points qui sont adhérents à A et à son complémentaire. Comme il s'agit de l'intersection de deux fermés, la frontière de A est un ensemble fermé. ◇

Exercice 7.9 Soient (E,d) un espace métrique et A un sous-ensemble non vide de E.
1 - Montrer que $A \subset \overline{A}$ mais qu'en général il n'y a pas égalité de ces ensembles.
2 - a) Montrer que si A est fermé alors $A = \overline{A}$.
b) Montrer que si $A = \overline{A}$ alors A est fermé.

On a également les propriétés suivantes.

Proposition 7.11 *Soient (E,d) un espace métrique et A et B deux sous-ensembles non vides de E.*

1. $\overline{\complement_E A} = \complement_E \mathrm{Int}(A)$ *et* $\mathrm{Int}(\complement_E A) = \complement_E \overline{A}$;

2. $A \subset B \Longrightarrow \overline{A} \subset \overline{B}$;

3. $\overline{A \cup B} = \overline{A} \cup \overline{B}$;

4. $\overline{A \cap B} \subset \overline{A} \cap \overline{B}$.

Démonstration $\trianglerighteq$ Montrons que $\overline{\complement_E A} = \complement_E \mathrm{Int}(A)$. Un élément x appartient à $\complement_E \mathrm{Int}(A)$ s'il n'appartient pas à $\mathrm{Int}(A)$, on a donc[43] :

$$\begin{aligned} x \in \complement_E \mathrm{Int}(A) &\iff \forall r \in \mathbb{R}_+^* \quad B(x,r) \not\subset A \\ &\iff \forall r \in \mathbb{R}_+^* \quad \exists y \in B(x,r) \quad y \notin A. \end{aligned}$$

[43] Il s'agit de la négation de l'assertion $\exists r \in \mathbb{R}_+^*\ B(x,r) \subset A$ traduisant que $x \in \mathrm{Int}(A)$.

Par ailleurs, d'après la définition 7.9,

$$\begin{aligned} x \in \overline{\complement_E A} &\iff \forall r \in \mathbb{R}_+^* \quad \exists y \in \complement_E A \quad d(x,y) < r \\ &\iff \forall r \in \mathbb{R}_+^* \quad \exists y \in \complement_E A \quad y \in B(x,r). \end{aligned}$$

On en déduit que $\overline{\complement_E A} = \complement_E \mathrm{Int}\,(A)$.

$\trianglerighteq$ Montrons que $\mathrm{Int}\,(\complement_E A) = \complement_E \overline{A}$. On a d'une part, d'après la définition 7.8,

$$\begin{aligned} x \in \mathrm{Int}\,(\complement_E A) &\iff \exists r \in \mathbb{R}_+^* \quad B(x,r) \subset \complement_E A \\ &\iff \exists r \in \mathbb{R}_+^* \quad \forall y \in B(x,r) \quad y \in \complement_E A. \end{aligned}$$

D'autre part, d'après la définition 7.9, on a[44]

$$\begin{aligned} x \in \complement_E \overline{A} \iff x \notin \overline{A} &\iff \text{non } (\forall r \in \mathbb{R}_+^* \quad \exists y \in B(x,r) \quad y \in A) \\ &\iff (\exists r \in \mathbb{R}_+^* \quad \forall y \in B(x,r) \quad y \in \complement_E A). \end{aligned}$$

On en déduit que $x \in \mathrm{Int}\,(\complement_E A)$ si et seulement si $x \in \complement_E \overline{A}$, ce qui signifie que $\mathrm{Int}\,(\complement_E A) = \complement_E \overline{A}$.

$\trianglerighteq$ D'après la proposition 7.10, pour tous sous-ensembles C et D non vides de E, on a $\mathrm{Int}\,(C \cap D) = \mathrm{Int}\,(C) \cap \mathrm{Int}\,(D)$. En prenant $C = \complement_E A$, $D = \complement_E B$ et en utilisant la première des relations que l'on vient de montrer, on obtient :

$$\complement_E \mathrm{Int}\,(C \cap D) = \overline{\complement_E (C \cap D)} = \overline{\complement_E C \cup \complement_E D} = \overline{A \cup B}$$

et $$\complement_E \big(\mathrm{Int}\,(C) \cap \mathrm{Int}\,(D)\big) = \complement_E \mathrm{Int}\,(C) \cup \complement_E \mathrm{Int}\,(D) = \overline{\complement_E C} \cup \overline{\complement_E D} = \overline{A} \cup \overline{B}.$$

On a donc $\overline{A \cup B} = \overline{A} \cup \overline{B}$. Les deux autres relations, qui se démontrent sur le même principe, sont laissées en exercice. □

Dans $\mathbb{R}^n$ muni des distances d_1, d_2 ou d_∞, définies à la proposition 7.2, l'adhérence d'une boule ouverte $B(x_0, r)$ est la boule fermée $\overline{B}(x_0, r)$, mais cette propriété n'est pas toujours vraie dans un espace métrique. Par exemple, dans $\mathbb{R}^n$ muni de la distance discrète (voir les exemples de la p. 346), la boule ouverte $B(x_0, 1)$ est réduite au singleton $\{x_0\}$ (qui est un ouvert et un fermé pour cette distance) et par conséquent l'adhérence de la boule ouverte qui est $\overline{B(x_0, 1)} = \{x_0\}$ n'est pas égale à la boule fermée $\overline{B}(x_0, 1) = \mathbb{R}^n$. On montre toutefois que dans le cas où l'espace métrique est un espace vectoriel normé (la distance étant subordonnée à la norme) l'adhérence d'une boule ouverte est toujours une boule fermée.

DÉFINITION 7.10 (Densité d'un ensemble)
Soient (E, d) un espace métrique et A et B deux sous-ensembles non vides de E. On dit que A est dense dans B si $\overline{A} = B$, autrement dit si[45]

$$\forall x \in B \quad \forall r \in \mathbb{R}_+^* \quad \exists y \in A \quad d(x,y) < r.$$

[44] On pourra revoir les règles de négation d'une assertion quantifiée dans le chap. 1 du *Cours de première année*.

Remarque La propriété de densité est une relation transitive : si A est dense dans B et si B est dense dans C alors A est dense dans C. En effet, A est dense dans B si et seulement si $\overline{A} = B$ et B est dense dans C si et seulement si $\overline{B} = C$; on a donc $\overline{A} = \overline{\overline{A}} = \overline{B} = C$. ◇

Exemples

1. L'ensemble des rationnels $\mathbb{Q}$ est dense dans $\mathbb{R}$, voir la proposition 3.13 p. 108 du *Cours de première année*. Plus généralement, l'ensemble $\mathbb{Q}^n$ est dense dans $\mathbb{R}^n$ pour les distances définies p. 354.

2. Montrons que l'ensemble $\mathcal{GL}_n(\mathbb{K})$ des matrices inversibles de $\mathcal{M}_n(\mathbb{K})$ est dense dans $\mathcal{M}_n(\mathbb{K})$ muni de la distance subordonnée à l'une des normes N_1 ou N_∞ définies à la proposition 7.5 p. 365 et notée ici $\|\cdot\|$. Soient M une matrice de $\mathcal{M}_n(\mathbb{K})$ et r un réel strictement positif. Considérons l'application p_M qui au scalaire $\lambda \in \mathbb{K}$ associe le déterminant de $M - \lambda I$ (p_M est la fonction polynomiale associée au polynôme caractéristique[46] de M). Le polynôme caractéristique étant de degré n, il admet au plus[47] n racines dans $\mathbb{K}$. On peut donc trouver un scalaire $\lambda_0 \in \mathbb{K}$ vérifiant $0 < |\lambda_0| < r$ et $p_M(\lambda_0) \neq 0$. Posons $M_0 = M - \lambda_0 I$; Comme $p_M(\lambda_0) = \det(M_0) \neq 0$, la matrice M est inversible. De plus, on a

$$\|M - M_0\| = \|\lambda_0 I\| = |\lambda_0| \, \|I\| = |\lambda_0| < r.$$

On a donc montré que pour tout $M \in \mathcal{M}_n(\mathbb{K})$ et pour tout $r \in \mathbb{R}_+^*$, il existe une matrice $M_0 \in \mathcal{GL}_n(\mathbb{K})$ tel que $d(M_0, M) = \|M - M_0\| < r$, ce qui permet d'affirmer que $\mathcal{GL}_n(\mathbb{K})$ est dense dans $\mathcal{M}_n(\mathbb{K})$.

3. Montrons que l'ensemble $\mathcal{S}_{00}(\mathbb{K})$ des suites numériques nulles à partir d'un certain rang est dense dans l'ensemble $\mathcal{S}_0(\mathbb{K})$ des suites numériques qui convergent vers 0, muni de la distance $d_\infty : ((u_n)_n, (v_n)_n) \in \mathcal{S}(\mathbb{K})^2 \mapsto \sup_{n\in\mathbb{N}} |u_n - v_n|$. Soient $(u_n)_n$ une suite numérique qui converge vers 0 et $r \in \mathbb{R}_+^*$. Puisque[48] $\lim_{n\to+\infty} u_n = 0$,

$$\exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \qquad \big(n \geqslant N \quad \Longrightarrow \quad |u_n| \leqslant \tfrac{1}{2} r\big).$$

La suite $(v_n)_n$ définie par $v_n = u_n$ si $n < N$ et $v_n = 0$ sinon, est une suite de $\mathcal{S}_{00}(\mathbb{K})$ et on a

$$d_\infty((u_n)_n, (v_n)_n) = \sup_{n<N} |u_n - v_n| + \sup_{n\geqslant N} |u_n - v_n| = \sup_{n\geqslant N} |u_n| \leqslant \tfrac{1}{2} r < r.$$

D'après la définition 7.10, on en déduit que $\mathcal{S}_{00}(\mathbb{K})$ est dense dans $\mathcal{S}_0(\mathbb{K})$ pour la distance d_∞.

---

[45] Voir la définition 7.9 p. 378.

[46] Voir la définition 12.5 p. 525 du *Cours de première année*.

[47] Plus précisément, si $\mathbb{K} = \mathbb{C}$, il admet exactement n racines comptées avec multiplicité et il admet au plus n racines si $\mathbb{K} = \mathbb{R}$, voir le théorème de D'Alembert-Gauss p. 248 du *Cours de première année*.

[48] Voir la définition 5.1 p. 168 du *Cours de première année*.

4. L'ensemble des fonctions polynomiales à coefficients réels est dense dans $\mathcal{C}^0([a,b],\mathbb{R})$ pour la distance d_∞ définie p. 348. Ce résultat est connu sous le nom de *théorème de Weierstrass*[49].

Soient (E,d) un espace métrique et A et B deux sous-ensembles non vides de E ; d'après la proposition 7.11 on a l'inclusion $\overline{A\cap B}\subset\overline{A}\cap\overline{B}$ mais en général il n'y a pas égalité. En particulier, si A et B sont denses dans E alors $\overline{A}=\overline{B}=E$ et $\overline{A}\cap\overline{B}=E$ mais $A\cap B$ n'est pas nécessairement dense dans E. Par exemple, l'ensemble $\mathbb{Q}$ des nombres rationnels est dense dans $\mathbb{R}$ de même que l'ensemble $\mathbb{R}\setminus\mathbb{Q}$ des nombres irrationnels mais $\mathbb{Q}\cap(\mathbb{R}\setminus\mathbb{Q})$ qui est l'ensemble vide n'est pas dense dans $\mathbb{R}$.

Toutefois, on montre que si A et B sont denses dans E et si l'un des deux ensembles est ouvert alors $A\cap B$ est dense dans E. Cette propriété fait l'objet de l'exercice suivant.

Exercice 7.10 Soient (E,d) un espace métrique et A, B deux sous-ensembles de E. Montrer que si A et B sont denses dans E et si A est ouvert alors $A\cap B$ est dense dans E.

7.3.3 Ensembles bornés

Définition 7.11 *Soient (E,d) un espace métrique et A un sous-ensemble de E. On dit que A est borné s'il existe un élément x_0 dans E et un réel strictement positif r tel que A soit inclus dans la boule fermée $\overline{B}(x_0,r)$, i.e. si*[50]

$$\exists x_0\in E\quad \exists r\in\mathbb{R}_+^*\quad \forall x\in A\quad d(x,x_0)\leqslant r.$$

Remarques

1. L'inégalité triangulaire permet de vérifier que la propriété d'être bornée ne dépend pas du choix de x_0. Cela signifie en particulier que si A est borné alors il existe un réel strictement positif r tel que A soit inclus dans la boule de centre 0 et de rayon r.

2. D'après la définition 7.11, l'ensemble A n'est pas borné si

$$\forall x_0\in E\quad \forall r\in\mathbb{R}_+^*\quad \exists x\in A\quad d(x,x_0)>r.$$

3. Si A et B sont deux sous-ensembles de E avec $A\subset B$ et si B est borné alors A est borné. Par ailleurs, si A est borné alors $\overline{A}$ est borné.

---

[49] Weierstrass, Karl (1815, Ostenfelde - 1897, Berlin).

[50] On obtient une définition similaire en considérant une boule ouverte au lieu d'une boule fermée.

4. L'intersection d'un nombre quelconque (fini ou infini) d'ensembles bornés et borné. La réunion d'un nombre fini d'ensembles bornés est bornée. Attention, la réunion d'un nombre infini d'ensembles bornés n'est pas nécessairement bornée. Ainsi, dans $\mathbb{R}$ muni de la distance usuelle $\bigcup_{n\in\mathbb{N}^*}[-1/n, 1/n] = [-1, 1]$ est borné mais $\bigcup_{n\in\mathbb{N}^*}[-n, n] = \mathbb{R}$ n'est pas borné. ◇

Exemples

1. Il résulte de la définition 7.11 que dans un espace métrique (E, d), toute boule fermée $\overline{B}(x_0, r)$ est bornée. Il en est de même des boules ouvertes.

2. Dans $\mathbb{R}$ muni de la distance usuelle, $\mathbb{R}$ n'est pas borné car pour tout $x_0 \in \mathbb{R}$ et pour tout $r \in \mathbb{R}_+^*$, le réel $x = x_0 + 2r$ est tel que $d(x, x_0) = |x - x_0| = 2r > r$. Par contre, si on munit $\mathbb{R}$ de la distance $d : (x, y) \in \mathbb{R}\times\mathbb{R} \mapsto |\arctan(x) - \arctan(y)|$ alors $\mathbb{R}$ est borné car pour tout $x \in \mathbb{R}$ on a $d(x, 0) = |\arctan(x)| \leqslant \frac{\pi}{2}$. Nous avions établi p. 353 que ces deux distances sur $\mathbb{R}$ n'étaient pas équivalentes.

3. Considérons l'ensemble $\mathcal{B}(I, \mathbb{R})$ des applications bornées sur l'intervalle I et munissons-le de la distance $d_\infty : (f, g) \in \mathcal{B}(I, \mathbb{R})^2 \mapsto \sup_{x\in I} |f(x) - g(x)|$. L'ensemble des applications bornées sur I n'est pas borné pour d_∞ car pour tout $f_0 \in \mathcal{B}(I, \mathbb{R})$ et pour tout $r \in \mathbb{R}_+^*$ la fonction $f : x \in I \mapsto 2r + f_0(x)$ appartient à $\mathcal{B}(I, \mathbb{R})$ et vérifie

$$d_\infty(f_0, f) = \sup_{x\in I} |f_0(x) - f(x)| = 2r > r.$$

Il n'y a pas de contradiction ; en considérant des applications de I dans $\mathbb{R}$ « bornées », on considère des applications pour lesquelles l'ensemble $f(I)$ est borné dans $\mathbb{R}$ pour la distance usuelle, voir le *Cours de première année* p. 565.

Il résulte de la définition 7.11 que si A est un ensemble borné dans l'espace métrique (E, d) alors A est borné dans l'espace métrique (E, δ) pour toute distance δ équivalente à la distance d. Mais un ensemble A peut être borné dans l'espace métrique (E, d) et non borné dans l'espace métrique (E, δ) si d et δ sont deux distances qui ne sont pas équivalente, voir les exemples précédents.

Puisque si deux distances sont équivalentes alors tout ensemble borné pour l'une est également borné pour l'autre, on dispose d'un moyen d'établir que deux distances ne sont pas équivalentes : il suffit de trouver un ensemble qui soit borné pour l'une des distances et pas pour l'autre. Ainsi, nous avons établi au cours de l'exercice 7.1 que l'application

$$d : (x, y) \in \mathbb{R}^2 \longmapsto \ln\left(1 + \frac{|x - y|}{1 + |x - y|}\right)$$

définit une distance dans $\mathbb{R}$. Pour cette distance, $\mathbb{R}$ est borné car car pour tout $x \in \mathbb{R}$ on a $d(x, 0) = \ln\left(1 + \frac{|x|}{1+|x|}\right) \leqslant \ln(2)$. Cette distance n'est donc pas équivalente à la distance usuelle de $\mathbb{R}$.

7.3.4 Espaces connexes

On appelle partition[51] d'un ensemble E toute famille $\{A_i\}_{i\in I}$, indicée par un sous-ensemble I de $\mathbb{R}$, de sous-ensembles A_i de E qui :

- sont non vides : $\forall i \in I \quad A_i \neq \varnothing$;
- sont deux à deux disjointes : $\forall (i,j) \in I^2 \quad (i \neq j \Rightarrow A_i \cap A_j = \varnothing)$;
- dont la réunion est l'ensemble E : $\bigcup_{i\in I} A_i = E$.

DÉFINITION 7.12 (Connexité)

✗ *Un espace métrique (E,d) est dit* connexe *s'il n'existe pas de partition de E en deux ouverts non vides disjoints.*

✗ *Un sous-ensemble A de E est appelée* partie connexe *de (E,d) si A muni de la distance induite par d est un espace métrique connexe.*

Remarques

1. De manière équivalente, un espace métrique est *connexe* s'il n'existe pas de partition de E en deux fermés non vides disjoints. On peut également caractériser un espace métrique connexe de la manière suivante : un espace métrique (E,d) est connexe si les seuls ensembles à la fois ouverts et fermés dans (E,d) sont l'ensemble vide et E. Intuitivement, une partie est connexe si elle est formée « d'un seul morceau », voir la figure 8.

2. En topologie, le terme « domaine » désigne un sous-ensemble d'un espace métrique qui est à la fois ouvert et connexe. ◇

Exemples

1. L'ensemble $\mathbb{R}$ muni de la distance usuelle est connexe. Les parties connexes de $\mathbb{R}$ sont des intervalles. Pour le vérifier, raisonnons par l'absurde et supposons qu'il existe un sous-ensemble X dans $\mathbb{R}$ qui ne soit pas un intervalle et qui soit connexe. Si X n'est pas un intervalle[52], il existe $(a,b,c) \in \mathbb{R}^3$ tel que

$$a \in X, \quad b \in X, \quad a \leqslant c \leqslant b \quad \text{et} \quad c \notin X.$$

Considérons les ensembles $U =]-\infty, c[\cap X$ et $V =]c, +\infty[\cap X$; U et V sont deux ouverts non vides ($a \in U$ et $b \in V$), disjoints et $U \cup V = X$. Par conséquent $\{U, V\}$ forme une partition de X, ce qui contredit l'hypothèse que X est connexe.

---

[51] Partitionner un ensemble E, c'est le « découper » en sous-ensembles disjoints. Par exemple, si A est sous-ensemble non vide de E, $\{A, \complement_E A\}$ constitue une partition de E. Une partition de $\mathbb{R}$ est $\{[n, n+1[\ ; \ n \in \mathbb{Z}\}$.

[52] On rappelle, voir la définition 3.9 p. 109 du *Cours de première année*, qu'un ensemble I dans $\mathbb{R}$ est un intervalle si $\forall (a,b,c) \in \mathbb{R}^3 \quad (a \in I \text{ et } b \in I \text{ et } a \leqslant c \leqslant b \Rightarrow c \in I)$. Par conséquent, un sous-ensemble I de $\mathbb{R}$ n'est pas un intervalle si

$$\exists (a,b,c) \in \mathbb{R}^3 \quad (a \in I \text{ et } b \in I \text{ et } a \leqslant c \leqslant b \text{ et } c \notin I).$$

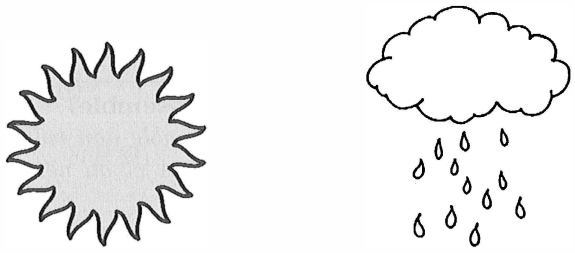


Fig. 8 Exemple d'une partie connexe de $\mathbb{R}^2$ muni de la distance euclidienne à gauche et d'une partie non connexe à droite.

2. Plus généralement, $\mathbb{R}^n$ muni de l'une des distances définies p. 354 est connexe. De même, $\mathbb{R}^n \setminus \{0\}$ est connexe pour la distance euclidienne.

Remarques

1. L'intersection de deux parties connexes n'est pas nécessairement connexe. On montre que l'union de deux parties connexes dont l'intersection est non vide est connexe. Si cette intersection est vide, l'union des deux parties connexes peut être ou non connexe ; considérer successivement $]-\infty, 0[\cup[0, +\infty[$ et $]-\infty, 0[\cup[1, +\infty[$.

2. L'adhérence d'une partie connexe est connexe, mais son intérieur ne l'est pas nécessairement.

3. Dans un espace vectoriel normé, les boules (ouvertes et fermées) sont connexes. Attention, cette propriété n'est pas toujours vraie dans un espace métrique. ◇

Il existe une notion plus restrictive[53] que la connexité : la connexité par arcs.

DÉFINITION 7.13 (Connexité par arcs)
Un espace métrique (E, d) *est dit* connexe par arcs *si pour tout* $(\alpha, \beta) \in E^2$ *il existe une application continue*[54] ϕ *de* $[a, b]$ *dans* E *telle que* $\phi(a) = \alpha$ *et* $\phi(b) = \beta$.

On appelle *chemin* dans un espace métrique (E, d) toute application continue ϕ d'un intervalle $[a, b]$ de $\mathbb{R}$ dans E. Un espace métrique (E, d) est donc connexe par arcs si deux quelconques de ses points peuvent être joints par un chemin.

---

[53] On montre qu'un espace métrique connexe par arcs est un espace connexe.

[54] *i.e.* vérifiant : $\forall x_0 \in [a, b]\ \forall \epsilon \in \mathbb{R}_+^*\ \exists \eta \in \mathbb{R}_+^*\ \forall x \in [a, b]\ \left(|x - x_0| \leqslant \eta \ \Rightarrow\ d(\phi(x), \phi(x_0)) \leqslant \epsilon\right)$.

7.3.5 Distance à un ensemble

DÉFINITION 7.14 (Distance d'un point à un ensemble)
Soient (E,d) un espace métrique et A un sous-ensemble non vide de E. On appelle distance de l'élément x_0 de E à l'ensemble A et on note $\delta(x,A)$ la borne inférieure de l'ensemble $\{d(x,y)\ ;\ y\in A\}$.

Remarques

1. Si A est non vide, la borne inférieure de l'ensemble $\{d(x,y)\ ;\ y\in A\}$ existe bien car[55] cet ensemble de réels est non vide et minoré par 0. De plus, cette borne inférieure $\delta(x,A)$ est caractérisée par

$$\big(\forall\varepsilon\in\mathbb{R}_+^*\quad \exists y_\varepsilon\in A\quad d(y_\varepsilon,x)<\delta(x,A)+\varepsilon\big)\quad\text{et}\quad\big(\forall y\in A\quad d(x,y)\geqslant\delta(x,A)\big).$$

2. Pour tout $(x,y)\in E$ on a $|\delta(x,A)-\delta(y,A)|\leqslant d(x,y)$. ◇

Nous allons voir avec les exemples qui suivent, que la définition 7.14 de la distance d'un élément à un ensemble recouvre bien la notion de distance d'un point à une droite ou à un plan définie en géométrie.

Exemples

1. Calculons dans $\mathbb{R}^2$ muni de la distance euclidienne, la distance du point P de coordonnées (x_0,y_0) à l'ensemble $\mathcal{D}=\{(x,y)\in\mathbb{R}^2\ ;\ ax+by+c=0\}$ où $(a,b,c)\in\mathbb{R}^3, b\neq 0$ (il s'agit d'une droite non parallèle à l'axe des ordonnées), voir la figure 9. On a

$$\begin{aligned}\delta(P,\mathcal{D}) &= \inf_{Q\in\mathcal{D}} d_2(P,Q)=\inf_{(x,y)\in\mathcal{D}}\sqrt{(x_0-x)^2+(y_0-y)^2}\\ &= \inf_{x\in\mathbb{R}}\sqrt{(x_0-x)^2+\left(y_0+\frac{ax+c}{b}\right)^2}.\end{aligned}$$

En étudiant l'application $f: x\in\mathbb{R}\mapsto (x_0-x)^2+(y_0+\frac{1}{b}(ax+c))^2$, on vérifie qu'elle atteint son maximum en $x^*=(x_0b^2-aby_0-ac)/(b^2+a^2)$. On en déduit que

$$\delta(P,\mathcal{D})=\sqrt{(x_0-x^*)^2+\left(y_0+\frac{ax^*+c}{b}\right)^2}=\frac{|ax_0+by_0+c|}{\sqrt{a^2+b^2}}.$$

2. On peut établir de la même manière que dans $\mathbb{R}^3$ muni de la distance euclidienne, la distance du point P de coordonnées (x_0,y_0,z_0) au plan $\mathcal{P}=\{(x,y,z)\in\mathbb{R}^3\ ;\ ax+by+cz+d=0\}$ (où $(a,b,c)\neq(0,0,0)$) est, voir la figure 9,

$$\delta(P,\mathcal{P})=\frac{|ax_0+by_0+cz_0+d|}{\sqrt{a^2+b^2+c^2}}.$$

[55] Voir la proposition 3.2 p. 95 du *Cours de première année*.

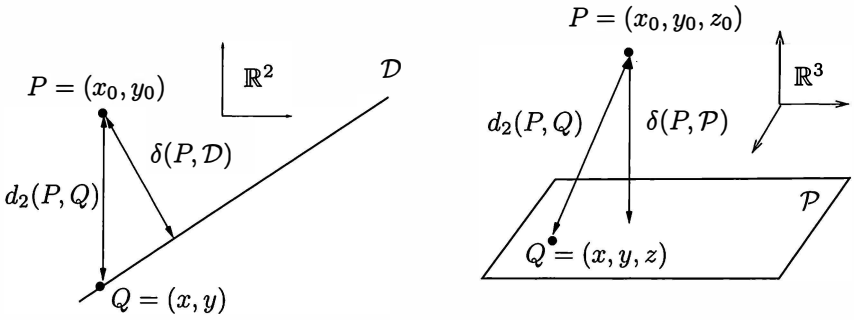


Fig. 9 Distance d'un point à une droite dans $\mathbb{R}^2$ muni de la distance euclidienne et distance d'un point à un plan dans $\mathbb{R}^3$.

PROPOSITION 7.12 *Soient (E, d) un espace métrique et A un sous-ensemble de E. On a*

$$\overline{A} = \{x \in E \ ; \ \delta(x, A) = 0\}$$

où $\delta(x, A)$ désigne la distance de l'élément x à l'ensemble A.

Démonstration $\trianglerighteq$ Soit x un élément de E tel que $\delta(x, A) = \inf_{y \in A} d(x, y) = 0$. Il résulte des propriétés de la borne inférieure[56] d'un ensemble de réels que :

$$\big(\forall \varepsilon \in \mathbb{R}_+^* \quad \exists y_\varepsilon \in A \quad d(x, y_\varepsilon) < \varepsilon\big) \qquad \text{et} \qquad \big(\forall y \in A \quad d(x, y) \geqslant 0\big).$$

On déduit de la première partie de cette assertion que toute boule ouverte de centre x et de rayon ε contient au moins un élément y_ε de A. Cela signifie d'après la définition 7.9 que $x \in \overline{A}$.

$\trianglerighteq$ Réciproquement, montrons que si $x \in \overline{A}$ alors $\delta(x, A) = 0$ ou, ce qui est équivalent[57], montrons que si $\delta(x, A) \neq 0$ alors $x \notin \overline{A}$. Supposons que $\delta(x, A) \neq 0$; il résulte des propriétés de la borne inférieure[56] d'un ensemble de réels que :

$$\big(\forall \varepsilon \in \mathbb{R}_+^* \quad \exists y_\varepsilon \in A \quad d(y_\varepsilon, x) < \delta(x, A) + \varepsilon\big) \quad \text{et} \quad \big(\forall y \in A \quad d(x, y) \geqslant \delta(x, A)\big).$$

Puisque pour tout $y \in A$ on a $d(x, y) \geqslant \delta(x, A)$, la boule ouverte de centre x et de rayon $\frac{1}{2}\delta(x, A)$ ne contient aucun élément de A. Cela signifie, voir la définition 7.9, que $x \notin \overline{A}$. □

[56] Voir la proposition 3.2 p. 95 du *Cours de première année*.

[57] Il s'agit en effet de la contraposée de l'implication que l'on cherche à démontrer, voir la proposition 1.2 p. 10 du *Cours de première année*.

Remarque La proposition 7.12 permet d'établir que dans un espace métrique (E,d), les singletons $\{x_0\}$ avec $x_0 \in E$ sont des ensembles fermés. En effet, on a

$$\overline{\{x_0\}} = \{x \in E \ ; \ d(\{x_0\},x) = 0\} = \{x \in E \ ; \ d(x_0,x) = 0\} = \{x_0\},$$

car d étant une distance $d(x_0,x) = 0 \iff x = x_0$. Par ailleurs, un sous-ensemble fini de E pouvant s'écrire comme la réunion finie des singletons formés par les éléments qui le composent, c'est un ensemble fermé dans E, voir la proposition 7.9. ◇

DÉFINITION 7.15 (Distance entre deux ensembles)
Soient (E,d) un espace métrique et A et B deux sous-ensembles non vides de E. On appelle distance de l'ensemble A à l'ensemble B et on note $\delta(A,B)$ la borne inférieure de l'ensemble $\{d(x,y) \ ; \ (x,y) \in A \times B\}$.

7.4 Suites dans un espace métrique

7.4.1 Définitions et propriétés

Une suite dans un ensemble E est une application définie sur[58] $\mathbb{N}$ et qui est à valeurs dans E :

$$\begin{array}{ccc} \mathbb{N} & \longrightarrow & E \\ n & \longmapsto & x_n \end{array}$$

que l'on note $(x_n)_n$ où x_n est appelé le terme général de la suite.

L'étude des propriétés liées à la convergence d'une suite nécessite de munir l'ensemble E d'une distance d, autrement dit de considérer un espace métrique.

DÉFINITION 7.16 *Soient (E,d) un espace métrique et $(x_n)_n$ une suite d'éléments de E. On dit que la suite $(x_n)_n$ converge s'il existe un élément ℓ dans E tel que :*

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists \eta_\varepsilon \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad (n \geqslant \eta_\varepsilon \implies d(x_n,\ell) \leqslant \varepsilon).$$

L'élément ℓ de E est appelé limite *de la suite $(x_n)_n$ et on dit que la suite $(x_n)_n$ converge vers ℓ.*

[58] Plus généralement, on peut considérer un sous-ensemble N_1 non borné de $\mathbb{N}$, par exemple $\mathbb{N}^*$.

Remarques

1. L'assertion quantifiée de la définition 7.16 indique que la suite $(x_n)_n$ converge vers ℓ si et seulement si la suite réelle de terme général $d(x_n, \ell)$ converge vers 0. On déduit donc du théorème d'encadrement[59] pour les suites réelles ; le résultat suivant : s'il existe une suite réelle $(\alpha_n)_n$ à termes positifs convergeant vers 0 telle que pour tout $n \in \mathbb{N}$ on ait $0 \leqslant d(x_n, \ell) \leqslant \alpha_n$ alors la suite $(x_n)_n$ converge vers ℓ.

2. On dit que la suite $(x_n)_n$ de l'espace métrique (E, d) *diverge* s'il n'existe pas d'élément ℓ dans E tel que la suite $(x_n)_n$ converge vers ℓ, c'est-à-dire si :

$$\forall \ell \in E \quad \exists \varepsilon \in \mathbb{R}^+ \quad \forall \eta_\varepsilon \in \mathbb{N} \quad \exists n \in \mathbb{N} \quad (n \leqslant \eta_\varepsilon \text{ et } d(x_n, \ell) > \varepsilon).$$

3. La notion de convergence pour les suites numériques[60] réelles (resp. complexes) étudiée dans le *Cours de première année* correspond à la définition 7.16 lorsque E est l'ensemble des réels muni de la distance subordonnée à la valeur absolue (resp. lorsque E est l'ensemble des complexes muni de la distance subordonnée au module).

4. La suite $(f_n)_n$ converge vers f dans $\mathcal{C}^0([a,b], \mathbb{R})$ muni de la distance subordonnée à la norme uniforme $N_\infty : f \in \mathcal{C}^0([0,1], \mathbb{R}) \mapsto \sup_{x \in [0,1]} |f(x)|$ si

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists \eta_\varepsilon \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad (n \geqslant \eta_\varepsilon \implies \sup_{x \in [a,b]} |f_n(x) - f(x)| \leqslant \varepsilon).$$

On retrouve la notion de convergence uniforme d'une suite de fonctions étudiée au chap. 2, voir la définition 2.2 p. 73. C'est la raison pour laquelle la norme N_∞ est appelée *norme de la convergence uniforme.* ◇

Exemples

1. Dans $\mathbb{R}$ muni de la distance $d : (x,y) \in \mathbb{R}^2 \mapsto |\arctan(x) - \arctan(y)|$, la suite de terme général $x_n = n$ est une suite qui diverge. En effet, quel que soit le réel ℓ, la suite de terme général $d(x_n, \ell) = |\arctan(n) - \arctan(\ell)|$ converge vers $\frac{\pi}{2} - \arctan(\ell)$ et comme la fonction arc-tangente est à valeurs dans $]-\frac{\pi}{2}, \frac{\pi}{2}[$, on a $\frac{\pi}{2} - \arctan(\ell) \neq 0$. Cela indique qu'il n'existe pas de réel ℓ tel que la suite de terme général $d(x_n, \ell)$ puisse converger vers 0.

2. Munissons l'espace vectoriel $\mathbb{K}[X]$ de la distance d subordonnée à la norme[61] $N : P \in \mathbb{K}[X] \mapsto \sup_{x \in [-1,1]} |P(x)|$. La suite $(P_n)_n$ de polynômes de $\mathbb{K}[X]$ de terme général $P_n = \frac{1}{n} X^n$ converge vers le polynôme nul $0_{\mathbb{K}[X]}$ dans $(\mathbb{K}[X], d)$

---

[59] Voir le théorème 5.1 p. 179 du *Cours de première année.*

[60] Voir la définition 5.1 p. 168 du *Cours de première année.*

[61] On pourra vérifier que l'application N définit bien une norme. Les conditions 1, 3 et 4 de la définition 7.3 résultent directement des propriétés de la valeur absolue et de la borne supérieure. Pour établir la deuxième condition, on remarquera que si $N(P) = 0$ alors $P(x) = 0$ pour tout $x \in [-1,1]$ et par conséquent le polynôme P admet une infinité de racines ; c'est donc le polynôme nul, voir le *Cours de première année* p. 248.

car pour tout $n \in \mathbb{N}^*$

$$d(P_n, 0_{\mathbb{K}[X]}) = \sup_{x\in[-1,1]} |P_n(x)| = \frac{1}{n} \quad \text{d'où} \quad \lim_{n\to 0} d(P_n, 0_{\mathbb{K}[X]}) = 0$$

3. La suite de fonctions $(f_n)_n$ de terme général $f_n : x \in [0,1] \mapsto x^n$ converge simplement vers la fonction qui prend la valeur 0 sur $[0,1[$ et la valeur 1 en 1. Dans $\mathcal{C}^0([0,1],\mathbb{R})$ muni de la distance subordonnée à la norme de la convergence uniforme $N_\infty : f \in \mathcal{C}^0([0,1],\mathbb{R}) \mapsto \sup_{x\in[0,1]} |f(x)|$ la suite $(f_n)_n$ diverge. En effet, cette suite de fonctions continues convergeant simplement vers une fonction qui n'est pas continue, d'après le théorème 2.3 p. 86 il ne peut y avoir convergence uniforme. Dans $\mathcal{C}^0([0,1],\mathbb{R})$ muni de la distance d_1 subordonnée à la norme de la convergence en moyenne $N_1 : f \in \mathcal{C}^0([0,1],\mathbb{R}) \mapsto \int_0^1 |f(x)|\,\mathrm{d}x$ la suite $(f_n)_n$ converge vers la fonction nulle car pour tout $n \in \mathbb{N}$

$$d_1(f_n, 0) = \int_0^1 |f_n(x)|\,\mathrm{d}x = \frac{1}{n+1} \quad \text{d'où} \quad \lim_{n\to 0} d_1(f_n, 0) = 0.$$

PROPOSITION 7.13 *Si une suite de l'espace métrique (E,d) converge alors la limite de cette suite est unique.*

Démonstration Raisonnons par l'absurde et supposons que la suite $(x_n)_n$ converge dans (E,d) et admet deux limites ℓ_1 et ℓ_2 distinctes. On a $d(\ell_1,\ell_2) \neq 0$ et le réel $\varepsilon = \frac{1}{3}d(\ell_2,\ell_1)$ est strictement positif. On déduit de la définition 7.16 que

$$\begin{aligned} & \exists \eta_1 \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad (n \geqslant \eta_1 \Longrightarrow d(x_n,\ell_1) \leqslant \varepsilon) \\ \text{et} \quad & \exists \eta_2 \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad (n \geqslant \eta_2 \Longrightarrow d(x_n,\ell_2) \leqslant \varepsilon). \end{aligned}$$

Notons $\eta = \max\{\eta_1, \eta_2\}$. En utilisant la quatrième propriété de la définition 7.1, on obtient

$$d(\ell_2,\ell_1) \leqslant d(\ell_2,x_\eta) + d(x_\eta,\ell_1) \leqslant 2\varepsilon = \tfrac{2}{3}d(\ell_2,\ell_1)$$

ce qui est absurde puisque $1 > 2/3$. On en conclut que si la suite $(x_n)_n$ converge dans (E,d), sa limite est unique. □

LEMME 7.2 *Si δ est une distance sur E équivalente à une distance d alors toute suite convergeant dans (E,d) converge aussi dans (E,δ) et admet la même limite.*

Démonstration Pour établir ce résultat, considérons une suite $(x_n)_n$ convergeant dans (E,d) vers ℓ. Si la distance δ est équivalente à la distance d alors il existe $(C_1,C_2) \in (\mathbb{R}_+^*)^2$ tel que pour tout $n \in \mathbb{N}$ on ait :

$$C_1\, d(x_n,\ell) \leqslant \delta(x_n,\ell) \leqslant C_2\, d(x_n,\ell).$$

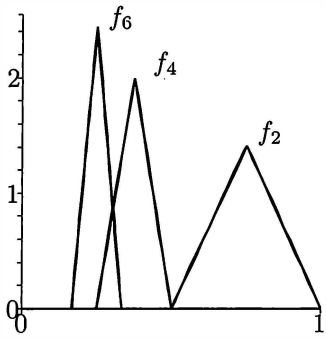


Fig. 10 Représentation graphique des premiers termes de la suite $(f_n)_n$.

La suite $(x_n)_n$ convergeant dans (E,d) vers ℓ on a $\lim_{n\to+\infty} d(x_n,\ell)=0$ d'où par le théorème d'encadrement : $\lim_{n\to+\infty}\delta(x_n,\ell)=0$. La suite $(x_n)_n$ converge donc vers ℓ dans (E,δ). □

Si d et δ sont deux distances sur E qui ne sont pas équivalentes alors une suite $(x_n)_n$ peut être de nature différente dans (E,d) et dans (E,δ). L'exemple suivant illustre cette mise en garde.

Exemple Considérons dans $\mathcal{C}^0([0,1],\mathbb{R})$ la suite $(f_n)_n$ de terme général :

$$f_n : x\in[0,1]\longmapsto\begin{cases}0 & \text{si } x\in[0,\frac{1}{n}]\cup[\frac{2}{n},1]\\ 2\sqrt{n}\,(nx-1) & \text{si } x\in]\frac{1}{n},\frac{3}{2n}[\\ -2\sqrt{n}\,(nx-2) & \text{si } x\in]\frac{3}{2n},\frac{2}{n}[\end{cases}$$

Cette suite de fonction converge simplement vers la fonction nulle, voir la figure 10. Munissons $\mathcal{C}^0([0,1],\mathbb{R})$ des distances subordonnées aux 3 normes définies à la proposition 7.6. Pour tout $n\in\mathbb{N}$ avec $n\geqslant 2$, on a

- $d_1(f_n,0)=\|f_n\|_1=\frac{1}{2\sqrt{n}}$ ce qui implique que dans $(\mathcal{C}^0([0,1],\mathbb{R}),d_1)$ la suite $(f_n)_n$ converge vers la fonction nulle ;
- $d_2(f_n,0)=\|f_n\|_2=\frac{1}{\sqrt{3}}$ ce qui implique que dans $(\mathcal{C}^0([0,1],\mathbb{R}),d_2)$ la suite $(f_n)_n$ ne converge pas vers la fonction nulle ;
- $d_\infty(f_n,0)=\|f_n\|_\infty=\sqrt{n}$ ce qui implique que dans $(\mathcal{C}^0([0,1],\mathbb{R}),d_\infty)$ la suite $(f_n)_n$ ne converge pas non plus vers la fonction nulle.

Exercice 7.11

1 - a) Soit $(f_n)_n$ une suite dans $\mathcal{C}^0([a,b],\mathbb{R})$. Montrer que si la suite $(f_n)_n$ converge uniformément vers f alors elle converge vers f dans $\mathcal{C}^0([a,b],\mathbb{R})$ muni de la distance d_2 subordonnée à la norme de la convergence quadratique[62]. La réciproque est-elle vraie ?

b) Nous avons établi, voir le théorème 2.1 p. 78, qu'une suite de fonctions convergeant uniformément sur un intervalle $[a,b]$ convergeait simplement sur cet intervalle. Montrer, en trouvant un contre-exemple, que la convergence d'une suite de fonctions dans $(\mathcal{C}^0([a,b],\mathbb{R}),d_2)$ n'implique pas nécessairement la convergence simple de la suite sur $[a,b]$.

2 - Montrer que si une suite $(f_n)_n$ converge dans $(\mathcal{C}^0([a,b],\mathbb{R}),d_2)$ alors elle converge aussi dans $(\mathcal{C}^0([a,b],\mathbb{R}),d_1)$ où d_1 est la distance subordonnée à la norme de la convergence en moyenne[62]. La réciproque est-elle vraie ?

Remarques

1. Par le même procédé que celui qui nous a permis d'introduire la notion de série numérique à partir de celle de suite numérique, on peut définir des séries dans un espace métrique.

2. Soient $(E_1,d_1),\dots,(E_p,d_p)$ des espaces métriques et (E,d) l'espace produit, voir p. 356. La suite $(\mathbf{x}_n)_n$ de terme général $\mathbf{x}_n=(x_n^{(1)},\dots,x_n^{(p)})$ converge vers $\ell=(\ell^{(1)},\dots,\ell^{(p)})$ dans (E,d) si et seulement si pour tout $k\in\{1,\dots,p\}$ la suite $(x_n^{(k)})_n$ converge dans (E_k,d_k) vers ℓ_k. ◇

Nous allons voir dans ce qui suit qu'un certain nombre de notions vues dans la cadre des suites numériques[63] se généralisent aux suites dans un espace métrique (E,d). On sera vigilant au fait que certaines des propriétés des suites numériques ne se généralisent pas aux suites dans un espace métrique. C'est en particulier le cas de la notion de suite monotone qui ne peut être définie que si E peut être muni d'une relation d'ordre total[64].

Définition 7.17 *Soient (E,d) un espace métrique et $(x_n)_n$ une suite d'éléments de E. La suite $(x_n)_n$ est dite* bornée *dans (E,d) si*

$$\exists e\in E\quad \exists K\in\mathbb{R}_+^*\quad \forall n\in\mathbb{N}\quad d(x_n,e)\leqslant K.$$

[62] Voir la proposition 7.6 p. 368.

[63] Voir le *Cours de première année*, chap. 5.

[64] Voir la définition 3.2 p. 89.

Remarques

1. Il résulte de la définition 7.17, qu'une suite $(x_n)_n$ d'éléments de E n'est pas bornée dans (E,d) si

$$\forall e \in E \quad \forall K \in \mathbb{R}_+^* \quad \exists n \in \mathbb{N} \quad d(x_n, e) > K. \tag{1}$$

2. Dire qu'une suite $(x_n)_n$ d'éléments de l'espace métrique (E,d) est bornée revient à dire que l'ensemble $U = \{x_n \ ; \ n \in \mathbb{N}\}$ des valeurs prises par la suite est un ensemble borné dans E, voir la définition 7.11. ◇

Exemples

1. Dans $\mathbb{R}$ muni de la distance usuelle, la suite $(x_n)_n$ de terme général $x_n = n$ n'est pas bornée car pour tout $e \in \mathbb{R}$ et pour tout $K \in \mathbb{R}_+^*$ le terme d'indice $n = \mathrm{E}(e+K)+1$ est tel que $|x_n - e| > K$. Dans $\mathbb{R}$ muni de la distance $d : (x,y) \in \mathbb{R}^2 \mapsto \frac{|x-y|}{1+|x-y|}$, voir l'exercice 7.2 p. 347, cette suite $(x_n)_n$ est bornée car pour tout $n \in \mathbb{N}$, on a $d(x_n, 0) = \frac{n}{1+n} \leqslant 1$.

2. Considérons l'espace vectoriel $\mathbb{K}[X]$ muni de la distance d subordonnée à la norme $P \in \mathbb{K}[X] \mapsto \sup_{x\in[-1,1]} |P(x)|$, voir p. 389. La suite $(P_n)_n$ de terme général $P_n = \frac{1}{n}X^n$ est bornée dans $(\mathbb{K}[X], d)$ car pour tout $n \in \mathbb{N}^*$

$$d(P_n, 0_{\mathbb{K}[X]}) = \sup_{x\in[-1,1]} |P_n(x)| = \frac{1}{n} \leqslant 1.$$

Munissons à présent l'espace vectoriel $\mathbb{K}[X]$ de la distance δ subordonnée à la norme[(65)] $N : P \in \mathbb{K}[X] \mapsto \sum_{k=0}^{+\infty} |P^{(k)}(0)|$ et considérons à nouveau la suite $(P_n)_n$ de terme général $P_n = \frac{1}{n}X^n$. Cette suite n'est pas bornée dans $(\mathbb{K}[X], \delta)$. En effet, soit $P = \sum_{k=0}^{m} a_k X^k$ un polynôme de $\mathbb{K}[X]$ et $K \in \mathbb{R}_+^*$; pour $n > m$ et $n > 2$, on a

$$\begin{aligned}\delta(P_n, P) &= \sum_{k=0}^{n} |P^{(k)}(0) - P_n^{(k)}(0)| = |P^{(m)}(0)| + |P_n^{(n)}(0)| \\ &= m!\,|a_m| + (n-1)! > (n-1)! > n.\end{aligned}$$

On en déduit qu'en choisissant $n = 1 + \max(2, m, \mathrm{E}(K))$, on a $\delta(P_n, P) > K$ ce qui prouve l'assertion (1).

Remarque Il résulte des définitions que si la suite $(x_n)_n$ est bornée dans (E,d) et que δ est une distance équivalente à d, alors la suite $(x_n)_n$ est bornée dans

(65) On vérifiera à titre d'exercice qu'il s'agit bien d'une norme. On pourra remarquer que la somme est finie car un polynôme de degré n à toutes ses dérivées d'ordre strictement supérieur à n qui sont nulles. Les conditions 1, 3 et 4 de la définition 7.3 résultent directement des propriétés de la valeur absolue et de la borne supérieure. Pour établir la deuxième condition, on remarquera que si P est un polynôme de degré n tel que $N(P) = 0$ alors pour tout $k \in \{0, \ldots, n\}$ on a $P^{(k)}(0) = 0$ ce qui signifie, voir la proposition 6.8 p. 244 du *Cours de première année*, que 0 est racine de multiplicité $n+1$ de P. Comme un polynôme de degré n qui est non nul admet au plus n racines, on en conclut que P est le polynôme nul.

(E,δ). Par contraposition, on dispose d'un moyen d'établir que deux distances ne sont pas équivalentes : il suffit de trouver une suite $(x_n)_n$ qui soit bornée pour l'une des distances et pas pour l'autre. Ainsi, les exemples précédents indiquent que sur $\mathbb{R}$ la distance $d : (x,y) \in \mathbb{R}^2 \mapsto \frac{|x-y|}{1+|x-y|}$ n'est pas équivalente à la distance usuelle et que sur $\mathbb{K}[X]$ les deux normes

$$P \in \mathbb{K}[X] \longmapsto \sup_{x\in[-1,1]} |P(x)| \quad \text{et} \quad P \in \mathbb{K}[X] \longmapsto \sum_{k=0}^{+\infty} |P^{(k)}(0)|$$

ne sont pas non plus équivalentes. ◇

Proposition 7.14 *Soient (E,d) un espace métrique et $(x_n)_n$ une suite d'éléments de E. Si la suite $(x_n)_n$ converge alors la suite $(x_n)_n$ est bornée.*

Démonstration Notons ℓ la limite dans (E,d) de la suite $(x_n)_n$; d'après la définition 7.16, on a :

$$\exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad \big(n \geqslant N \implies d(x_n,\ell) \leqslant 1\big).$$

Posons $K = \max(d(x_1,\ell),\dots,d(x_N,\ell),1)$; pour tout $n \in \mathbb{N}$, on a $d(x_n,\ell) \leqslant K$ ce qui d'après la définition 7.17 permet d'affirmer que la suite $(x_n)_n$ est bornée. □

Remarque Par contraposition, la proposition 7.14 indique que si une suite n'est pas bornée dans l'espace métrique (E,d) alors elle ne converge pas dans (E,d). Ainsi, on déduit des exemples de la p. 393 que la suite $(P_n)_n$ de terme général $P_n = \frac{1}{n}X^n$ ne converge pas dans l'espace vectoriel $\mathbb{K}[X]$ muni de la distance δ subordonnée à la norme $N : P \in \mathbb{K}[X] \mapsto \sum_{k=0}^{+\infty} |P^{(k)}(0)|$. En effet, on a établi que cette suite n'était pas bornée. ◇

Définition 7.18 (Valeur d'adhérence)
Soient (E,d) un espace métrique et $(x_n)_n$ une suite d'éléments de E. On dit que l'élément a de E est une valeur d'adhérence *de la suite $(x_n)_n$ si*

$$\forall r \in \mathbb{R}_+^* \quad \forall N \in \mathbb{N} \quad \exists n \in \mathbb{N} \quad \big(n \geqslant N \quad \text{et} \quad x_n \in B(a,r)\big).$$

Remarques

1. On peut montrer que l'ensemble des valeurs d'adhérence d'une suite est toujours un ensemble fermé.

2. D'après la définition 7.18, a est une valeur d'adhérence de la suite $(x_n)_n$ dans (E,d) si pour toute boule ouverte $B(a,r)$ centrée en a et pour tout élément x_N de la suite $(x_n)_n$, on peut trouver un élément de la suite $(x_n)_n$ de rang supérieur à N qui appartienne à la boule ouverte $B(a,r)$, voir la figure 11. ◇

Exemple Dans $\mathbb{R}^2$ muni de la distance euclidienne, la suite $(x_n)_n$ de terme général

$$x_n = \left((-1)^n \cos(\pi n^{-\frac{1}{2}}),\ (-1)^n n^{-\frac{1}{2}} \sin(\pi n^{\frac{1}{2}})\right)$$

possède deux valeurs d'adhérence qui sont $(-1,0)$ et $(1,0)$.

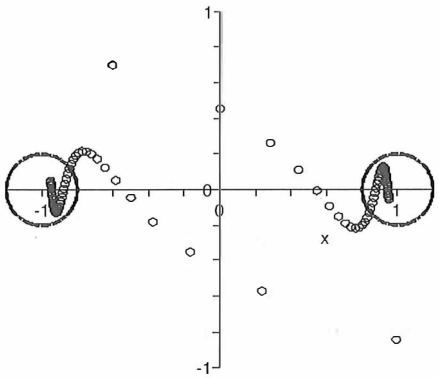


Fig. 11 La suite $(x_n)_n$ possède deux valeurs d'adhérence.

Montrons que $a = (1,0)$ est valeur d'adhérence de la suite. Pour tout $k \in \mathbb{N}^*$, on a

$$d_2(x_{2k}, a)^2 = \left(\cos(\pi/\sqrt{2k}) - 1\right)^2 + \frac{\sin^2(\pi\sqrt{2k})}{2k}.$$

Puisque $\lim_{k\to+\infty} d_2(x_{2k}, a) = 0$, pour tout $r > 0$ on a[(66)]

$$\exists K \in \mathbb{N} \quad \forall k \in \mathbb{N} \quad \left(k \geqslant K \quad \Longrightarrow \quad d_2(x_{2k}, a) \leqslant \tfrac{1}{2}r\right).$$

On en déduit que pour tout $N \in \mathbb{N}$, l'entier $n = \max(2K, N)$ vérifie $n \geqslant N$ et $d_2(x_n, a) \leqslant \frac{1}{2}r < r$, autrement dit $x_n \in B(a,r)$. On déduit de la définition 7.18 que $a = (1,0)$ est valeur d'adhérence de la suite $(x_n)_n$. On procède de manière similaire pour montrer que $(-1,0)$ est aussi valeur d'adhérence de la suite $(x_n)_n$.

Nous allons voir dans ce qui suit que la notion de valeur d'adhérence d'une suite est très liée à la notion de sous-suite et à la notion de limite.

7.4.2 Suites extraites

DÉFINITION 7.19 *Soient E un ensemble et $(x_n)_n$ une suite d'éléments de E. On dit que la suite $(y_n)_n$ est une* suite extraite *ou une* sous-suite *de la suite $(x_n)_n$ s'il existe une application h de $\mathbb{N}$ dans $\mathbb{N}$, strictement croissante, telle que pour tout $n \in \mathbb{N}$ on ait $y_n = x_{h(n)}$.*

(66) D'après la définition de la convergence d'une suite, voir la définition 5.1 p. 168 du *Cours de première année*.

Remarques

1. L'application h de la définition 7.19 est appelée *extractrice*. On vérifie par récurrence que si h est une extractrice alors pour tout $n \in \mathbb{N}$ on a $h(n) \geqslant n$.

2. Toute sous-suite d'une sous-suite de la suite $(x_n)_n$ est une sous-suite de la suite $(x_n)_n$ car la composée de deux applications strictement croissantes est une application strictement croissante (et par conséquent la composée de deux extractrices est une extractrice). ◇

PROPOSITION 7.15 *Soient (E, d) un espace métrique et $(x_n)_n$ une suite d'éléments de E. Si la suite $(x_n)_n$ converge vers ℓ dans (E, d) alors toute sous-suite de la suite $(x_n)_n$ converge également vers ℓ.*

Démonstration Supposons que la suite $(x_n)_n$ converge vers ℓ dans (E, d) ; d'après la définition 7.16, pour tout réel ε strictement positif fixé :

$$\exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad (n \geqslant N \Longrightarrow d(x_n, \ell) \leqslant \varepsilon).$$

Considérons une extractrice h et montrons que la suite $(x_{h(n)})_{n\in\mathbb{N}}$ converge vers ℓ. Comme h est une extractrice, h est strictement croissante et $h(N) \geqslant N$. Donc, si $n \geqslant N$ on a $h(n) \geqslant h(N) \geqslant N$ et par conséquent $d(x_{h(n)}, \ell) \leqslant \varepsilon$. Ainsi, pour tout réel strictement positif ε, il existe un entier N tel que si $n \geqslant N$ alors $d(x_{h(n)}, \ell) \leqslant \varepsilon$. D'après la définition 7.16, la suite $(x_{h(n)})_{n\in\mathbb{N}}$ converge vers ℓ dans (E, d). □

PROPOSITION 7.16 *Soient (E, d) un espace métrique et $(x_n)_n$ une suite d'éléments de E. Une condition nécessaire et suffisante pour que la suite $(x_n)_n$ converge dans (E, d) est que la sous-suite des termes d'indice pair* et *la sous-suite des termes d'indice impair convergent dans (E, d) et aient* même *limite. Dans ce cas, cette limite commune est la limite de la suite $(x_n)_n$.*

Démonstration $\trianglerighteq$ D'après la proposition 7.15 si la suite $(x_n)_n$ converge vers ℓ dans (E, d), la sous-suite des termes d'indice pairs et la sous-suite des termes d'indice impairs convergent vers ℓ.

$\trianglerighteq$ Réciproquement, supposons que la sous-suite des termes d'indice pairs et la sous-suite des termes d'indice impairs convergent toutes les deux vers ℓ dans (E, d). Pour tout réel ε strictement positif on a

$$\exists N_1 \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad (n \geqslant N_1 \Longrightarrow d(x_{2n}, \ell) \leqslant \varepsilon),$$
$$\exists N_2 \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad (n \geqslant N_2 \Longrightarrow d(x_{2n+1}, \ell) \leqslant \varepsilon).$$

Soient $N = \max(2N_1, 2N_2 + 1)$ et $p \in \mathbb{N}$ tel que $p \geqslant N$.

1. Si p est pair, il existe un entier k tel que $p = 2k$ et $k \geqslant N_1$.
2. Si p est impair, il existe un entier k tel que $p = 2k + 1$ et $k \geqslant N_2$.

Dans les deux cas, on a $d(x_p, \ell) \leqslant \varepsilon$. Ainsi, pour tout réel ε strictement positif, il existe un entier N tel que $d(x_p, \ell) \leqslant \varepsilon$ pour tout entier p supérieur à N. D'après la définition 7.16, la suite $(x_n)_n$ converge vers ℓ dans (E, d). □

THÉORÈME 7.3 *Soient (E, d) un espace métrique et $(x_n)_n$ une suite d'éléments de E. L'élément a de E est valeur d'adhérence de la suite $(x_n)_n$ si et seulement si il est limite d'une suite extraite de la suite $(x_n)_n$.*

Démonstration $\trianglerighteq$ Supposons que a soit limite de la suite $(x_{h(n)})_n$ extraite de la suite $(x_n)_n$. Pour montrer que a est valeur d'adhérence de la suite $(x_n)_n$ dans (E, d), il faut montrer que :

$$\forall r \in \mathbb{R}_+^* \quad \forall N \in \mathbb{N} \quad \exists n \in \mathbb{N} \quad \big(n \geqslant N \text{ et } x_{h(n)} \in B(a, r)\big).$$

Soient $r \in \mathbb{R}_+^*$ et $N \in \mathbb{N}$ fixés ; montrons que l'on peut trouver un entier naturel n tel que $n \geqslant N$ et $x_{h(n)} \in B(a, r)$. Comme $a = \lim_{n \to +\infty} x_{h(n)}$, d'après la définition 7.16 considérée avec $\varepsilon = \frac{1}{2}r$,

$$\exists N_0 \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad \big(n \geqslant N_0 \implies d(x_{h(n)}, a) \leqslant \tfrac{1}{2}r\big),$$

ce qui signifie qu'à partir du rang N_0 tous les termes de la suite sont dans la boule fermée $\overline{B}(a, \frac{1}{2}r)$ et donc dans $B(a, r)$. Considérons les entiers $n_0 = \max(N, N_0)$ et $n_1 = h(n_0)$; on a $n_1 \geqslant n_0$ (car h est une extractrice) donc, d'une part $n_1 \geqslant N$, et d'autre part $n_1 \geqslant N_0$, ce qui implique que x_{n_1} appartient à $B(a, r)$. L'entier n_1 vérifie donc les conditions requises. On en conclut que l'élément a est bien valeur d'adhérence de la suite $(x_n)_n$ dans (E, d).

$\trianglelefteq$ Réciproquement, supposons que a est valeur d'adhérence de la suite $(x_n)_n$ et construisons une suite extraite de la suite $(x_n)_n$ dont il est la limite. Construire une suite extraite de la suite $(x_n)_n$ qui converge vers a revient à trouver une extractrice h (c'est-à-dire une application strictement croissante de $\mathbb{N}$ dans $\mathbb{N}$) telle que la suite $(x_{h(n)})_n$ obtenue converge vers a. On construit cette extractrice par récurrence. On pose $h(0) = 0$. Supposons connue la valeur de $h(n-1)$ pour $n \in \mathbb{N}^*$; comme a est valeur d'adhérence de la suite $(x_n)_n$, d'après la définition 7.18 considérée avec $r = \frac{1}{n}$ et $N = h(n-1) + 1$

$$\exists m \in \mathbb{N} \quad \big(m \geqslant h(n-1) + 1 \text{ et } x_m \in B(a, \tfrac{1}{n})\big).$$

En posant $h(n) = m$, on a d'une part $h(n) > h(n-1)$ et d'autre part $0 \leqslant d(x_{h(n)}, a) < 1/n$. La première relation assure que h est bien une application strictement croissante de $\mathbb{N}$ dans $\mathbb{N}$ et par la théorème d'encadrement, la seconde relation indique que la suite $(x_{h(n)})_n$ extraite de la suite $(x_n)_n$ converge vers a. □

On a le résultat suivant qui est une simple conséquence du théorème 7.3 et de la proposition 7.15.

PROPOSITION 7.17 *Soient (E,d) un espace métrique et $(x_n)_n$ une suite d'éléments de E. Si la suite $(x_n)_n$ possède plus d'une valeur d'adhérence dans (E,d) alors elle diverge dans (E,d).*

Par contraposition, la proposition 7.17 fournit le résultat suivant : si une suite converge alors elle possède une unique valeur d'adhérence ; de plus cette valeur d'adhérence est sa limite car d'après la proposition 7.15, toutes les sous-suites convergent vers la limite. On sera toutefois attentif au fait que la réciproque est fausse. Par exemple la suite réelle de terme général $x_n = (1-(-1)^n)\,\mathrm{e}^n$ admet 0 pour seule valeur d'adhérence et pourtant elle diverge.

7.4.3 Suites de Cauchy

DÉFINITION 7.20 (Suite de Cauchy)
Soient (E,d) un espace métrique et $(x_n)_n$ une suite d'éléments de E. On dit que la suite $(x_n)_n$ est une suite de Cauchy dans (E,d) si :

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N_\varepsilon \in \mathbb{N} \quad \forall (m,n) \in \mathbb{N}^2 \quad (n \geqslant m \geqslant N_\varepsilon \implies d(x_n,x_m) \leqslant \varepsilon). \quad (2)$$

Remarques

1. On peut également exprimer que la suite $(x_n)_n$ est une suite de Cauchy dans (E,d) de la manière suivante :

$$\forall \varepsilon \in \mathbb{R}^+ \quad \exists N_\varepsilon \in \mathbb{N} \quad \forall (n,p) \in \mathbb{N}^2 \quad \big(n \geqslant N_\varepsilon \implies d(x_{n+p},x_n) \leqslant \varepsilon\big).$$

2. Pour $\mathbb{K}$ muni de la distance usuelle, la définition 7.20 coïncide avec celle de suite numérique de Cauchy donnée dans le *Cours de première année* p. 191. Par ailleurs, pour $\mathcal{C}^0([a,b],\mathbb{R})$ muni de la distance subordonnée à la norme $N_\infty : f \in \mathcal{C}^0([a,b],\mathbb{R}) \mapsto \sup_{x\in[a,b]} |f(x)|$ la définition 7.20 coïncide avec celle d'une suite d'applications uniformément de Cauchy donnée au chap. 2, voir p. 82.

3. En prenant la négation[67] de l'assertion (2), on obtient qu'une suite $(x_n)_n$ n'est pas une suite de Cauchy dans (E,d) si :

$$\exists \varepsilon \in \mathbb{R}_+^* \quad \forall N \in \mathbb{N} \quad \exists (m,n) \in \mathbb{N}^2 \quad (n \geqslant m \geqslant N \text{ et } d(x_n,x_m) > \varepsilon).$$

4. Toute suite extraite d'une suite de Cauchy est elle-même une suite de Cauchy. Cela résulte du fait que si h est une extractrice alors $h(n) \geqslant n$ pour tout $n \in \mathbb{N}$. ◇

[67] Voir le chap. 1 du *Cours de première année* pour les règles de négation d'une assertion quantifiée.

Exemples

1. Dans $\mathbb{R}$ muni de la distance $d : (x, y) \in \mathbb{R}^2 \mapsto |\arctan(x) - \arctan(y)|$, la suite $(u_n)_n$ de terme général $u_n = n$ est une suite de Cauchy. En effet, la relation[(68)] $\arctan(x) + \arctan(1/x) = \pi/2$ valable pour tout réel $x > 0$ implique que pour tout $(m, n) \in \mathbb{N}^2$ on a :

$$d(u_n, u_m) = |\arctan(n) - \arctan(m)| = |\arctan(1/n) - \arctan(1/m)|.$$

Dans le cas où $n = m$ on a $d(u_n, u_m) = 0$. Supposons que l'on ait $n > m$; le théorème des accroissements finis appliqué à la fonction arc-tangente sur l'intervalle $[\frac{1}{n}, \frac{1}{m}]$ établit l'existence d'un réel $c \in]\frac{1}{n}, \frac{1}{m}[$ tel que :

$$d(u_n, u_m) = \frac{1}{1+c^2}\left(\frac{1}{m} - \frac{1}{n}\right).$$

Puisque $1 + c^2 > 1$ et $0 < 1/n < 1/m$, on a $0 \leqslant d(u_n, u_m) \leqslant 1/m$. Ainsi, pour tout $\varepsilon \in \mathbb{R}_+^*$, en considérant l'entier $N = \mathrm{E}(\varepsilon^{-1}) + 1$, on obtient pour tout $(m, n) \in \mathbb{N}^2$ avec $n \geqslant m \geqslant N_\varepsilon$ la relation $d(u_n, u_m) \leqslant \varepsilon$. L'assertion (2) est démontrée et la suite $(u_n)_n$ est une suite de Cauchy.

2. La suite de terme général $u_n = n$ n'est pas une suite de Cauchy dans $\mathbb{R}$ muni de la distance usuelle car si $\varepsilon = \frac{1}{2}$ alors pour tout $N \in \mathbb{N}$ les entiers $n = N + 1$ et $m = N$ vérifient $n \geqslant m \geqslant N$ et $d(x_n, x_m) = |n - m| = 1 > \varepsilon$.

3. Considérons la suite $(f_n)_n$ de $\mathcal{C}^0([0,1])$ définie pour $n \geqslant 3$ par

$$f_n : x \in [0,1] \mapsto \begin{cases} 1 & \text{si } x \in [0, \frac{1}{2}] \\ n\left(\frac{1}{2} - \frac{1}{n} - x\right) & \text{si } x \in]\frac{1}{2}, \frac{1}{2} + \frac{1}{n}[\\ 0 & \text{si } x \in [\frac{1}{2} + \frac{1}{n}, 1] \end{cases}$$

et montrons qu'il s'agit d'une suite de Cauchy dans $(\mathcal{C}^0([0,1], \mathbb{R}), d_1)$ où d_1 est la distance subordonnée à la norme N_1 de la convergence en moyenne : $f \in \mathcal{C}^0([0,1], \mathbb{R}) \mapsto \int_0^1 |f(x)|\, \mathrm{d}x$. Soient ε un réel strictement positif fixé et m, n deux entiers tels que $n \geqslant m$. On a

$$\begin{aligned} d_1(f_n, f_m) &= m \int_{\frac{1}{2}}^{\frac{1}{2}+\frac{1}{m}} \left(x - \frac{1}{2} + \frac{1}{m}\right) \mathrm{d}x - n \int_{\frac{1}{2}}^{\frac{1}{2}+\frac{1}{n}} \left(x - \frac{1}{2} + \frac{1}{n}\right) \mathrm{d}x \\ &= \frac{1}{2m} - \frac{1}{2n} \leqslant \frac{1}{2m}. \end{aligned}$$

On en déduit que $d_1(f_n, f_m) \leqslant \varepsilon$ pour tout $(m, n) \in \mathbb{N}^2$ avec $n \geqslant m \geqslant N$ où $N = \mathrm{E}(\frac{1}{2\varepsilon}) + 1$.

Soit $(x_n)_n$ une suite d'éléments de E qui est de Cauchy dans (E, d_1) ; si d_2 est une distance équivalente à d_1 sur E alors la suite $(x_n)_n$ est aussi une suite de

(68) Voir la proposition 14.19 p. 655 du *Cours de première année*.

Cauchy dans (E, d_2). La vérification de cette propriété, qui découle de manière simple des définitions 7.20 et 7.2, est laissée en exercice.

Si d_1 et d_2 sont deux distances équivalentes sur E, les espaces métriques (E, d_1) et (E, d_2) ont les mêmes suites de Cauchy. Mais si les distances ne sont pas équivalentes, les suites de Cauchy ne sont pas nécessairement les mêmes pour les deux distances. On a vu dans les exemples précédents que dans $\mathbb{R}$ muni de la distance $(x, y) \in \mathbb{R}^2 \mapsto |\arctan(x) - \arctan(y)|$, la suite de terme général n est une suite de Cauchy, mais elle n'est pas de Cauchy avec la distance usuelle.

THÉORÈME 7.4 *Soient (E, d) un espace métrique et $(x_n)_n$ une suite d'éléments de E. Si la suite $(x_n)_n$ converge dans (E, d) alors c'est une suite de Cauchy dans (E, d).*

Démonstration Supposons que la suite $(x_n)_n$ converge vers ℓ dans (E, d) et montrons qu'elle est de Cauchy en établissant l'assertion (2). Soit ε un réel strictement positif ; montrons que l'on peut trouver un entier naturel N tel que pour tout $(m, n) \in \mathbb{N}^2$ avec $m \geqslant n \geqslant N$ on ait $d(x_n, x_m) \leqslant \varepsilon$. Puisque la suite $(x_n)_n$ converge vers ℓ dans (E, d), il résulte de la définition 7.16 que

$$\exists N_\varepsilon \in \mathbb{N} \quad \forall k \in \mathbb{N} \quad (k \geqslant N_\varepsilon \implies d(x_k, \ell) \leqslant \tfrac{1}{2}\varepsilon).$$

Pour tout $(m, n) \in \mathbb{N}^2$, la première inégalité triangulaire indique que :

$$d(x_n, x_m) \leqslant d(x_n, \ell) + d(\ell, x_m).$$

On en déduit que si $m \geqslant n \geqslant N_\varepsilon$ alors,

$$d(x_n, x_m) \leqslant \tfrac{1}{2}\varepsilon + \tfrac{1}{2}\varepsilon = \varepsilon.$$

La suite $(x_n)_n$ est donc bien une suite de Cauchy. □

Exercice 7.12 Soient (E, d) un espace métrique et $(x_n)_n$ une suite de Cauchy dans (E, d). Soit $(y_n)_n$ une suite d'éléments de E telle que $y_0 = x_0$ et telle que pour tout $n \in \mathbb{N}^*$ on ait $d(x_n, y_n) \leqslant 1/n$.
1 - Montrer que la suite $(y_n)_n$ est une suite de Cauchy dans (E, d).
2 - Montrer que la suite $(x_n)_n$ converge vers ℓ dans (E, d) si et seulement si la suite $(y_n)_n$ converge vers ℓ dans (E, d).

Le théorème 7.4 n'admet pas de réciproque : dans un espace métrique, il existe des suites de Cauchy qui ne convergent pas. Par exemple, on a vu que dans $\mathbb{R}$ muni de la distance $d : (x, y) \in \mathbb{R}^2 \mapsto |\arctan(x) - \arctan(y)|$, la suite de terme général $u_n = n$ est une suite de Cauchy et que cette suite diverge, voir p. 389.

Cette mise en garde motive la définition suivante.

DÉFINITION 7.21 (Espace métrique complet)
Un espace métrique (E, d) est qualifié d'espace métrique complet *si toute suite de Cauchy dans (E, d) est une suite convergente dans (E, d).*

Exemples

1. Dans $\mathbb{R}$ et dans $\mathbb{C}$ munis de leur distance usuelle $d : (x, y) \in \mathbb{R}^2 \mapsto |x - y|$, toute suite de Cauchy converge[69]. Il s'agit donc d'espaces métriques complets.

2. L'ensemble $\mathbb{R}$ muni de la distance[70] $\delta : (x, y) \in \mathbb{R}^2 \mapsto \frac{|x-y|}{1+|x-y|}$ est complet. Pour établir ce résultat, nous allons utiliser un moyen détourné et commencer par montrer que toute suite de Cauchy dans $(\mathbb{R}, \delta)$ est aussi une suite de Cauchy dans $(\mathbb{R}, d)$ où d désigne la distance usuelle. Considérons une suite $(x_n)_n$ de Cauchy dans $(\mathbb{R}, \delta)$; d'après la définition 7.20 on a[71] :

$$\forall \varepsilon \in \mathbb{R}_+^* \ \exists N_\varepsilon \in \mathbb{N} \ \forall (m, n) \in \mathbb{N}^2 \quad \left(n \geqslant m \geqslant N_\varepsilon \implies \frac{|x_n - x_m|}{1 + |x_n - x_m|} \leqslant \frac{\varepsilon}{1+\varepsilon} \right).$$

Comme

$$\begin{aligned} \frac{|x_n - x_m|}{1 + |x_n - x_m|} \leqslant \frac{\varepsilon}{1+\varepsilon} &\iff |x_n - x_m|\,(1+\varepsilon) \leqslant \varepsilon\,(1 + |x_n - x_m|) \\ &\iff |x_n - x_m| \leqslant \varepsilon, \end{aligned}$$

on en déduit que :

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N_\varepsilon \in \mathbb{N} \quad \forall (m, n) \in \mathbb{N}^2 \quad (n \geqslant m \geqslant N_\varepsilon \implies |x_n - x_m| \leqslant \varepsilon),$$

ce qui indique que la suite $(x_n)_n$ est de Cauchy pour la distance usuelle. Comme $\mathbb{R}$ muni de la distance usuelle est complet, la suite $(x_n)_n$ converge dans $(\mathbb{R}, d)$ vers un réel ℓ. Montrons à présent que la suite $(x_n)_n$ converge également vers ℓ dans $(\mathbb{R}, \delta)$; pour tout $n \in \mathbb{N}$, on a

$$\delta(x_n, \ell) = \frac{|x_n - \ell|}{1 + |x_n - \ell|} = \frac{d(x_n, \ell)}{1 + d(x_n, \ell)}.$$

Puisque $\lim_{n \to +\infty} d(x_n, \ell) = 0$, lorsque n tend vers $+\infty$, le numérateur dans l'expression précédente tend vers 0 et le dénominateur vers 1. On en déduit que $\lim_{n \to +\infty} \delta(x_n, \ell) = 0$, ce qui signifie que la suite $(x_n)_n$ converge vers ℓ dans $(\mathbb{R}, \delta)$.

3. L'ensemble $\mathbb{Q}$ muni de la distance usuelle n'est pas complet[72]. Cela prouve qu'un sous-espace métrique d'un espace métrique complet n'est pas nécessairement complet.

---

[69] Voir le théorème 5.4 p. 191 du *Cours de première année* pour la démonstration de la complétude de $\mathbb{R}$. La démonstration de la complétude de $\mathbb{C}$ est analogue en considérant le module à la place de la valeur absolue.

[70] Le fait que δ est une distance a été établi à l'exercice 7.2.

[71] On considère l'assertion de la définition 7.20 avec $\varepsilon = \frac{\varepsilon}{1+\varepsilon}$ (sic).

[72] Voir le *Cours de première année* p. 193.

4. L'intervalle $]0,1[$ muni de la distance usuelle sur $\mathbb{R}$ n'est pas complet ; on vérifiera par exemple que la suite de terme général $x_n = \frac{1}{n+2}$ est une suite de Cauchy dans $]0,1[$ mais qu'elle n'a pas de limite dans $]0,1[$. Par contre $[0,1]$ muni de la distance usuelle sur $\mathbb{R}$ est complet[73].

5. L'espace vectoriel $\mathcal{S}(\mathbb{R})$ des suites réelles muni de la distance[74] d définie pour $(x,y) = ((x_n)_n, (y_n)_n) \in \mathcal{S}(\mathbb{R})^2$ par $d(x,y) = \sum_{n=0}^{+\infty} \frac{1}{2^n} \frac{|x_n - y_n|}{1+|x_n-y_n|}$ est complet. Pour le montrer, considérons une suite $(X_n)_n$ de Cauchy dans $\mathcal{S}(\mathbb{R})$: c'est une suite dont les termes sont des suites réelles... ; le terme général de la suite $(X_n)_n$ est noté $X_n = (x_{n,k})_k$. La suite $(X_n)_n$ étant supposée de Cauchy, on a :

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N_\varepsilon \in \mathbb{N} \quad \forall (m,n) \in \mathbb{N}^2$$
$$\left(n \geqslant m \geqslant N_\varepsilon \implies \sum_{k=0}^{+\infty} \frac{1}{2^k} \frac{|x_{n,k} - x_{m,k}|}{1+|x_{n,k}-x_{m,k}|} \leqslant \varepsilon\right). \qquad (3)$$

La série étant à termes positifs, on en déduit que pour tout $k \in \mathbb{N}$:

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N_\varepsilon \in \mathbb{N} \quad \forall (m,n) \in \mathbb{N}^2$$
$$\left(n \geqslant m \geqslant N_\varepsilon \implies \frac{1}{2^k} \frac{|x_{n,k} - x_{m,k}|}{1+|x_{n,k}-x_{m,k}|} \leqslant \varepsilon\right).$$

Cette assertion indique que pour tout k fixé la suite $(x_{n,k})_n$ est une suite de Cauchy dans $\mathbb{R}$ muni de la distance[75] $\delta : (x,y) \in \mathbb{R}^2 \mapsto \frac{|x-y|}{1+|x-y|}$. Or comme $\mathbb{R}$ muni de la distance δ est complet (voir l'exemple précédent), pour tout k fixé la suite $(x_{n,k})_n$ converge dans $(\mathbb{R}, \delta)$ vers un réel que nous notons ℓ_k. En considérant la limite quand n tend vers $+\infty$ dans l'assertion (3), on obtient :

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N_\varepsilon \in \mathbb{N} \quad \forall m \in \mathbb{N} \quad \left(m \geqslant N_\varepsilon \implies \sum_{k=0}^{+\infty} \frac{1}{2^k} \frac{|\ell_k - x_{m,k}|}{1+|\ell_k - x_{m,k}|} \leqslant \varepsilon\right),$$

autrement dit que :

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N_\varepsilon \in \mathbb{N} \quad \forall m \in \mathbb{N} \quad \left(m \geqslant N_\varepsilon \implies d((X_n)_n, (\ell_n)_n) \leqslant \varepsilon\right).$$

Cela indique que la suite $(X_n)_n$ converge vers la suite $(\ell_n)_n$ dans $\mathcal{S}(\mathbb{R})$ muni de la distance d. L'espace vectoriel $\mathcal{S}(\mathbb{R})$ est donc complet pour la distance d.

Remarque Si d_1 et d_2 sont deux distances équivalentes sur E et si l'espace métrique (E,d_1) est complet alors (E,d_2) est lui aussi un espace métrique

---

[73] C'est en effet un fermé dans $\mathbb{R}$ muni de la distance usuelle et nous établirons à la proposition 7.19 p. 412 que tout sous-ensemble fermé d'un espace métrique complet est complet.

[74] Il a été établi à l'exercice 7.2 que l'application d était une distance.

[75] On vérifie sans peine que si d est une distance sur un espace métrique E alors pour tout réel strictement positif α, l'application $(x,y) \in E^2 \mapsto \alpha d(x,y)$ est aussi une distance et que cette distance est équivalente à d. Ceci explique pourquoi on « abandonne » le coefficient $\frac{1}{2^k}$.

BANACH, Stefan (1892, Cracovie - 1945, Lvov).



BANACH était un mathématicien polonais qui a posé les bases de l'analyse fonctionnelle. C'est dans sa thèse soutenue en 1922 qu'apparaît pour la première fois la notion d'espace de Banach et que sont démontrés les théorèmes fondamentaux concernant ces espaces. Cette thèse marque la naissance de « l'analyse fonctionnelle ». En 1929, il fonde avec HUGO STEINHAUS la revue mathématique[76] *Studia Mathematica*, consacrée au développement de l'analyse fonctionnelle.

complet. Attention ceci est faux si les distances ne sont pas équivalentes (voir les mises en garde précédentes pour des contre-exemples). ◇

Exercice 7.13 Soit (E, d) un espace métrique.
1 - Montrer que toute suite de Cauchy dans (E, d) est bornée.
2 - Montrer qu'une suite de Cauchy dans (E, d) qui possède une sous-suite convergente est une suite convergente.

DÉFINITION 7.22 *On appelle* espace de Banach *un espace vectoriel normé qui est complet (pour la distance subordonnée à sa norme).*

Exemples

1. L'espace vectoriel normé $(\mathcal{C}^0([a,b],\mathbb{R}), \|\cdot\|_\infty)$ est un espace de Banach. Considérons une suite $(f_n)_n$ qui est de Cauchy dans $(\mathcal{C}^0([a,b],\mathbb{R}), \|\cdot\|_\infty)$; elle vérifie :

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N_\varepsilon \in \mathbb{N} \quad \forall (m,n) \in \mathbb{N}^2 \quad \big(n \geqslant m \geqslant N_\varepsilon \implies \sup_{x\in[a,b]} |f_n(x) - f_m(x)| \leqslant \varepsilon\big), \tag{4}$$

autrement dit :

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N_\varepsilon \in \mathbb{N} \quad \forall (m,n) \in \mathbb{N}^2 \quad \forall x \in [a,b] \quad (n \geqslant m \geqslant N_\varepsilon \implies |f_n(x) - f_m(x)| \leqslant \varepsilon). \tag{5}$$

Cela implique, compte tenu de la place du quantificateur $(\forall x \in [a,b])$, que pour tout $x \in [a,b]$ on a :

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N_\varepsilon \in \mathbb{N} \quad \forall (m,n) \in \mathbb{N}^2 \quad (n \geqslant m \geqslant N_\varepsilon \implies |f_n(x) - f_m(x)| \leqslant \varepsilon).$$

[76] Voir `http ://journals.impan.gov.pl/sm`

Cette assertion se traduit par : pour tout $x \in [a,b]$, la suite numérique $(f_n(x))_n$ est une suite de Cauchy dans $\mathbb{R}$ muni de la distance usuelle d. Comme $(\mathbb{R}, d)$ est complet, cette suite converge vers un réel que nous notons $f(x)$. En considérant la limite lorsque n tend vers $+\infty$ dans l'assertion (5), on obtient :

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N_\varepsilon \in \mathbb{N} \quad \forall m \in \mathbb{N} \quad \forall x \in [a,b] \quad (m \geqslant N_\varepsilon \implies |f(x) - f_m(x)| \leqslant \varepsilon).$$

Ainsi, la suite $(f_n)_n$ converge uniformément vers f. Comme les applications f_n sont continues sur $[a,b]$, d'après le théorème 2.3 p. 86, f est continue sur $[a,b]$. On en conclut que toute suite de Cauchy dans $(\mathcal{C}^0([a,b],\mathbb{R}), \|\cdot\|_\infty)$ converge dans $(\mathcal{C}^0([a,b],\mathbb{R}), \|\cdot\|_\infty)$.

2. L'espace vectoriel $\mathcal{C}^0([a,b],\mathbb{R})$ muni de la norme $\|\cdot\|_1$ de la convergence en moyenne définie par $f \in \mathcal{C}^0([a,b],\mathbb{R}) \mapsto \int_0^1 |f(t)|\,\mathrm{d}t$ n'est pas un espace de Banach. Nous avons établi p. 399 que la suite $(f_n)_n$ de terme général

$$f_n : x \in [0,1] \mapsto \begin{cases} 1 & \text{si } x \in [0, \frac{1}{2}] \\ n\left(\frac{1}{2} - \frac{1}{n} - x\right) & \text{si } x \in]\frac{1}{2}, \frac{1}{2} + \frac{1}{n}[\\ 0 & \text{si } x \in [\frac{1}{2} + \frac{1}{n}, 1] \end{cases}$$

est de Cauchy dans $(\mathcal{C}^0([0,1],\mathbb{R}), \|\cdot\|_1)$. Cette suite converge simplement sur $[0,1]$ vers l'application f définie par $f(x) = 1$ si $x \in [0, \frac{1}{2}]$ et $f(x) = 0$ si $x \in]\frac{1}{2}, 1]$. L'application f n'est pas continue sur $[0,1]$; toutefois, elle est Riemann-intégrable sur $[0,1]$ car elle est continue par morceaux. Dans l'espace vectoriel de $\mathcal{RI}([0,1],\mathbb{R})$ des fonctions Riemann-intégrables sur $[0,1]$ muni de la norme $\|\cdot\|_1$, on a pour tout $n \in \mathbb{N}$:

$$\|f_n - f\|_1 = \int_0^1 |f_n(t) - f(t)|\,\mathrm{d}t = n \int_{\frac{1}{2}}^{\frac{1}{2}+\frac{1}{n}} \left(t - \frac{1}{2} + \frac{1}{n}\right)\,\mathrm{d}t = \frac{1}{n}$$

ce qui implique que $\lim_{n\to+\infty} \|f_n - f\|_1 = 0$. On en déduit que la suite $(f_n)_n$ converge dans $(\mathcal{RI}([0,1],\mathbb{R}), \|\cdot\|_1)$ vers f. Comme $\mathcal{C}^0([0,1],\mathbb{R})$ est un sous-espace vectoriel de $\mathcal{RI}([0,1])$, si $(f_n)_n$ convergeait dans $(\mathcal{C}^0([0,1],\mathbb{R}), \|\cdot\|_1)$ ce serait nécessairement vers f. Puisque f n'est pas continue sur $[0,1]$, on en conclut que la suite $(f_n)_n$ ne converge pas dans $(\mathcal{C}^0([0,1],\mathbb{R}), \|\cdot\|_1)$.

3. Montrons que muni de la norme $f \mapsto \sup_{x\in[a,b]} |f(x)| + \sup_{x\in[a,b]} |f'(x)|$, que l'on notera $\|\cdot\|$, l'espace vectoriel $\mathcal{C}^1([a,b],\mathbb{R})$ est un espace de Banach[77]. Soit $(f_n)_n$ une suite de Cauchy dans $(\mathcal{C}^1([a,b],\mathbb{R}), \|\cdot\|)$; on a

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N \in \mathbb{N} \quad \forall (m,n) \in \mathbb{N}^2 \quad \left(n \geqslant m \geqslant N \implies \|f_n - f_m\| \leqslant \varepsilon\right).$$

Puisque $\|f_n - f_m\| = \sup_{x\in[a,b]} |f_n(x) - f_m(x)| + \sup_{x\in[a,b]} |f_n'(x) - f_m'(x)|$, on en déduit que :

$$\forall \varepsilon \in \mathbb{R}_+^* \; \exists N \in \mathbb{N} \; \forall (m,n) \in \mathbb{N}^2 \quad \left(n \geqslant m \geqslant N \implies \sup_{x\in[a,b]} |f_n(x) - f_m(x)| \leqslant \varepsilon\right)$$

[77] L'espace vectoriel $\mathcal{C}^1([a,b],\mathbb{R})$ muni de la norme de la convergence uniforme n'est pas un espace de Banach.

c'est-à-dire que la suite $(f_n)_n$ est uniformément de Cauchy[78]. D'après le théorème 2.2 p. 82, la suite $(f_n)_n$ converge uniformément vers une application $f \in \mathcal{C}^0([a,b],\mathbb{R})$ et on a

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad (n \geqslant N \implies \sup_{x\in[a,b]} |f_n(x) - f(x)| \leqslant \varepsilon). \qquad (6)$$

Un raisonnement analogue permet d'établir que la suite $(f_n')_n$ converge uniformément vers une application g dans $\mathcal{C}^0([a,b],\mathbb{R})$. Le théorème 2.5 p. 92 permet d'affirmer que $g = f'$ ce qui en particulier implique que $f \in \mathcal{C}^1([a,b],\mathbb{R})$. On a ainsi,

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad (n \geqslant N \implies \sup_{x\in[a,b]} |f_n'(x) - f'(x)| \leqslant \varepsilon) \qquad (7)$$

et on déduit des assertions (6) et (7), considérées avec $\frac{1}{2}\varepsilon$, que

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad (n \geqslant N \implies \|f_n - f\| \leqslant \varepsilon).$$

Ceci permet de conclure que $(f_n)_n$ converge vers f dans $(\mathcal{C}^1([a,b],\mathbb{R}), \|\cdot\|)$.

Exercice 7.14 En considérant la suite de fonctions $(f_n)_n$ définies p. 404, montrer que l'espace vectoriel $\mathcal{C}^0([0,1],\mathbb{R})$ muni de la norme $\|\cdot\|_2$ de la converge quadratique n'est pas un espace de Banach.

Lemme 7.3 *Pour tout $p \in \mathbb{N}^*$, les espaces vectoriels $\mathbb{R}^p$ et $\mathbb{C}^p$ munis de la norme infinie sont des espaces de Banach.*

Démonstration Désignons par $\mathbb{K}$ l'espace vectoriel $\mathbb{R}$ ou $\mathbb{C}$ et par $|\cdot|$ la valeur absolue lorsque $\mathbb{K} = \mathbb{R}$ ou le module lorsque $\mathbb{K} = \mathbb{C}$. Considérons une suite de Cauchy $(\mathbf{x}_n)_n$ dans $\mathbb{K}^p$ muni de la norme infinie $\|\cdot\|_\infty$; on a[79]

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N \in \mathbb{N} \quad \forall (n,m) \in \mathbb{N}^2 \quad (n \geqslant m \geqslant N \implies \|\mathbf{x}_n - \mathbf{x}_m\|_\infty \leqslant \varepsilon).$$

Pour tout $n \in \mathbb{N}$, désignons par $(x_n^{(1)}, \ldots, x_n^{(p)})$ les composantes de $\mathbf{x}_n$. Puisque

$$\|\mathbf{x}_n - \mathbf{x}_m\|_\infty = \max_{i\in\{1,\ldots,p\}} |x_n^{(i)} - x_m^{(i)}|,$$

on a $\|\mathbf{x}_n - \mathbf{x}_m\|_\infty \leqslant \varepsilon$ si et seulement si $|x_n^{(i)} - x_m^{(i)}| \leqslant \varepsilon$ pour tout $i \in \{1,\ldots,p\}$. On en déduit que la suite $(\mathbf{x}_n)_n$ est de Cauchy si et seulement si :

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N \in \mathbb{N} \quad \forall (n,m) \in \mathbb{N}^2 \quad \forall i \in \{1,\ldots,p\} \quad (n \geqslant m \geqslant N \implies |x_n^{(i)} - x_m^{(i)}| \leqslant \varepsilon).$$

[78] Voir la définition 2.3 p. 82.

[79] Voir la définition 7.20 p. 398.

Cette assertion indique que la suite $(x_n^{(i)})_n$, dont les termes sont formés par la i–ième composante des termes de la suite $(\mathbf{x}_n)_n$, est une suite de Cauchy dans $\mathbb{K}$. Comme $\mathbb{R}$ (resp. $\mathbb{C}$) muni de la distance usuelle est complet[80], on en déduit que la suite $(x_n^{(i)})_n$ converge. En désignant par $\ell^{(i)}$ la limite de la suite $(\mathbf{x}_n)_n$, on a donc[81]

$$\forall i \in \{1,\dots,p\} \quad \forall \varepsilon \in \mathbb{R}_+^* \quad \exists N_i \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad (n \geqslant N_i \implies |x_n^{(i)} - \ell^{(i)}| \leqslant \varepsilon).$$

En notant $N = \max_{i\in\{1,\dots,p\}} N_i$, on en déduit que

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad \forall i \in \{1,\dots,p\} \quad (n \geqslant N \implies |x_n^{(i)} - \ell^{(i)}| \leqslant \varepsilon).$$

Si pour tout $i \in \{1,\dots,p\}$ on a $|x_n^{(i)} - \ell^{(i)}| \leqslant \varepsilon$ alors $\max_{i\in\{1,\dots,p\}} |x_n^{(i)} - \ell^{(i)}| \leqslant \varepsilon$; en notant $\ell = (\ell_1,\dots,\ell_p) \in \mathbb{K}^p$, l'assertion précédente se réécrit

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad (n \geqslant N \implies \|x_n - \ell\|_\infty \leqslant \varepsilon).$$

D'après la définition 7.16, on en conclut que la suite $(\mathbf{x}_n)_n$ converge vers ℓ. □

Nous sommes à présent en mesure de démontrer le résultat suivant dont un corollaire est le théorème 7.1 énoncé p. 367.

LEMME 7.4 *Dans l'espace vectoriel $\mathbb{K}^p$ ($p \in \mathbb{N}^*$), toutes les normes sont équivalentes à la norme infinie*

$$(x_1,\dots,x_p) \in \mathbb{K}^p \longmapsto \max_{i\in\{1,\dots,p\}} |x_i|$$

Démonstration Nous allons montrer par récurrence[82] que pour tout $p \in \mathbb{N}^*$, une norme ν sur $\mathbb{K}^p$ est équivalente à la norme infinie sur $\mathbb{K}^p$ que l'on notera $\|\cdot\|_{\infty,\mathbb{K}^p}$.

⊵ Si $p = 1$ alors pour tout $x \in \mathbb{K}$ on a : $\nu(x) = \nu(x\cdot 1) = |x|\,\nu(1) = \nu(1)\,\|x\|_{\infty,\mathbb{K}}$. L'équivalence des deux normes est établie.

⊵ Supposons que sur $\mathbb{K}^{p-1}$ toutes les normes sont équivalentes à $\|\cdot\|_{\infty,\mathbb{K}^{p-1}}$. Considérons une norme ν sur $\mathbb{K}^p$ et montrons qu'elle est équivalente à $\|\cdot\|_{\infty,\mathbb{K}^p}$, *i.e.* montrons qu'il existe $C_1, C_2 \in \mathbb{R}_+^*$ tels que $C_1\,\nu(\mathbf{x}) \leqslant \|\mathbf{x}\|_{\infty,\mathbb{K}^p} \leqslant C_2\,\nu(\mathbf{x})$.

▷ Munissons $\mathbb{K}^p$ de sa base canonique $(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_p)$ et considérons un élément $\mathbf{x}$ de $\mathbb{K}^p$ de composantes $(x_1, x_2, \dots, x_p)$, *i.e.* $\mathbf{x} = \sum_{i=1}^p x_i\,\mathbf{e}_i$. On a :

$$\nu(\mathbf{x}) = \nu\Big(\sum_{i=1}^p x_i\,\mathbf{e}_i\Big) \leqslant \sum_{i=1}^p \nu(x_i\,\mathbf{e}_i) = \sum_{i=1}^p |x_i|\,\nu(\mathbf{e}_i) \leqslant \|\mathbf{x}\|_{\infty,\mathbb{K}^p} \sum_{i=1}^p \nu(\mathbf{e}_i),$$

[80] Voir le théorème 5.4 p. 191 du *Cours de première année*.

[81] Voir la définition 7.16 p. 388.

[82] D'autres démonstrations existent. Cette démonstration par récurrence s'inspire de S. Lang, *Undergraduate analysis*, Springer Verlag, New York, 1983.

la première des inégalités résultant de l'inégalité triangulaire, la seconde du fait que pour tout $i \in \{1, \dots, p\}$ on a $|x_i| \leqslant \|\mathbf{x}\|_{\infty,\mathbb{K}^p}$. Ainsi, si l'on note C le réel strictement positif $\sum_{i=1}^{p} \nu(\mathbf{e}_i)$ on a $\nu(\mathbf{x}) \leqslant C\,\|\mathbf{x}\|_{\infty,\mathbb{K}^p}$ et $\|\mathbf{x}\|_{\infty,\mathbb{K}^p} \geqslant \frac{1}{C}\,\nu(\mathbf{x})$. On remarquera que pour obtenir cette première inégalité, il n'a pas été fait usage de l'hypothèse de récurrence.

▷ Montrons à présent qu'il existe un réel strictement positif C_2 tel que pour tout $\mathbf{x} \in \mathbb{K}^p$ on ait $\|\mathbf{x}\|_{\infty,\mathbb{K}^p} \leqslant C_2\,\nu(\mathbf{x})$. Pour cela raisonnons par l'absurde et supposons qu'un tel réel n'existe pas, c'est-à-dire supposons que :

$$\forall C_2 \in \mathbb{R}_+^* \quad \exists \mathbf{x} \in \mathbb{K}^p \quad \|\mathbf{x}\|_{\infty,\mathbb{K}^p} > C_2 \nu(\mathbf{x}).$$

Remarquons que nécessairement $\mathbf{x}$ est non nul car $\|0\|_{\infty,\mathbb{K}^p} = \nu(0) = 0$. On déduit (en prenant des réels C_2 de la forme $C_2 = n$ où $n \in \mathbb{N}^*$) que :

$$\forall n \in \mathbb{N}^* \quad \exists \mathbf{x}^{(n)} \in \mathbb{K}^p \setminus \{0\} \quad \|\mathbf{x}^{(n)}\|_{\infty,\mathbb{K}^p} > n\nu(\mathbf{x}^{(n)}). \tag{8}$$

Fixons n et notons i_0 l'entier vérifiant $|x_{i_0}^{(n)}| = \max_{i\in\{1,\dots,p\}} |x_i^{(n)}|$; on a $|x_{i_0}^{(n)}| = \|\mathbf{x}^{(n)}\|_{\infty,\mathbb{K}^p}$. Le scalaire $x_{i_0}^{(n)}$ est non nul (car $\mathbf{x}^{(n)}$ est non nul) et si l'on note $\hat{\mathbf{x}}^{(n)} = \mathbf{x}^{(n)}/x_{i_0}^{(n)}$ on obtient

$$\|\hat{\mathbf{x}}^{(n)}\|_{\infty,\mathbb{K}^p} = \frac{\|\mathbf{x}^{(n)}\|_{\infty,\mathbb{K}^p}}{|x_{i_0}^{(n)}|} = 1 \quad \text{et} \quad \nu(\hat{\mathbf{x}}^{(n)}) = \frac{\nu(\mathbf{x}^{(n)})}{|x_{i_0}^{(n)}|}.$$

On déduit alors de l'assertion (8) que :

$$\forall n \in \mathbb{N}^* \quad \exists \hat{\mathbf{x}}^{(n)} \in \mathbb{K}^p \setminus \{0\} \quad \left(\|\hat{\mathbf{x}}^{(n)}\|_{\infty,\mathbb{K}^p} = 1 \ \text{ et } \ \nu(\hat{\mathbf{x}}^{(n)}) < 1/n\right).$$

L'entier i_0 dépend bien entendu de n et prend ses valeurs dans l'ensemble $I = \{1, \dots, p\}$. Lorsque n décrit $\mathbb{N}$, comme I est fini, il y a (au moins) une valeur de l'ensemble I qui correspond à l'indice i_0 un nombre infini de fois. Autrement dit, il existe $j_0 \in \{1, \dots, p\}$ tel que l'ensemble $S = \{n \in \mathbb{N}^* \ ; \ i_0(n) = j_0\}$ est infini. On a donc démontré l'assertion suivante :

$$\exists j_0 \in \{1, \dots, p\} \quad \forall n \in S \quad \exists \hat{\mathbf{x}}^{(n)} \in \overline{B'}(0,1) \quad \left(\hat{x}_{j_0}^{(n)} = 1 \ \text{ et } \ \nu(\hat{\mathbf{x}}^{(n)}) < 1/n\right),$$

où $\overline{B'}(0,1)$ désigne la boule unité fermée privée de 0 dans $\mathbb{K}^p$ muni de la norme $\|\cdot\|_{\infty,\mathbb{K}^p}$. Comme S est un sous-ensemble infini de $\mathbb{N}$, on peut extraire de la suite $(\hat{\mathbf{x}}^{(n)})_n$ une sous-suite $(\tilde{\mathbf{x}}^{(n)})_n$ construite en ne considérant que les termes d'indice n appartenant à S. On obtient alors l'assertion suivante :

$$\exists j_0 \in \{1, \dots, p\} \quad \forall n \in \mathbb{N} \quad \exists \tilde{\mathbf{x}}^{(n)} \in \overline{B'}(0,1) \quad \left(\tilde{x}_{j_0}^{(n)} = 1 \ \text{ et } \ \nu(\tilde{\mathbf{x}}^{(n)}) < 1/n\right).$$

Considérons à présent le sous-espace vectoriel $F = \{\mathbf{x} \in \mathbb{K}^p \ ; \ x_{j_0} = 0\}$ de dimension $p-1$ dans $\mathbb{K}^p$ que l'on identifie[83] à $\mathbb{K}^{p-1}$. En restriction à F, les

[83] Voir p. 363.

deux applications ν et $\|\cdot\|_{\infty,\mathbb{K}^p}$ induisent des normes qui, par hypothèse de récurrence, sont équivalentes. Désignons par ν_F la norme qui est issue de la restriction à F de ν.

Comme pour tout $n \in \mathbb{N}$ on a $\tilde{x}_{j_0}^{(n)} = 1$, il existe un élément $\mathbf{y}^{(n)}$ dans F tel que $\tilde{\mathbf{x}}^{(n)} = \mathbf{e}_{j_0} + \mathbf{y}^{(n)}$. Intéressons-nous à la suite $(\mathbf{y}^{(n)})_n$. On a pour tout $(m,n) \in \mathbb{N}^* \times \mathbb{N}^*$,

$$\begin{aligned} \nu_F(\mathbf{y}^{(n)} - \mathbf{y}^{(m)}) &= \nu_F(\mathbf{y}^{(n)} + \mathbf{e}_{j_0} - (\mathbf{y}^{(m)} + \mathbf{e}_{j_0})) = \nu_F(\mathbf{x}_n - \mathbf{x}_m) \\ &\leqslant \nu_F(\mathbf{x}_n) + \nu_F(\mathbf{x}_m) < 1/n + 1/m. \end{aligned}$$

Pour tout $\varepsilon \in \mathbb{R}_+^*$, en prenant $N = \mathrm{E}(2/\varepsilon) + 1$, on déduit de l'assertion précédente que :

$$\forall (m,n) \in \mathbb{N}^* \times \mathbb{N}^* \quad \Big(n \geqslant m \geqslant N \implies \nu_F(\mathbf{y}^{(n)} - \mathbf{y}^{(m)}) \leqslant \varepsilon\Big),$$

autrement dit, que la suite $(\mathbf{y}^{(n)})_n$ est de Cauchy dans $\mathbb{K}^{p-1}$ pour la norme ν_F. Comme sur $\mathbb{K}^{p-1}$ toutes les normes sont équivalentes à la norme $\|\cdot\|_{\infty,\mathbb{K}^{p-1}}$, la suite $(\mathbf{y}^{(n)})_n$ est aussi de Cauchy pour la norme $\|\cdot\|_{\infty,\mathbb{K}^{p-1}}$. Puisque $\mathbb{K}^{p-1}$ est complet pour la norme $\|\cdot\|_{\infty,\mathbb{K}^{p-1}}$, la suite $(\mathbf{y}^{(n)})_n$ converge pour la norme $\|\cdot\|_{\infty,\mathbb{K}^{p-1}}$ dans F vers $\mathbf{y} \in F$. Elle converge aussi vers $\mathbf{y}$ pour la norme ν_F car ces deux normes sont équivalentes sur F. Pour tout $n \in \mathbb{N}^*$ on obtient,

$$\nu_F(\mathbf{y} + \mathbf{e}_{j_0}) = \nu_F(\mathbf{y} - \mathbf{y}^{(n)} + \mathbf{y}^{(n)} + \mathbf{e}_{j_0}) \leqslant \nu_F(\mathbf{y} - \mathbf{y}^{(n)}) + \underbrace{\nu_F(\mathbf{y}^{(n)} + \mathbf{e}_{j_0})}_{= \nu_F(\tilde{\mathbf{x}}^{(n)})}$$

soit : $0 \leqslant \nu_F(\mathbf{y} + \mathbf{e}_{j_0}) \leqslant \nu_F(\mathbf{y} - \mathbf{y}^{(n)}) + 1/n$.

Comme cette inégalité est vraie pour tout $n \in \mathbb{N}$, en considérant la limite du membre de droite lorsque n tend vers $+\infty$, on obtient $\nu_F(\mathbf{y} + \mathbf{e}_{j_0}) = 0$. Puisque ν_F est une norme cela implique que $\mathbf{y} + \mathbf{e}_{j_0} = 0$ soit $y_{j_0} = -1$. On a obtenu la contradiction : comme $\mathbf{y} \in F$ on devrait avoir $y_{j_0} = 0$. On en conclut qu'il existe bien un réel strictement positif C_2 tel que pour tout $\mathbf{x} \in \mathbb{K}^p$ on ait $\|\mathbf{x}\|_{\infty,\mathbb{K}^p} \leqslant C_2\, \nu(\mathbf{x})$, ce qui achève d'établir l'équivalence des deux normes. □

Il résulte du lemme précédent que pour l'espace vectoriel $\mathbb{K}^p$ ($p \in \mathbb{N}^*$) toutes les normes sont équivalentes. Comme par ailleurs nous avons établi que l'espace vectoriel $\mathbb{K}^p$ muni de la norme infinie est complet, il en résulte que $\mathbb{K}^p$ muni de l'une quelconque de ses normes est complet.

Proposition 7.18 *Les espaces vectoriels $\mathbb{R}^p$ et $\mathbb{C}^p$ ($p \in \mathbb{N}^*$) munis de l'une quelconque de leurs normes sont des espaces de Banach.*

Plus généralement, on a le résultat suivant qui indique que tout espace vectoriel normé de dimension finie est complet.

Théorème 7.5 (Complétude en dimension finie)
Tous les espaces vectoriels normés de dimension finie sont complets.

Démonstration Soient E un $\mathbb{K}$-espace vectoriel de dimension $p \in \mathbb{N}^*$ et $\mathcal{B} = (\mathbf{e}_1, \dots, \mathbf{e}_p)$ la base canonique de E. On munit E de la norme infinie définie pour tout vecteur $\mathbf{x}$ de E de composantes $(x^{(1)}, \dots, x^{(p)}) \in \mathbb{K}^p$ par

$$\|\mathbf{x}\|_{\infty,E} = \max_{i=1,\dots,p} |x^{(i)}| = \|(x^{(1)}, \dots, x^{(p)})\|_{\infty,\mathbb{K}^n}.$$

Considérons une suite de Cauchy $(\mathbf{x}_n)_n$ dans $(E, \|\cdot\|_{\infty,E})$ et désignons par $(x_n^{(1)}, \dots, x_n^{(p)})$ les composantes du terme général $\mathbf{x}_n$ dans la base $\mathcal{B}$. On a[84]

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N \in \mathbb{N} \quad \forall (n,m) \in \mathbb{N}^2 \quad (n \geqslant m \geqslant N \implies \|\mathbf{x}_n - \mathbf{x}_m\|_{\infty,E} \leqslant \varepsilon).$$

Puisque

$$\|\mathbf{x}_n - \mathbf{x}_m\|_{\infty,E} = \max_{i\in\{1,\dots,p\}} |x_n^{(i)} - x_m^{(i)}| = \|(x_n^{(1)}, \dots, x_n^{(p)}) - (x_m^{(1)}, \dots, x_m^{(p)})\|_{\infty,\mathbb{K}^n},$$

on déduit de cette assertion que la suite de terme général $(x_n^{(1)}, \dots, x_n^{(p)})$ est une suite de Cauchy dans $\mathbb{K}^p$ muni de la norme infinie $\|\cdot\|_{\infty,\mathbb{K}^p}$. Comme $\mathbb{K}^p$ est complet, cette suite converge vers un élément $(\ell^{(1)}, \dots, \ell^{(p)})$ de $\mathbb{K}^p$ pour la norme infinie, ce qui ce traduit par :

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad (n \geqslant N \implies \max_{i\in\{1,\dots,p\}} |x_n^{(i)} - \ell^{(i)}| \leqslant \varepsilon).$$

Notons ℓ l'élément de E de composantes $(\ell^{(1)}, \dots, \ell^{(p)})$ dans la base $\mathcal{B}$; on a $\max_{i\in\{1,\dots,p\}} |x_n^{(i)} - \ell^{(i)}| = \|\mathbf{x} - \ell\|_{\infty,E}$ et l'assertion précédente qui se réécrit

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad (n \geqslant N \implies \|\mathbf{x} - \ell\|_{\infty,E} \leqslant \varepsilon)$$

indique que la suite $(\mathbf{x}_n)_n$ converge vers ℓ dans E pour la norme $\|\cdot\|_{\infty,E}$. □

Remarque En adaptant la démonstration qui vient d'être faite, on peut montrer que si $(E_1, \delta_1), \dots, (E_p, \delta_p)$ sont des espaces métriques complets alors le produit cartésien $E = E_1 \times \cdots \times E_p$ muni d'une des distances produit (voir p. 356) est un espace métrique complet. ◇

Exercice 7.15 On considère l'espace vectoriel réel E des applications de $[0,1]$ dans $\mathbb{R}$, de classe $\mathcal{C}^1$, s'annulant en 0.
1 - Montrer que $N : f \in E \mapsto \sup_{x\in[0,1]} |f(x)| + \sup_{x\in[0,1]} |f'(x)|$ est une norme sur E.
2 - On rappelle que $(\mathcal{C}^0([0,1],\mathbb{R}), \|\cdot\|_\infty)$ est un espace de Banach, voir p. 403.
a) Montrer que si $(f_n)_n$ est une suite de Cauchy dans (E,N), elle converge dans $(\mathcal{C}^0([0,1],\mathbb{R}), \|\cdot\|_\infty)$.
b) Montrer que (E,N) est un espace de Banach.

[84] Voir la définition 7.20 p. 398.

7.4.4 Caractérisation des fermés dans un espace métrique

THÉORÈME 7.6 *Soient (E,d) un espace métrique et A un sous-ensemble de E.*

✗ *Un élément a de E appartient à l'adhérence $\overline{A}$ de A si et seulement si il est la limite d'une suite d'éléments de A.*

✗ *Une condition nécessaire et suffisante pour que A soit un ensemble fermé dans (E,d) est que toutes les suites $(x_n)_n$ d'éléments de A convergentes aient pour limite un élément de A.*

Démonstration ⊵ Supposons que a soit la limite d'une suite $(x_n)_n$ d'éléments de A et montrons que $a \in \overline{A}$. Pour cela, voir la définition 7.9, montrons que toute boule ouverte de centre a contient au moins un élément de A, c'est-à-dire que pour tout $r \in \mathbb{R}_+^*$ l'ensemble $B(a,r) \cap A$ est non vide. En prenant $\varepsilon = \frac{r}{2}$ dans la définition 7.16, on obtient

$$\exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad (n \geqslant N \implies d(x_n, a) \leqslant \tfrac{r}{2}).$$

ce qui signifie que pour tout $n \geqslant N$, $x_n \in \overline{B}(a, \frac{r}{2})$. Puisque $\overline{B}(a, \frac{r}{2}) \subset B(a,r)$, à partir du rang N tous les éléments de la suite $(x_n)_n$ (qui sont des éléments de A) appartiennent à $B(a,r)$. La boule $B(a,r)$ contient donc des éléments de A.

⊵ Inversement, supposons que $a \in \overline{A}$ et construisons une suite $(x_n)_n$ d'éléments de A qui converge vers a. Puisque $a \in \overline{A}$, toute boule ouverte de centre a et de rayon $\frac{1}{n}$ où $n \in \mathbb{N}^*$ contient au moins un élément de A ; notons x_n l'un de ces éléments. Comme $x_n \in B(a, \frac{1}{n})$, on a $d(x_n, a) < \frac{1}{n}$ et par conséquent la suite $(x_n)_n$ converge vers a.

⊵ La seconde propriété résulte de la première propriété qui vient d'être démontré et du fait que l'ensemble A est fermé dans (E,d) si et seulement si $\overline{A} = A$. □

Exemples

1. Le théorème 7.6 va nous permettre de montrer facilement que si (F, d_F) est un espace métrique alors l'ensemble $B = \{(y_1, y_2) \in F^2 \ ; \ y_1 = y_2\}$ est un fermé dans F^2 (muni d'une des distances produit). Considérons une suite $(u_n)_n$ d'éléments de B qui converge vers l'élément $\ell = (\ell_1, \ell_2)$ de F^2 et montrons que ℓ appartient à B, c'est-à-dire que $\ell_1 = \ell_2$. Pour tout $n \in \mathbb{N}$ on a $u_n = (a_n, b_n)$ avec $(a_n, b_n) \in F^2$ et $a_n = b_n$. Puisque la suite $(u_n)_n$ converge vers $\ell = (\ell_1, \ell_2)$, la suite $(a_n)_n$ converge vers ℓ_1 et la suite $(b_n)_n$ converge vers ℓ_2. Comme les deux suites $(a_n)_n$ et $(b_n)_n$ sont égales, par unicité de la limite, elles convergent vers le même élément : on a donc $\ell_1 = \ell_2$.

2. L'espace vectoriel $\mathcal{C}^0([a,b], \mathbb{R})$ n'est pas un sous-espace vectoriel fermé dans l'espace vectoriel $\mathcal{RI}([0,1], \mathbb{R})$ des fonctions Riemann-intégrables sur $[0,1]$ muni

de la distance subordonnée à la norme $N_1 : f \mapsto \int_0^1 |f(t)|\,\mathrm{d}t$ de la convergence en moyenne. En effet, nous avons vu (voir p. 403) que la suite $(f_n)_n$ de $\mathcal{C}^0([0,1],\mathbb{R})$ définie pour $n \geqslant 3$ par

$$f_n : x \in [0,1] \mapsto \begin{cases} 1 & \text{si } t \in [0, \frac{1}{2}] \\ n\left(\frac{1}{2} - \frac{1}{n} - x\right) & \text{si } x \in]\frac{1}{2}, \frac{1}{2} + \frac{1}{n}[\\ 0 & \text{si } t \in [\frac{1}{2} + \frac{1}{n}, 1] \end{cases}$$

converge dans $(\mathcal{RI}([0,1],\mathbb{R}), N_1)$ vers la fonction f définie par $f(x) = 1$ si $x \in [0, \frac{1}{2}]$ et $f(x) = 0$ si $x \in]\frac{1}{2}, 1[$. Cette fonction n'étant pas continue en $\frac{1}{2}$, elle n'appartient pas à $\mathcal{C}^0([0,1],\mathbb{R})$.

3. L'ensemble $\mathcal{B}(I,\mathbb{R})$ des applications de I dans $\mathbb{R}$ qui sont bornées est fermé dans l'ensemble $\mathcal{A}(I,\mathbb{R})$ des applications de I dans $\mathbb{R}$ muni de la distance d_∞ subordonnée à la norme $\| \ \|_\infty$ de la convergence uniforme. Pour le montrer, considérons une suite $(f_n)_n$ d'éléments de l'ensemble $\mathcal{B}(I,\mathbb{R})$ qui converge vers une application f dans $(\mathcal{A}(I,\mathbb{R}), \| \ \|_\infty)$ et montrons que f est bornée. Puisque pour tout $n \in \mathbb{N}$, l'application f_n est bornée[85],

$$\forall n \in \mathbb{N} \quad \exists M_n \in \mathbb{R}^+ \quad \forall x \in I \quad |f_n(x)| \leqslant M_n$$

et puisque[86] la suite $(f_n)_n$ converge uniformément vers f,

$$\exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad (n \geqslant N \implies \|f_n - f\|_{\infty,I} \leqslant 1).$$

On en déduit, en ayant recours à l'inégalité triangulaire, que :

$$\|f\|_{\infty,I} \leqslant \|f - f_N\|_{\infty,I} + \|f_N\|_{\infty,I} \leqslant 1 + M_N,$$

ce qui prouve que l'application f est bornée.

4. Dans un espace vectoriel normé $(E, \|\cdot\|)$, la sphère $S = \{\mathbf{x} \in E \ ; \ \|\mathbf{x}\| = 1\}$ est un fermé. En effet, considérons une suite $(\mathbf{x}_n)_n$ d'éléments de S qui converge vers $\mathbf{x}$ dans $(E, \|\cdot\|)$. Pour tout $n \in \mathbb{N}$, on obtient en ayant recours à la seconde inégalité triangulaire :

$$\big|\|\mathbf{x}_n\| - 1\big| = \big|\|\mathbf{x}_n\| - \|\mathbf{x}\|\big| \leqslant \|\mathbf{x}_n - \mathbf{x}\|,$$

et par conséquent $1 - \|\mathbf{x}_n - \mathbf{x}\| \leqslant \|\mathbf{x}\| \leqslant 1 + \|\mathbf{x}_n - \mathbf{x}\|$. Puisque la suite $(\mathbf{x}_n)_n$ converge vers $\mathbf{x}$ on a $\lim_{n\to+\infty} \|\mathbf{x}_n - \mathbf{x}\| = 0$ et on conclut par le théorème d'encadrement que $\|\mathbf{x}\| = \lim_{n\to+\infty} \|\mathbf{x}\| = 1$, autrement dit que $\mathbf{x} \in S$.

Corollaire 7.2 *Soient (E,d) un espace métrique et A un sous-ensemble de E. A est dense dans E si et seulement si tout élément de E est la limite d'une suite d'éléments de A.*

[85] Voir le *Cours de première année* p. 566.

[86] Voir la définition 2.2 p. 73 ; on prend ici $\varepsilon = 1$.

Démonstration D'après la définition 7.10, A est dense dans E si $\overline{A} = E$ et d'après le théorème 7.6, un élément a de E appartient à l'adhérence $\overline{A}$ de A si et seulement s'il est la limite d'une suite d'éléments de A. □

Nous avons vu p. 401 qu'un sous-espace métrique d'un espace métrique complet n'était pas nécessairement complet. La proposition suivante fournit une condition nécessaire et suffisante pour qu'un sous-espace métrique d'un espace métrique complet soit lui-même complet.

PROPOSITION 7.19 *Soient (E,d) un espace métrique complet et A un sous-ensemble de E. Une condition nécessaire et suffisante pour que (A,d) soit un sous-espace métrique complet est que A soit fermé dans (E,d).*

Démonstration $\trianglerighteq$ Nous allons établir l'assertion : « (A,d) est un espace métrique complet implique que A est fermé dans (E,d) », en montrant que sa contraposée est vraie. Supposons donc que A n'est pas un ensemble fermé dans (E,d). D'après le théorème 7.6, il existe alors (au moins) une suite $(x_n)_n$ d'éléments de A qui converge dans (E,d) et dont la limite n'appartient pas à A (elle appartient à $\complement_E A$). Comme la suite $(x_n)_n$ converge dans (E,d), c'est une suite de Cauchy dans (E,d) et par conséquent dans (A,d) car $A \subset E$. On en conclut que (A,d) n'est pas un espace métrique complet car on dispose d'une suite d'éléments de A qui est de Cauchy dans (A,d) mais qui ne converge pas dans (A,d).

$\trianglerighteq$ Montrons l'implication réciproque : « A est fermé dans (E,d) implique que (A,d) est un espace métrique complet ». Soit $(x_n)_n$ une suite d'éléments de A qui est de Cauchy dans (A,d). Puisque $A \subset E$, cette suite est aussi une suite d'éléments de E et, comme (E,d) est complet, cette suite de Cauchy converge dans (E,d). Comme il s'agit d'une suite d'éléments de A et que A est fermé, sa limite est un élément de A. On en déduit que la suite $(x_n)_n$ converge dans (A,d) et par conséquent que (A,d) est complet. □

Remarque La première partie de la démonstration, qui n'utilise pas l'hypothèse de complétude de (E,d), établit le résultat suivant. Dans un espace métrique (E,d), un sous-espace métrique complet (A,d) est nécessairement fermé. La contraposée nous indique qu'un sous-espace métrique qui n'est pas fermé ne peut être complet. On établit par ce moyen la non-complétude de $\mathbb{Q}$ pour la distance usuelle : il n'est pas fermé dans $\mathbb{R}$ (on a $\overline{\mathbb{Q}} = \mathbb{R}$) donc il ne peut pas être complet. ◇

7.4.5 Ensembles compacts

Dans les problèmes de « passage à la limite » une classe d'ensembles joue un rôle important, ce sont les ensembles dits « compacts ».

DÉFINITION 7.23 *Soient (E, d) un espace métrique et K un sous-ensemble de E. L'ensemble K est dit* compact *dans (E, d) si de toute suite d'éléments de K on peut extraire une sous-suite qui converge vers un élément de K.*

Remarques

1. Si un sous-ensemble K est compact pour E muni de la distance d, il est compact pour toute distance équivalente à d sur E puisque toute suite convergeant dans (E, d) converge également dans E muni d'une distance équivalente à d.

2. D'après la définition 7.18, un ensemble K est compact si toute suite d'éléments de K admet au moins une valeur d'adhérence appartenant à K. ◇

Exemples

1. Dans un espace métrique (E, d), tout sous-ensemble K de E possédant un nombre fini d'éléments est compact. En effet, une suite $(x_n)_n$ d'éléments de K a nécessairement un nombre infini de ses termes qui prennent une même valeur a de K. On dispose ainsi d'une suite extraite de la suite $(x_n)_n$ qui est constante. Comme une suite constante converge, on en déduit que la suite $(x_n)_n$ possède une sous-suite qui converge et dont la limite qui est a appartient à K.

2. L'ensemble $\mathbb{R}$ muni de la distance usuelle n'est pas un ensemble compact car il n'existe pas de suite extraite de la suite de terme général $u_n = n$ qui soit convergente.

3. Dans $\mathbb{R}$ muni de la distance usuelle, tout intervalle fermé et borné $[a, b]$, où a, b désignent deux réels tels que $a < b$, est un ensemble compact. Considérons en effet une suite $(u_n)_n$ d'éléments de $[a, b]$; cette suite est bornée (elle est majorée par b et minorée par a). Or, d'après le théorème de Bolzano-Weierstrass[87], de toute suite réelle bornée on peut extraire une sous-suite convergente. Les termes de cette sous-suite sont des éléments de $[a, b]$; comme l'intervalle $[a, b]$ est fermé, d'après le théorème 7.6 sa limite est nécessairement un élément de $[a, b]$. On peut donc extraire de la suite $(u_n)_n$ une sous-suite convergeant vers un élément de $[a, b]$; l'intervalle $[a, b]$ est compact. Cela justifie la terminologie *d'intervalle compact* souvent utilisée pour désigner un intervalle fermé et borné de $\mathbb{R}$.

4. Dans $\mathbb{R}^2$ muni d'une distance équivalente à la distance euclidienne, tout pavé $P = [a_1, b_1] \times [a_2, b_2]$, où a_1, b_1, a_2, b_2 désignent quatre réels tels que $a_1 < b_1$ et $a_2 < b_2$, est compact. Considérons une suite $(u_n)_n$ d'éléments de P ; pour tout $n \in \mathbb{N}$, on a $u_n = (x_n, y_n)$ avec $x_n \in [a_1, b_1]$ et $y_n \in [a_2, b_2]$. La suite $(x_n)_n$ est une suite d'éléments de l'intervalle $[a_1, b_1]$; comme cet intervalle est compact, on peut extraire de la suite $(x_n)_n$ une sous-suite $(x_{h(n)})_n$ (h désignant l'extractrice[88]) qui converge vers un élément ℓ_1 de $[a_1, b_1]$. De manière similaire, la suite $(y_n)_n$ étant une suite d'éléments de l'intervalle compact $[a_1, b_1]$, on peut

[87] Voir le théorème 5.3 p. 189 du *Cours de première année*.

[88] On rappelle, voir la définition 5.7 p. 186 du *Cours de première année* qu'une extractrice est une application de $\mathbb{N}$ dans $\mathbb{N}$ strictement croissante.

extraire de cette suite une sous-suite $(y_{g(n)})_n$ (g désignant l'extractrice) qui converge vers un élément ℓ_2 de $[a_2, b_2]$. En posant $f = g \circ h$, on en déduit que la suite $(v_n)_n$ de terme général $v_n = (x_{f(n)}, y_{f(n)})$ est une suite extraite de la suite $(u_n)_n$ qui converge vers $(\ell_1, \ell_2) \in [a_1, b_1] \times [a_2, b_2]$. L'ensemble P est donc bien un ensemble compact.

Remarque En utilisant un raisonnement analogue à celui développé dans l'exemple précédent, on montre que si K_1 et K_2 sont deux ensembles compacts dans les espaces métriques (E_1, d_1) et (E_2, d_2) respectivement, alors l'ensemble $K_1 \times K_2$ est un ensemble compact dans $E_1 \times E_2$ muni de l'une des distances produit. Par récurrence, on étend ce résultat au produit cartésien de n ensembles compacts. En particulier, dans $\mathbb{R}^n$ muni d'une distance équivalente à la distance euclidienne, tout pavé $\prod_{i=1}^n [a_i, b_i]$ (où pour $i \in \{1, \ldots, n\}$ $(a_i, b_i) \in \mathbb{R}^2$ avec $a_i < b_i$) est un ensemble compact. ◇

PROPOSITION 7.20 *Soient (E, d) un espace métrique et K un sous-ensemble de E. Si K est compact alors K est fermé et borné.*

Démonstration ⊵ Raisonnons par l'absurde et supposons que K n'est pas fermé. Il existe dans ce cas une suite $(x_n)_n$ d'éléments de K qui converge et dont la limite ℓ n'appartient pas à K. Comme K est compact, on peut extraire de cette suite une sous-suite $(x_{h(n)})_n$ qui converge et dont la limite ℓ' est un élément de K. On aboutit à une contradiction : comme la suite $(x_n)_n$ converge vers ℓ, toute suite extraite de cette suite converge obligatoirement vers ℓ ; on a donc $\ell = \ell'$ avec $\ell' \in K$ et $\ell \notin K$.

⊵ Raisonnons par l'absurde et supposons que K n'est pas borné, c'est-à-dire supposons que[89] :

$$\forall \ell \in K \quad \forall C \in \mathbb{R}^+ \quad \exists x \in K \quad d(x, a) > C.$$

En considérant des réels $C = 1/n$ où $n \in \mathbb{N}^*$, cette assertion implique qu'il existe une suite $(x_n)_n$ d'éléments de K tel que

$$\forall \ell \in K \quad \forall n \in \mathbb{N}^* \qquad d(x_n, \ell) > n.$$

Considérons une sous-suite quelconque $(x_{h(n)})_n$ extraite de la $(x_n)_n$, h désignant l'extractrice. D'après l'assertion précédente,

$$\forall \ell \in K \quad \forall N \in \mathbb{N}^* \qquad d(x_{h(N)}, \ell) > h(N)$$

et comme h est une extractrice $h(N) \geqslant N > \frac{1}{2}$. On en déduit que

$$\forall \ell \in K \quad \forall N \in \mathbb{N}^* \qquad d(x_{h(N)}, \ell) > \tfrac{1}{2},$$

[89] Voir la définition 7.11 p. 382.

puis en posant $\varepsilon = \frac{1}{2}$ et $n = N$, que

$$\forall \ell \in K \quad \exists \varepsilon \in \mathbb{R}_+^* \quad \forall N \in \mathbb{N} \quad \exists n \in \mathbb{N} \quad \big(n \leqslant N \text{ et } d(x_{h(n)}, \ell) > \varepsilon\big).$$

Cette assertion exprime le fait que la suite $(x_{h(n)})_n$ diverge dans (K, d), voir p. 389. On obtient une contradiction : puisque K est compact, on devrait pouvoir extraire de la suite $(x_n)_n$ une sous-suite qui converge vers un élément de K. □

Remarque Par contraposition, la proposition 7.20 indique qu'un ensemble qui n'est pas fermé ou qui n'est pas borné, ne peut pas être un compact. ◇

La proposition suivante indique que pour qu'un sous-ensemble d'un ensemble compact soit lui-même compact, il suffit qu'il soit fermé.

PROPOSITION 7.21 *Soient (E, d) un espace métrique et K un sous-ensemble de E qui est compact. Si A est un sous-ensemble fermé de K alors A est compact dans (E, d).*

Démonstration Considérons une suite $(x_n)_n$ d'éléments de A. Pour montrer que A est un ensemble compact, d'après la définition 7.23, il faut montrer que l'on peut extraire de la suite $(x_n)_n$ une sous-suite qui converge dans (E, d) vers un élément ℓ de A. Comme $A \subset K$, les éléments de la suite $(x_n)_n$ appartiennent à K. L'ensemble K étant par hypothèse compact, on peut extraire de la suite $(x_n)_n$ une sous-suite $(x_{h(n)})_n$ qui converge dans (E, d) vers un élément ℓ de l'ensemble K. Comme les termes de la suite extraite $(x_n)_n$ appartiennent à l'ensemble A, les termes de la suite $(x_{h(n)})_n$ appartiennent aussi à l'ensemble A. On peut donc extraire de la suite $(x_n)_n$ une sous-suite $(x_{h(n)})_n$ d'éléments de A qui converge dans (E, d). Par ailleurs, comme A est par hypothèse un ensemble fermé, d'après le théorème 7.6, la limite ℓ de la suite $(x_{h(n)})_n$ est un élément de A. On a donc montré que l'on peut extraire de la suite $(x_n)_n$ une sous-suite $(x_{h(n)})_n$ d'éléments de A qui converge dans (E, d) vers un élément de A. L'ensemble A est donc compact dans (E, d). □

La proposition 7.20 indique qu'un ensemble compact est nécessairement fermé et borné. Le théorème suivant indique que dans $\mathbb{R}^n$, la réciproque est vraie : tout ensemble fermé et borné est compact. Il y a donc coïncidence entre les ensembles compacts et les ensembles fermés et bornés dans $\mathbb{R}^n$.

THÉORÈME 7.7 (Caractérisation des compacts de $\mathbb{R}^n$)
Dans le $\mathbb{R}$-espace vectoriel $\mathbb{R}^n$, un sous-ensemble est compact si et seulement s'il est fermé et borné.

Démonstration D'après la proposition 7.20, tout ensemble compact de $\mathbb{R}^n$ est fermé et borné. Réciproquement, considérons un ensemble A fermé et borné

dans $\mathbb{R}^n$ muni de la norme infinie ; comme A est borné[90], il existe $\mathbf{a} \in A$ et $r \in \mathbb{R}^+$ tel que $A \subset \overline{B}(\mathbf{a}, r)$. En désignant par $(a_1, \ldots, a_n)$ les composantes de $\mathbf{a}$, on a : $\overline{B}(\mathbf{a}, r) = \prod_{i=1}^n [a_i - r, a_i + r]$. Puisque le pavé $P = \prod_{i=1}^n [a_i - r, a_i + r]$ est un ensemble compact de $\mathbb{R}^n$ (voir p. 414) et que A est un sous-ensemble fermé de P, d'après la proposition 7.21, A est compact. □

On montre également que dans le $\mathbb{C}$-espace vectoriel $\mathbb{C}^n$, un sous-ensemble est compact si et seulement s'il est fermé et borné et plus généralement, on a le résultat suivant que nous admettrons.

PROPOSITION 7.22 *Soit $(E, \|\cdot\|)$ un espace vectoriel normé de dimension finie et K un sous-ensemble de E. K est compact si et seulement si K est fermé et borné dans $(E, \|\cdot\|)$.*

7.5 Exercices de synthèse

Exercice 7.16

1 - a) Montrer que l'application $N_2 : f \in \mathcal{C}^1([a,b], \mathbb{R}) \mapsto \left(\int_a^b |f(t)|^2 \, dt\right)^{\frac{1}{2}}$ est une norme sur $\mathcal{C}^1([a,b], \mathbb{R})$.

b) Montrer que l'application $N_2' : f \in \mathcal{C}^1([a,b], \mathbb{R}) \mapsto \left(\int_a^b |f'(t)|^2 \, dt\right)^{\frac{1}{2}}$ est une semi-norme sur $\mathcal{C}^1([a,b], \mathbb{R})$ mais qu'il ne s'agit pas d'une norme.

2 - a) Montrer que $E = \{f \in \mathcal{C}^1([a,b], \mathbb{R}) \ ; \ f(a) = 0\}$ est un espace vectoriel.

b) Montrer que N_2 et N_2' sont deux normes sur E.

3 - Montrer que pour tout $f \in E$ on a

$$N_2(f) \leqslant \frac{b-a}{\sqrt{2}} \, N_2'(f)$$

mais que les deux normes ne sont pas équivalentes (on pourra considérer les fonctions puissances pour l'établir).

Exercice 7.17 *Norme de Holder[91] sur* $\mathbb{R}^n$

1 - a) Pour $a \in \mathbb{R}_+$ fixé, étudier l'application $f : x \in \mathbb{R}^+ \mapsto \frac{1}{3}a^3 + \frac{2}{3}x^{\frac{3}{2}} - ax$.

b) En déduire que pour tous $a, b \in \mathbb{R}^+$ on a $ab \leqslant \frac{1}{3}a^3 + \frac{2}{3}b^{\frac{3}{2}}$.

---

[90] Voir la définition 7.11 p. 382.

[91] HÖLDER, Otto Ludwig (1859, Stuttgart - 1937, Leipzig). Mathématicien allemand qui a notamment travaillé sur la convergence des séries de Fourier et en théorie des groupes.

2 - Pour tout $p \in]1, +\infty[$, on considère l'application $N_p : \mathbf{x} \in \mathbb{R}^n \mapsto (\sum_{k=1}^n |x_k|^p)^{\frac{1}{p}}$ où $(x_1, \ldots, x_n)$ désignent les composantes de $\mathbf{x}$.

a) Soient $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ tels que $N_3(\mathbf{u}) = 1$ et $N_{\frac{3}{2}}(\mathbf{v}) = 1$; montrer que $\sum_{k=1}^n |u_k v_k| \leqslant 1$.

b) En déduire que pour tous $\mathbf{U}, \mathbf{V} \in \mathbb{R}^n$ on a $\sum_{k=1}^n |U_k V_k| \leqslant N_3(\mathbf{U}) N_{\frac{3}{2}}(\mathbf{V})$.

3 - a) Montrer que pour tous $\mathbf{x} = (x_1, \ldots, x_n)$ et $\mathbf{y} = (y_1, \ldots, y_n)$ dans $\mathbb{R}^n$ on a

$$N_3(\mathbf{x} + \mathbf{y})^3 \leqslant \sum_{k=1}^n |x_k| \, |x_k + y_k|^2 + \sum_{k=1}^n |y_k| \, |x_k + y_k|^2.$$

b) En utilisant la question 2, en déduire que $N_3(\mathbf{x} + \mathbf{y}) \leqslant N_3(\mathbf{x}) + N_3(\mathbf{y})$.

4 - Montrer que l'application $N_3 : \mathbf{x} \in \mathbb{R}^n \mapsto \left(\sum_{k=1}^n |x_k|^3\right)^{\frac{1}{3}}$ est une norme sur $\mathbb{R}^n$.

Plus généralement, on montre que pour tout $p \in]1, +\infty[$ l'application $\mathbf{x} \in \mathbb{R}^n \mapsto (\sum_{k=1}^n |x_k|^p)^{\frac{1}{p}}$ est une norme sur $\mathbb{R}^n$.

7.6 Solutions des exercices

Solution de l'exercice 7.1

1 - Pour tout $(x,y) \in \mathbb{R}^2$ on a $\frac{|x-y|}{1+|x-y|} \geqslant 0$. Comme l'application logarithme est strictement croissante sur $]0,+\infty[$ et que $\ln(1) = 0$, on en déduit que l'on a bien $d(x,y) \geqslant 0$. Par ailleurs, puisque $|x-y| = |y-x|$ on peut affirmer que $d(x,y) = d(y,x)$.

2 - Puisque $\ln(z) = 0 \Leftrightarrow z = 1$, on a les équivalences suivantes pour tout $(x,y) \in \mathbb{R}^2$:

$$d(x,y) = 0 \iff \frac{|x-y|}{1+|x-y|} = 0 \iff |x-y| = 0 \iff x = y.$$

3 - Pour montrer que $\forall (x,y,z) \in \mathbb{R}^3 \quad d(x,y) + d(y,z) \leqslant d(x,z)$, on peut montrer que $d(x,y)+d(y,z)-d(x,z) \geqslant 0$ ou encore, l'application exponentielle étant strictement croissante, montrer que $Q = \exp\left(d(x,y) + d(y,z) - d(x,z)\right) \geqslant 1$. On a,

$$\begin{aligned} Q &= \left(1+\frac{|x-y|}{1+|x-y|}\right) \times \left(1+\frac{|y-z|}{1+|y-z|}\right) \times \left(1+\frac{|x-z|}{1+|x-z|}\right)^{-1} \\ &= \frac{(1+2|x-y|)(1+2|y-z|)(1+|x-z|)}{(1+|x-y|)(1+|y-z|)(1+2|x-z|)} \overset{not.}{=} \frac{N}{D}. \end{aligned}$$

Remarquons que le dénominateur D est une quantité strictement positive. Montrer que $Q \geqslant 1$ revient à montrer que $N \geqslant D$. En développant et en simplifiant l'expression $N \geqslant D$, on obtient que cette inégalité est satisfaite si et seulement si

$$\underbrace{-|x-z| + |y-z| + |x-y|}_{=A} + \underbrace{|x-y||y-z|(3+2|z-x|)}_{=B} \geqslant 0.$$

Il est clair que B est une quantité positive. Par ailleurs d'après la première inégalité triangulaire

$$|x-z| = |(x-y)+(y-z)| \leqslant |x-y| + |y-z|,$$

ce qui permet d'affirmer que A est également une quantité positive.

Remarque On peut aussi établir que A est une quantité positive en invoquant le fait que $\delta : (x,y) \in \mathbb{R}^2 \mapsto |x-y|$ est la distance usuelle de $\mathbb{R}$. On a donc $\delta(x,z) = |x-z| \leqslant \delta(x,y) + \delta(y,z) = |x-y| + |y-z|$ d'après la troisième condition intervenant dans la définition d'une distance. ◇

On conclut de ce qui précède que l'application d est bien une distance dans $\mathbb{R}$.

Solution de l'exercice 7.2

1 - Montrons que, compte tenu des hypothèses vérifiées par ϕ, l'application $d : (x,y) \in \mathbb{R}^2 \mapsto \phi(|x-y|)$ satisfait aux quatre conditions de la définition 7.1.

a) Comme ϕ est strictement croissante et que $\phi(0) = 0$, l'application d est bien à valeurs positives.

b) Si $(x, y) \in \mathbb{R}^2$ sont tels que $x = y$ alors on a $d(x, y) = 0$ car $\phi(0) = 0$. Réciproquement, si $d(x, y) = 0$ alors $\phi(|x - y|) = 0$ et comme ϕ est strictement croissante avec $\phi(0) = 0$, on en déduit que $|x - y| = 0$, autrement dit que $x = y$.

c) Pour tout $(x, y) \in \mathbb{R}^2$ on a $|x - y| = |y - x|$ donc $d(x, y) = d(y, x)$.

d) Pour tout $(x, y, z) \in \mathbb{R}^3$ on a $|x - y| \leqslant |x - z| + |y - z|$ et comme ϕ est strictement croissante, on a $\phi(|x - y|) \leqslant \phi(|x - z| + |y - z|)$. Or par hypothèse, pour tout $(u, v) \in [0, +\infty[^2$ on a $\phi(u + v) \leqslant \phi(u) + \phi(v)$, donc $\phi(|x - z| + |y - z|) \leqslant \phi(|x - z|) + \phi(|y - z|)$. On en conclut que :

$$d(x, y) = \phi(|x - y|) \leqslant \phi(|x - z|) + \phi(|y - z|) = d(x, z) + d(z, y).$$

2 - Pour établir que l'application $d : (x, y) \in \mathbb{R}^2 \mapsto \frac{|x-y|}{1+|x-y|}$ est une distance sur $\mathbb{R}$, il suffit de remarquer que pour tout $(x, y) \in \mathbb{R}^2$ on a $d(x, y) = \phi(|x - y|)$ où ϕ est l'application $u \in \mathbb{R}^+ \mapsto \frac{u}{1+u}$ et montrer que cette application ϕ satisfait les conditions de la question 1.

a) L'application ϕ est dérivable sur $\mathbb{R}^+$ car c'est une fraction rationnelle sans pôle dans $\mathbb{R}^+$; pour tout $u \in \mathbb{R}^+$ on a $\phi'(u) = \frac{1}{1+u^2} > 0$ donc ϕ est strictement croissante sur $\mathbb{R}^+$.

b) On a bien $\phi(0) = 0$.

c) Pour montrer que pour tout $(u, v) \in [0, +\infty[^2$ on a $\phi(u + v) \leqslant \phi(u) + \phi(v)$, il suffit d'étudier le signe de $\phi(u + v) - \phi(u) - \phi(v)$. On a

$$\phi(u + v) - \phi(u) - \phi(v) = -\frac{(2 + u + v)uv}{(1 + u)(1 + v)(1 + u + v)}.$$

Comme $u \geqslant 0$ et $v \geqslant 0$, il est clair que $\phi(u + v) - \phi(u) - \phi(v) \leqslant 0$.

3 - Reprenons et généralisons le raisonnement de la question 1. Montrons que compte tenu des hypothèses vérifiées par ϕ et du fait que δ est une distance sur $\mathbb{R}$, l'application $d : (x, y) \in \mathbb{R}^2 \mapsto \phi(\delta(x, y))$ satisfait aux quatre conditions de la définition 7.1.

a) Comme ϕ est strictement croissante et que $\phi(0) = 0$, l'application d est bien à valeurs positives.

b) Si $(x, y) \in \mathbb{R}^2$ sont tels que $x = y$ alors on a $d(x, y) = 0$ car $\phi(0) = 0$. Réciproquement, si $d(x, y) = 0$ alors $\phi(\delta(x, y)) = 0$ et comme ϕ est strictement croissante avec $\phi(0) = 0$, on en déduit que $\delta(x, y) = 0$, autrement dit que $x = y$ car δ est une distance.

c) Pour tout $(x, y) \in \mathbb{R}^2$, puisque δ est une distance on a $\delta(x, y) = \delta(y, x)$ et par conséquent $d(x, y) = \phi(\delta(x, y)) = \phi(\delta(y, x)) = d(y, x)$.

d) Comme δ est une distance, pour tout $(x, y, z) \in \mathbb{R}^3$ on a l'inégalité triangulaire $\delta(x, y) \leqslant \delta(x, z) + \delta(y, z)$ et comme ϕ est strictement croissante, on a $\phi(\delta(x, y)) \leqslant \phi(\delta(x, z) + \delta(y, z))$. Or, par hypothèse, pour tout $(u, v) \in [0, +\infty[^2$ on a $\phi(u + v) \leqslant \phi(u) + \phi(v)$, donc

$$\phi(\delta(x, z) + \delta(y, z)) \leqslant \phi(\delta(x, z)) + \phi(\delta(y, z)).$$

On en conclut que : $d(x,y) \leqslant d(x,z) + d(z,y)$.

Solution de l'exercice 7.3

Remarquons que pour tous f et g appartenant à $\mathcal{C}^0([a,b],\mathbb{R})$, l'application $f : x \in [a,b] \mapsto (f(x) - g(x))^2$ est continue et donc Riemann-intégrable[92] sur $[a,b]$. Par conséquent $d_2(f,g)$ est bien définie. Vérifions que d_2 satisfait aux quatre conditions de la définition 7.1.

1 - Si f et g sont deux applications continues sur $[a,b]$ alors l'application $x \in [a,b] \mapsto (f(x) - g(x))^2$ étant continue et positive sur l'intervalle $[a,b]$, on a[93] d'une part $\int_a^b (f(x) - g(x))^2 \, \mathrm{d}x \geqslant 0$ et d'autre part

$$\int_a^b (f(x) - g(x))^2 \, \mathrm{d}x = 0 \iff \big(\forall x \in [a,b] \quad (f(x) - g(x))^2 = 0\big) \iff f = g.$$

On en déduit que $d_2(f,g) \geqslant 0$ et $(d_2(f,g) = 0 \Leftrightarrow f = g)$.

2 - Soit $(f,g) \in \mathcal{C}^0([a,b],\mathbb{R})^2$; Comme $(f(x) - g(x))^2 = (g(x) - f(x))^2$ pour tout réel $x \in [a,b]$, on en déduit que $d_2(f,g) = d_2(g,f)$.

3 - Pour $(f,g,h) \in \mathcal{C}^0([a,b],\mathbb{R})^3$ on a

$$\begin{aligned} d_2(f,h)^2 &= \int_a^b (f(x) - h(x))^2 \, \mathrm{d}x = \int_a^b \Big((f(x) - g(x)) + (g(x) - h(x))\Big)^2 \, \mathrm{d}x \\ &= \int_a^b (f(x) - g(x))^2 \, \mathrm{d}x + \int_a^b (g(x) - h(x))^2 \, \mathrm{d}x \\ &\qquad + 2 \int_a^b \big(f(x) - g(x)\big) \times \big(g(x) - h(x)\big) \, \mathrm{d}x. \end{aligned}$$

Or, d'après l'inégalité de Cauchy-Schwarz[94], on a

$$\begin{aligned} \int_a^b \big(f(x) - g(x)\big) \times \big(g(x) - h(x)\big) \, \mathrm{d}x &\leqslant \left(\int_a^b (f(x) - g(x))^2 \, \mathrm{d}x\right)^{\frac{1}{2}} \\ &\quad \times \left(\int_a^b (g(x) - h(x))^2 \, \mathrm{d}x\right)^{\frac{1}{2}} = d_2(f,g)\, d_2(g,h). \end{aligned}$$

On en déduit que :

$$d_2(f,h)^2 \leqslant d_2(f,g)^2 + d_2(g,h)^2 + 2\, d_2(f,g)\, d_2(g,h) = (d_2(f,g) + d_2(g,h))^2,$$

autrement dit que : $d_2(f,h) \leqslant d_2(f,g) + d_2(g,h)$.

---

[92] Voir la proposition 18.2 p. 842 du *Cours de première année*.

[93] Voir la proposition 18.5 p. 844 du *Cours de première année*.

[94] Voir la proposition 18.6 p. 846 du *Cours de première année*.

Solution de l'exercice 7.4

1 - La série $\sum_n v_n$ de terme général $v_n = \frac{1}{2^n}\frac{|u_n|}{1+|u_n|}$ est une série à termes positifs. Pour tout $n \in \mathbb{N}$, on a $|u_n| \leqslant |u_n| + 1$ donc $v_n \leqslant \frac{1}{2^n}$. Comme la série géométrique de raison $\frac{1}{2}$ converge, on déduit des critères de comparaison pour les séries à termes positifs (voir le théorème 1.2 p. 14) que la série $\sum_n v_n$ converge.

2 - Vérifions que d satisfait aux quatre conditions de la définition 7.1.

– Soit $((u_n)_n, (v_n)_n) \in \mathcal{S}(\mathbb{K})^2$. On a clairement, compte tenu des propriétés de la valeur absolue dans le cas où $\mathbb{K} = \mathbb{R}$ ou du module dans le cas où $\mathbb{K} = \mathbb{C}$,

$$d((u_n)_n, (v_n)_n) \geqslant 0 \quad \text{et} \quad d((u_n)_n, (v_n)_n) = d((v_n)_n, (u_n)_n).$$

– Par ailleurs $d((u_n)_n, (v_n)_n) = 0$ si et seulement si pour tout $n \in \mathbb{N}$ on a $\frac{1}{2^n}\frac{|u_n - v_n|}{1+|u_n-v_n|} = 0$. Il est évident que si les deux suites $(u_n)_n$ et $(v_n)_n$ sont égales alors pour tout $n \in \mathbb{N}$ on a $\frac{1}{2^n}\frac{|u_n-v_n|}{1+|u_n-v_n|} = 0$. Inversement, Comme $2^n(1+|u_n-v_n|) > 0$, la relation $\frac{1}{2^n}\frac{|u_n-v_n|}{1+|u_n-v_n|} = 0$ pour tout $n \in \mathbb{N}$ ne peut avoir lieu que si $|u_n - v_n| = 0$ *i.e.* si $u_n = v_n$. On en conclut que $d((u_n)_n, (v_n)_n) = 0$ si et seulement si les deux suites $(u_n)_n$ et $(v_n)_n$ sont égales.

– Reste à établir la première inégalité triangulaire pour d. L'application $\phi : x \in \mathbb{R}^+ \mapsto \frac{x}{1+x}$ est continue et dérivable de dérivée l'application $x \in \mathbb{R}^+ \mapsto \frac{1}{(1+x)^2}$ qui est strictement positive. L'application ϕ est donc strictement croissante, ce qui implique que si $(x, y) \in \mathbb{R}^+ \times \mathbb{R}^+$ vérifie $x \leqslant y$ alors

$$\phi(x) = \frac{x}{1+x} \leqslant \phi(y) = \frac{y}{1+y}.$$

Soient $(u_n)_n$, $(v_n)_n$ et $(w_n)_n$ trois suites numériques. Pour tout $n \in \mathbb{N}$, on a : $|u_n - v_n| \leqslant |u_n - w_n| + |w_n - v_n|$. En appliquant la relation précédente avec $x = |u_n - v_n|$ et $y = |u_n - w_n| + |w_n - v_n|$, on obtient :

$$\begin{aligned}\frac{|u_n - v_n|}{1+|u_n - v_n|} &\leqslant \frac{|u_n - w_n| + |w_n - v_n|}{1 + |u_n - w_n| + |w_n - v_n|} \\ &= \frac{|u_n - w_n|}{1 + |u_n - w_n| + |w_n - v_n|} + \frac{|w_n - v_n|}{1 + |u_n - w_n| + |w_n - v_n|}.\end{aligned}$$

Mais

$$\begin{aligned}1 + |u_n - w_n| + |w_n - v_n| &\geqslant 1 + |u_n - w_n|, \\ 1 + |u_n - w_n| + |w_n - v_n| &\geqslant 1 + |w_n - v_n|\end{aligned}$$

donc

$$\frac{|u_n - w_n|}{1 + |u_n - w_n| + |w_n - v_n|} \leqslant \frac{|u_n - w_n|}{1 + |u_n - w_n|}$$

et

$$\frac{|w_n - v_n|}{1 + |u_n - w_n| + |w_n - v_n|} \leqslant \frac{|w_n - v_n|}{1 + |w_n - v_n|}.$$

On en déduit que pour tout $n \in \mathbb{N}$

$$\frac{1}{2^n}\frac{|u_n - v_n|}{1 + |u_n - v_n|} \leqslant \frac{1}{2^n}\frac{|u_n - w_n|}{1 + |u_n - w_n|} + \frac{1}{2^n}\frac{|w_n - v_n|}{1 + |w_n - v_n|}$$

et par conséquent, en sommant sur $n \in \mathbb{N}$, que

$$\begin{aligned} d((u_n)_n,(v_n)_n) &= \sum_{n=0}^{+\infty} \frac{1}{2^n} \frac{|u_n - v_n|}{1+|u_n - v_n|} \\ &\leqslant \sum_{n=0}^{+\infty} \frac{1}{2^n} \frac{|u_n - w_n|}{1+|u_n - w_n|} + \sum_{n=0}^{+\infty} \frac{1}{2^n} \frac{|w_n - v_n|}{1+|w_n - v_n|} \\ &= d((u_n)_n,(w_n)_n) + d((w_n)_n,(v_n)_n). \end{aligned}$$

Solution de l'exercice 7.5

La vérification du fait que d définit bien une distance sur $\mathbb{R}^3$ est très semblable à ce qui a été fait dans le preuve de la proposition 7.2. Montrons l'équivalence des deux distances. On a

$$d^2(x,y) = (x_1-y_1)^2 + (x_2-y_2)^2 + \underbrace{2|x_1-y_1||x_2-y_2|}_{\geqslant 0} + (x_3-y_3)^2,$$

donc $d(x,y) \geqslant d_2(x,y)$. Par ailleurs la relation $2ab \leqslant a^2+b^2$ indique que

$$d^2(x,y) \leqslant 2(x_1-y_1)^2 + 2(x_2-y_2)^2 + (x_3-y_3)^2 \leqslant 2d_2^2(x,y).$$

Les deux distances d et d_2 sont donc équivalentes avec $d_2 \leqslant d \leqslant \sqrt{2}\, d_2$.

Solution de l'exercice 7.6

1 - Montrons que $\widetilde{N} : \mathbf{x} \in E \mapsto aN(\mathbf{x})$ vérifie les 4 conditions de la définition 7.3.

i) Pour tout $\mathbf{x} \in E$ on a $\widetilde{N}(\mathbf{x}) = aN(\mathbf{x}) \geqslant 0$ car $a > 0$ et $N(\mathbf{x}) \geqslant 0$ puisque N est une norme.

ii) Pour tout $\mathbf{x} \in E$ on a $\widetilde{N}(\mathbf{x}) = 0 \Leftrightarrow aN(\mathbf{x}) = 0 \Leftrightarrow N(\mathbf{x}) = 0$ car $a > 0$. Or N étant une norme $N(\mathbf{x}) = 0 \Leftrightarrow \mathbf{x} = 0$.

iii) Pour tout $\mathbf{x} \in E$ et pour tout $\lambda \in \mathbb{K}$ on a $\widetilde{N}(\lambda\mathbf{x}) = aN(\lambda\mathbf{x}) = a|\lambda|N(\mathbf{x})$ car N est une norme. On a donc $\widetilde{N}(\lambda\mathbf{x}) = |\lambda|(aN(\mathbf{x})) = |\lambda|\widetilde{N}(\mathbf{x})$.

iv) Pour tout $(\mathbf{x},\mathbf{y}) \in E^2$ on a $\widetilde{N}(\mathbf{x}+\mathbf{y}) = aN(\mathbf{x}+\mathbf{y}) \leqslant a(N(\mathbf{x}) + N(\mathbf{y}))$ en utilisant la quatrième condition de la définition 7.3 pour N. On en déduit que $\widetilde{N}(\mathbf{x}+\mathbf{y}) \leqslant aN(\mathbf{x}) + aN(\mathbf{y}) = \widetilde{N}(\mathbf{x}) + \widetilde{N}(\mathbf{y})$.

2 - Montrons que $N : \mathbf{x} \in E \mapsto a\,N_1(\mathbf{x}) + b\,N_2(\mathbf{x})$ est une norme sur E.

i) Pour tout $\mathbf{x} \in E$, $N(\mathbf{x}) = a\,N_1(\mathbf{x}) + b\,N_2(\mathbf{x}) \geqslant 0$ car $a > 0, b > 0$ et $N_1(\mathbf{x}) \geqslant 0, N_2(\mathbf{x}) \geqslant 0$ puisque N_1 et N_2 sont deux normes.

ii) La somme de deux réels positifs étant nulle si et seulement si ces deux réels sont nuls, on a $N(\mathbf{x}) = 0$ si et seulement si $a\,N_1(\mathbf{x}) = 0$ et $b\,N_2(\mathbf{x}) = 0$. Comme $a \neq 0$ et $b \neq 0$ ceci équivaut encore à $N_1(\mathbf{x}) = N_2(\mathbf{x}) = 0$. Puisque N_1 et N_2 sont des normes, cette relation équivaut à $\mathbf{x} = \mathbf{0}$.

iii) Pour tout $\mathbf{x} \in E$ et pour tout $\lambda \in \mathbb{K}$ on a

$$N(\lambda \mathbf{x}) = a\, N_1(\lambda \mathbf{x}) + b N_2(\lambda \mathbf{x}) = a\,|\lambda|\, N_1(\mathbf{x}) + b|\lambda|\, N_2(\mathbf{x}) = |\lambda|\, N(\mathbf{x}).$$

iv) Pour tout $(\mathbf{x}, \mathbf{y}) \in E^2$ on a

$$\begin{aligned} N(\mathbf{x}+\mathbf{y}) &= a\, N_1(\mathbf{x}+\mathbf{y}) + b\, N_2(\mathbf{x}+\mathbf{y}) \\ &\leqslant a\, (N_1(\mathbf{x}) + N_2(\mathbf{y})) + b\, (N_2(\mathbf{x}) + N_2(\mathbf{y})) \\ &= \underbrace{a\, N_1(\mathbf{x}) + b\, N_2(\mathbf{x})}_{= N(\mathbf{x})} + \underbrace{a\, N_1(\mathbf{y}) + b\, N_2(\mathbf{y})}_{= N(\mathbf{y})}. \end{aligned}$$

On en déduit que $N(\mathbf{x}+\mathbf{y}) \leqslant N(\mathbf{x}) + N(\mathbf{y})$.

Solution de l'exercice 7.7

1 - Montrons que l'application $N : \mathbf{x} \in \mathbb{K}^n \mapsto \|\mathrm{A}\,\mathbf{x}\|$ est une semi-norme sur $\mathbb{K}^n$. Comme $\|\cdot\|$ est une norme, pour tout $\mathbf{x} \in \mathbb{K}^n$ on a $N(\mathbf{x}) \geqslant 0$. De plus, pour tout $\mathbf{x} \in \mathbb{K}^n$ et pour tout $\lambda \in \mathbb{K}$, on a

$$N(\lambda \cdot \mathbf{x}) = \|\mathrm{A}\,(\lambda \cdot \mathbf{x})\| = \|\lambda \mathrm{A}\,\mathbf{x}\|.$$

Comme $\|\cdot\|$ est une norme, $\|\lambda \mathrm{A}\,\mathbf{x}\| = |\lambda|\ \|\mathrm{A}\,\mathbf{x}\| = |\lambda|\, N(\mathbf{x})$, ce qui permet de conclure que $N(\lambda \cdot \mathbf{x}) = |\lambda|\, N(\mathbf{x})$. Reste à établir l'inégalité triangulaire. Pour tout $(\mathbf{x}, \mathbf{y}) \in E^2$ on a

$$N(\mathbf{x}+\mathbf{y}) = \|\mathrm{A}\,(\mathbf{x}+\mathbf{y})\| = \|\mathrm{A}\,\mathbf{x} + \mathrm{A}\,\mathbf{y}\|.$$

Comme $\|\cdot\|$ est une norme, $\|\mathrm{A}\,\mathbf{x} + \mathrm{A}\,\mathbf{y}\| \leqslant \|\mathrm{A}\,\mathbf{x}\| + \|\mathrm{A}\,\mathbf{y}\| = N(\mathbf{x}) + N(\mathbf{y})$. On a donc $N(\mathbf{x}+\mathbf{y}) \leqslant N(\mathbf{x}) + N(\mathbf{y})$ et on peut conclure que N est une semi-norme sur $\mathbb{K}^n$.

2 - a) L'application N n'est pas une norme puisque, par exemple, en prenant pour A la matrice dont tous les coefficients sont nuls sauf les coefficients diagonaux de position 2 à n qui valent 1, *i.e.* en prenant

$$\mathrm{A} = \begin{pmatrix} 0 & 0 & 0 & & 0 \\ 0 & 1 & 0 & & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & & 0 & 1 & 0 \\ 0 & & 0 & 0 & 1 \end{pmatrix} \quad \text{et} \quad \mathbf{x} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}$$

on obtient $\mathrm{A}\,\mathbf{x} = \mathbf{0}$. On a donc $N(\mathbf{x}) = 0$ mais $\mathbf{x} \neq 0$.

b) Comme $\|\cdot\|$ est une norme, pour tous $\mathrm{A} \in \mathcal{M}_n(\mathbb{K})$ et $\mathbf{x} \in \mathbb{K}^n$ on a

$$N(\mathbf{x}) = 0 \iff \|\mathrm{A}\,\mathbf{x}\| = 0 \iff \mathrm{A}\,\mathbf{x} = \mathbf{0}.$$

L'application N sera une norme si $\mathrm{A}\,\mathbf{x} = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$. Or $\mathrm{A}\,\mathbf{x}$ signifie que $\mathbf{x}$ appartient au noyau de A. Donc pour que l'on ait obligatoirement $\mathbf{x} = \mathbf{0}$, il faut que $\mathrm{Ker}(\mathrm{A})$ soit égal à $\{\mathbf{0}\}$. Cette condition équivaut[95] à ce que A soit une matrice inversible.

En conclusion, l'application $N : \mathbf{x} \in \mathbb{K}^n \mapsto \|\mathrm{A}\,\mathbf{x}\|$ est une norme sur $\mathbb{K}^n$ si et seulement si A est une matrice inversible.

Solution de l'exercice 7.8

1 - a) L'application N_1 est clairement à valeurs positives. Il est par ailleurs immédiat, d'après la définition de l'application N_1, que d'une part $N_1(\mathrm{A}) = 0$ si et seulement si A est la matrice nulle et que d'autre part pour tout $\lambda \in \mathbb{K}$ on a $N_1(\lambda \mathrm{A}) = |\lambda|\ N_1(\mathrm{A})$. Soient $\mathrm{A} = (a_{ij})_{i,j}$ et $\mathrm{B} = (b_{ij})_{i,j}$ deux matrices de $\mathcal{M}_{m,n}(\mathbb{K})$; on a

$$N_1(\mathrm{A}+\mathrm{B}) = \max_{1\leqslant j\leqslant m}\Big(\sum_{i=1}^{n} |a_{ij}+b_{ij}|\Big).$$

Pour tout $(i,j) \in \mathbb{N}^2$ on a $|a_{ij}+b_{ij}| \leqslant |a_{ij}| + |b_{ij}|$ et par conséquent

$$\gamma_j = \sum_{i=1}^{n} |a_{ij}+b_{ij}| \leqslant \sum_{i=1}^{n} (|a_{ij}| + |b_{ij}|) = \underbrace{\sum_{i=1}^{n} |a_{ij}|}_{=\alpha_j} + \underbrace{\sum_{i=1}^{n} |b_{ij}|}_{=\beta_j}.$$

On peut remarquer que : $\max\limits_{1\leqslant j\leqslant m} (\alpha_j+\beta_j) \leqslant \max\limits_{1\leqslant j\leqslant m} \alpha_j + \max\limits_{1\leqslant j\leqslant m} \beta_j$, ce qui entraîne que :

$$\begin{aligned} N_1(\mathrm{A}+\mathrm{B}) &= \max_{1\leqslant j\leqslant m} \sum_{i=1}^{n} |a_{ij}+b_{ij}| \\ &\leqslant \max_{1\leqslant j\leqslant m} \sum_{i=1}^{n} |a_{ij}| + \max_{1\leqslant j\leqslant m} \sum_{i=1}^{n} |b_{ij}| \;=\; N_1(\mathrm{A}) + N_1(\mathrm{B}). \end{aligned}$$

On en déduit que $N_1(\mathrm{A}+\mathrm{B}) \leqslant N_1(\mathrm{A}) + N_1(\mathrm{B})$. On a donc établi que l'application N_1 est une norme sur $\mathcal{M}_n(\mathbb{K})$.

b) Pour montrer que N_∞ est une norme sur $\mathcal{M}_n(\mathbb{K})$, on peut remarquer que pour tout $\mathrm{A} \in \mathcal{M}_n(\mathbb{K})$ on a $N_\infty(\mathrm{A}) = N_1(\mathrm{A}^{\mathrm{T}})$. On en déduit que N_∞ est bien à valeurs positives et que $N_\infty(\mathrm{A}) = 0$ si et seulement si A^T est la matrice nulle autrement dit si et seulement $\mathrm{A} = 0$. Par ailleurs, pour tout $\lambda \in \mathbb{K}$ on a $N_\infty(\lambda \mathrm{A}) = N_1(\lambda \mathrm{A}^T) = |\lambda|\ N_1(\mathrm{A}^T) = |\lambda|\ N_\infty(\mathrm{A})$. Enfin, pour toutes matrices A et B de $\mathcal{M}_{m,n}(\mathbb{K})$ on a

$$\begin{aligned} N_\infty(\mathrm{A}+\mathrm{B}) &= N_1((\mathrm{A}+\mathrm{B})^T) \;=\; N_1(\mathrm{A}^T + \mathrm{B}^T) \\ &\leqslant N_1(\mathrm{A}^T) + N_1(\mathrm{B}^T) \;=\; N_\infty(\mathrm{A}) + N_\infty(\mathrm{B}). \end{aligned}$$

[95] Voir le *Cours de première année* p. 425 et suivantes.

2 - Soient $A = (a_{ij})_{1\leqslant i\leqslant n, 1\leqslant j\leqslant m}$ et $B = (b_{ij})_{1\leqslant i\leqslant n, 1\leqslant j\leqslant m}$ deux matrices de $\mathcal{M}_{m,n}(\mathbb{K})$ et $C = A \times B$. Les coefficients c_{ij}, $(i,j) \in \{1,\dots,n\}^2$ de la matrice C sont données par[96] : $c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$. On a $N_s(A \times B)^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} |c_{ij}|^2$ et d'après l'inégalité de Cauchy-Schwarz[97],

$$|c_{ij}|^2 = \left|\sum_{k=1}^{n} a_{ik}b_{kj}\right|^2 \leqslant \left(\sum_{k=1}^{n} |a_{ik}|^2\right) \times \left(\sum_{k=1}^{n} |b_{kj}|^2\right).$$

On en déduit que

$$\begin{aligned} N_s(A \times B)^2 &\leqslant \sum_{i=1}^{n}\sum_{j=1}^{n} \underbrace{\left(\sum_{k=1}^{n} |a_{ik}|^2\right)}_{\text{indépendant de } j} \left(\sum_{k=1}^{n} |b_{kj}|^2\right) \\ &= \left(\sum_{i=1}^{n}\sum_{k=1}^{n} |a_{ik}|^2\right) \left(\sum_{j=1}^{n}\sum_{k=1}^{n} |b_{kj}|^2\right) = N_s(A)^2\, N_s(B)^2. \end{aligned}$$

Solution de l'exercice 7.9

1 - Soit x un élément quelconque de A ; pour tout $r \in \mathbb{R}_+^*$ le point x appartient à $B(x,r)$ et il appartient A ; on a donc $x \in \overline{A}$. On en déduit que $A \subset \overline{A}$. Il n'y a toutefois pas égalité de ces ensembles car par exemple dans $\mathbb{R}$ muni de la distance usuelle, l'ensemble $A = [0,1[$ a pour adhérence $[0,1]$ donc $A \neq \overline{A}$.

2 - a) Montrons que si A est fermé alors $\overline{A} \subset A$ ou compte tenu de la première question que $\overline{A} = A$. Supposons que A est fermé, autrement dit supposons que $\complement_E A$ est ouvert. Pour montrer que $\overline{A} \subset A$, raisonnons par l'absurde et supposons que que cette inclusion n'a pas lieu : il existe donc un élément $x_0 \in \overline{A}$ qui n'appartient pas à A (autrement dit qui appartient à $\complement_E A$). D'après la définition 7.9, puisque $x_0 \in \overline{A}$

$$\forall r \in \mathbb{R}_+^* \quad \exists y \in B(x_0,r) \cap A. \tag{9}$$

Comme par hypothèse l'ensemble $\complement_E A$ est ouvert, il existe $r' \in \mathbb{R}_+^*$ tel que $B(x_0,r') \subset \complement_E A$. Cela constitue une contradiction avec l'assertion (9) : il n'existe pas d'élément y de l'ensemble A dans la boule $B(x_0,r')$. On en conclut que si A est fermé alors $\overline{A} \subset A$.

---

[96] Voir la définition 10.8 p. 398 du *Cours de première année*.

[97] Voir la proposition 3.6 p. 101 du *Cours de première année* pour le cas où $\mathbb{K} = \mathbb{R}$; on établit de la même manière l'inégalité de Cauchy-Schwarz dans le cas où $\mathbb{K} = \mathbb{C}$.

b) Réciproquement, montrons que si $\overline{A} = A$ alors A est fermé (autrement dit $\complement_E A$ est ouvert), ou, de manière équivalente[98], montrons que si $\complement_E A$ n'est pas ouvert alors $\overline{A} \neq A$. On a les équivalences suivantes :

$$\begin{aligned}
\complement_E A \text{ n'est pas ouvert} &\iff \text{non } \big(\forall x \in \complement_E A \quad \exists r \in \mathbb{R}_+^* \quad B(x,r) \subset \complement_E A\big) \\
&\iff \text{non } \big(\forall x \in \complement_E A \quad \exists r \in \mathbb{R}_+^* \quad \forall y \in B(x,r) \quad y \in \complement_E A\big) \\
&\iff \exists x \in \complement_E A \quad \forall r \in \mathbb{R}_+^* \quad \exists y \in B(x,r) \quad y \notin \complement_E A \\
&\iff \exists x \in \complement_E A \quad \forall r \in \mathbb{R}_+^* \quad \exists y \in B(x,r) \quad y \in A \\
&\iff \exists x \in \complement_E A \quad x \in \overline{A}.
\end{aligned}$$

Ainsi si $\complement_E A$ n'est pas ouvert alors il existe un élément x appartenant à $\overline{A}$ mais pas à A ; cela signifie que $\overline{A} \neq A$. On en conclut que si $\overline{A} = A$ alors A est fermé. On a donc établi l'équivalence,

$$\boxed{A \text{ est fermé si et seulement si } A = \overline{A}.}$$

Solution de l'exercice 7.10

Montrons que si A et B sont denses dans E et si A est ouvert alors $A \cap B$ est dense dans E. D'après la définition 7.10, il s'agit de montrer que

$$\forall x \in E \quad \forall r \in \mathbb{R}_+^* \quad \exists y \in (A \cap B) \quad d(x,y) < r, \tag{10}$$

Soit x un élément de E et r un réel strictement positif ; comme A est dense dans E, il existe $y \in A$ tel que $d(x,y) < r$, *i.e.* tel que $y \in A \cap B(x,r)$. Puisque A et $B(x,r)$ sont des ouverts, $A \cap B(x,r)$ est un voisinage de y (un ensemble ouvert est voisinage de chacun de ses points) et par conséquent, il existe un réel strictement positif r' tel que $B(y,r') \subset A \cap B(x,r)$. Or B étant dense dans E, il existe $y' \in B$ tel que $y' \in B(y,r')$. Cet élément y' appartient à B et à $B(y,r')$, donc à B et à A et à $B(x,r)$. On a par conséquent $y' \in A \cap B$ et $d(x,y') < r$. L'assertion (10) est démontrée et on peut conclure que si A et B sont denses dans E et si A est ouvert alors $A \cap B$ est dense dans E.

Solution de l'exercice 7.11

1 - a) Supposons que la suite $(f_n)_n$ converge uniformément vers $f \in \mathcal{C}^0([a,b],\mathbb{R})$, autrement dit qu'elle converge vers f pour la distance

$$d_\infty : (f,g) \in \mathcal{C}^0([a,b],\mathbb{R}) \times \mathcal{C}^0([a,b],\mathbb{R}) \longmapsto \sup_{x\in[a,b]} |f(x) - g(x)|$$

et montrons qu'elle converge vers f pour la distance d_2. Pour tout $n \in \mathbb{N}$, on a

$$\begin{aligned}
d_2(f_n,f)^2 = \int_a^b \big(f_n(x) - f(x)\big)^2 \,\mathrm{d}x \;&\leqslant\; \int_a^b \Big(\sup_{y\in[a,b]} |f_n(y) - f(y)|\Big)^2 \,\mathrm{d}x \\
&= (b-a)\, d_\infty(f_n,f)^2.
\end{aligned}$$

[98] Il s'agit de la contraposée de l'assertion à démontrer, voir la proposition 1.2 p. 10 du *Cours de première année*.

On en déduit que pour tout $n \in \mathbb{N}$, on a0 $\leqslant d_2(f_n, f) \leqslant \sqrt{b-a}\; d_\infty(f_n, f)$. Comme la suite $(f_n)_n$ converge uniformément vers $f \in \mathcal{C}^0([a,b], \mathbb{R})$, on a $\lim_{n\to+\infty} d_\infty(f_n, f) = 0$ et le théorème d'encadrement permet de conclure que $\lim_{n\to+\infty} d_2(f_n, f) = 0$, autrement dit que la suite $(f_n)_n$ converge vers f dans $(\mathcal{C}^0([a,b], \mathbb{R}), d_2)$.

La réciproque est fausse. Considérons la suite de fonctions $(\phi_n)_n$ de terme général

$$\phi_n : x \in [0,1] \longmapsto \begin{cases} -n^{\frac{5}{2}}x + \sqrt{n} & \text{si } x \in [0, 1/n^2] \\ 0 & \text{si } x \in]1/n^2, 1] \end{cases}.$$

Cette suite converge dans $(\mathcal{C}^0([a,b], \mathbb{R}), d_2)$ vers la fonction nulle car

$$\begin{aligned} d_2(\phi_n, 0)^2 &= \int_0^1 \phi_n^2(x)\, \mathrm{d}x = \int_0^{1/n^2} \phi_n^2(x)\, \mathrm{d}x = \int_0^{1/n^2} \left(-n^{\frac{5}{2}}x + \sqrt{n}\right)^2 \mathrm{d}x \\ &= \left[-\frac{1}{3n^{\frac{5}{2}}}\left(-n^{\frac{5}{2}}x + \sqrt{n}\right)^3\right]_0^{1/n^2} = \frac{1}{3n}, \end{aligned}$$

donc $\lim\limits_{n\to+\infty} d_2(\phi_n, 0) = 0$. Par contre, on a $d_\infty(\phi_n, \phi) = \sup_{x\in[0,1]} |\phi_n(x)| = \sqrt{n}$ et cette quantité ne tend pas vers 0 quand n tend vers $+\infty$. La suite $(\phi_n)_n$ ne converge donc pas vers la fonction nulle dans $(\mathcal{C}^0([a,b], \mathbb{R}), d_\infty)$.

b) On a vu que la suite $(\phi_n)_n$ converge vers la fonction nulle dans $\mathcal{C}^0([0,1], \mathbb{R})$ pour la distance d_2. Elle converge simplement vers la fonction nulle sur $]0,1]$ mais ne converge pas simplement en 0 puisque pour tout $n \in \mathbb{N}$ on a $\phi_n(0) = \sqrt{n}$ et la suite numérique réelle $(\sqrt{n})_n$ diverge.

2 - Supposons que la suite $(f_n)_n$ converge vers f dans $(\mathcal{C}^0([0,1], \mathbb{R}), d_2)$; on a

$$\lim_{n\to+\infty} d_2(f_n, f) = \lim_{n\to+\infty} \left(\int_a^b \left(f_n(x) - f(x)\right)^2 \mathrm{d}x\right)^{1/2} = 0.$$

En utilisant l'inégalité de Cauchy-Schwarz[(99)], on obtient pour tout $n \in \mathbb{N}$,

$$\begin{aligned} 0 \leqslant \left(\int_a^b |f_n(x) - f(x)|\, \mathrm{d}x\right)^2 &\leqslant \int_a^b 1\, \mathrm{d}x \times \int_a^b \left(f_n(x) - f(x)\right)^2 \mathrm{d}x \\ &= (b-a)\; d_2^2(f_n, f). \end{aligned}$$

Le théorème d'encadrement permet de déduire de cette inégalité que $\lim_{n\to+\infty} d_1(f_n, f)$, autrement dit que la suite $(f_n)_n$ converge vers f dans $(\mathcal{C}^0([0,1], \mathbb{R}), d_1)$.

La réciproque est fausse. Considérons la suite de fonctions $(\psi_n)_n$ de terme général

$$\psi_n : x \in [0,1] \longmapsto \begin{cases} -n^{\frac{3}{2}}x + \sqrt{n} & \text{si } x \in [0, 1/n] \\ 0 & \text{si } x \in]1/n, 1] \end{cases}.$$

(99) Voir la proposition 18.6 p. 846 du *Cours de première année*.

Cette suite converge dans $(\mathcal{C}^0([a,b],\mathbb{R}),d_1)$ vers la fonction nulle car

$$d_1(\psi_n,0)=\int_0^1|\psi_n(x)|\,\mathrm{d}x=\frac{1}{2\sqrt{n}},$$

donc $\lim_{n\to+\infty}d_1(\psi_n,0)=0$. Par contre, on a

$$\begin{aligned}d_2(\psi_n,0)^2&=\int_0^1\psi_n^2(x)\,\mathrm{d}x=\int_0^{1/n}\psi_n^2(x)\,\mathrm{d}x=\int_0^{1/n}\left(-n^{\frac{3}{2}}x+\sqrt{n}\right)^2\mathrm{d}x\\&=\Big[-\frac{1}{3n^{\frac{3}{2}}}\left(-n^{\frac{3}{2}}x+\sqrt{n}\right)^3\Big]_0^{1/n}=\frac{1}{3},\end{aligned}$$

donc $\lim_{n\to+\infty}d_2(\psi_n,\psi)\neq 0$, ce qui signifie que la suite $(\psi_n)_n$ ne converge pas vers la fonction nulle dans $(\mathcal{C}^0([0,1],\mathbb{R}),d_2)$.

Solution de l'exercice 7.12

1 - Pour montrer que la suite $(y_n)_n$ est de Cauchy dans (E,d), il s'agit de d'établir l'assertion suivante :

$$\forall\varepsilon\in\mathbb{R}_+^*\quad\exists N\in\mathbb{N}\quad\forall(m,n)\in\mathbb{N}^2\quad(n\geqslant m\geqslant N\implies d(y_n,y_m)\leqslant\varepsilon).\quad(11)$$

Soit ε un réel strictement positif et m, n deux entiers naturels ; on a en vertu de l'inégalité triangulaire :

$$d(y_n,y_m)\leqslant d(y_n,x_n)+d(x_n,x_m)+d(x_m,y_m)\leqslant\tfrac{1}{n}+\tfrac{1}{m}+d(x_m,y_m).\quad(12)$$

Par hypothèse, la suite $(x_n)_n$ est de Cauchy dans (E,d) donc

$$\exists N_1\in\mathbb{N}\quad\forall(m,n)\in\mathbb{N}^2\quad(n\geqslant m\geqslant N_1\implies d(x_n,x_m)\leqslant\tfrac{1}{3}\varepsilon).$$

Par ailleurs, la suite $(x_n)_n$ de terme général $x_n=1/n$ converge dans $\mathbb{R}$ muni de la distance euclidienne vers 0, donc

$$\exists N_2\in\mathbb{N}\quad\forall k\in\mathbb{N}\quad(k\geqslant N_2\implies\tfrac{1}{k}\leqslant\tfrac{1}{3}\varepsilon).$$

On en déduit que si m et n sont choisis tels que $n\geqslant m\geqslant N=\max(N_1,N_2)$ alors d'après (12) on a $d(y_n,y_m)\leqslant 3\,\frac{1}{3}\varepsilon=\varepsilon$. On a donc établi l'assertion (11).

2 - Compte tenu du rôle identique joué par les suites $(x_n)_n$ et $(y_n)_n$ (elles sont toutes les deux des suites de Cauchy et l'hypothèse « $y_0=x_0$ et pour tout $n\in\mathbb{N}^*$ on ait $d(x_n,y_n)\leqslant 1/n$ » donne le même rôle à $(x_n)_n$ et $(y_n)_n$), pour montrer que la suite $(x_n)_n$ converge vers ℓ dans (E,d) si et seulement si la suite $(y_n)_n$ converge vers ℓ dans (E,d), il suffit de montrer que si la suite $(x_n)_n$ converge vers ℓ dans (E,d) alors la suite $(y_n)_n$ converge vers ℓ dans (E,d).

Pour tout $n\in\mathbb{N}$, compte tenu de l'inégalité triangulaire, on obtient

$$0\leqslant d(y_n,\ell)\leqslant d(y_n,x_n)+d(x_n,\ell)\leqslant d(x_n,\ell)+1/n.$$

Si on suppose que la suite $(x_n)_n$ converge vers ℓ dans (E,d), alors en vertu du théorème d'encadrement, on obtient $\lim_{n\to+\infty} d(y_n,\ell)=0$. Cela signifie que la suite $(y_n)_n$ converge vers ℓ dans (E,d).

SOLUTION DE L'EXERCICE 7.13

1 - Considérons une suite $(x_n)_n$ de Cauchy dans (E,d) et montrons que cette suite est bornée, *i.e.* montrons que[100]

$$\exists e\in E \quad \exists K\in\mathbb{R}^+ \quad \forall n\in\mathbb{N} \quad d(x_n,e)\leqslant K.$$

D'après la définition 7.20 (on prend $\varepsilon=1$),

$$\exists N\in\mathbb{N} \quad \forall n\in\mathbb{N} \quad \forall m\in\mathbb{N} \quad \big((n\geqslant N \text{ et } m\geqslant N)\implies d(x_n,x_m)\leqslant 1\big).$$

On en déduit en prenant $m=N$ que pour tout $n\in\mathbb{N}$ avec $n\geqslant N$,

$$d(x_n,x_0)\leqslant d(x_0,x_N)+d(x_N,x_n)\leqslant d(x_0,x_N)+1.$$

Par ailleurs, pour $n\in\mathbb{N}$ avec $n<N$ on a

$$d(x_n,x_0)\leqslant M \quad \text{où} \quad M=\max_{k=1,\dots,N-1} d(x_k,x_0).$$

En posant $e=x_0$ et $K=\max(d(x_0,x_N)+1,M)$, on obtient que pour tout $n\in\mathbb{N}$, $d(x_n,x_0)\leqslant K$, ce qui signifie que la suite $(x_n)_n$ est bornée.

2 - Soit $(x_n)_n$ une suite de Cauchy dans (E,d) possédant une sous-suite convergente. D'après la définition 7.20,

$$\begin{aligned}\forall\varepsilon_1\in\mathbb{R}_+^* \quad \exists N_1\in\mathbb{N} \quad \forall n\in\mathbb{N} \quad \forall m\in\mathbb{N}\\ \big((n\geqslant N_1 \text{ et } m\geqslant N_1)\implies d(x_n,x_m)\leqslant\varepsilon_1\big).\end{aligned} \tag{13}$$

Par hypothèse, la suite $(x_n)_n$ possède une sous-suite $(x_{h(n)})_n$ qui converge. En désignant par $\ell\in E$ la limite de cette sous-suite, on a d'après la définition 7.16,

$$\forall\varepsilon_2\in\mathbb{R}_+^* \quad \exists N_2\in\mathbb{N} \quad \forall n\in\mathbb{N} \quad \big(n\geqslant N_2\implies d(x_{h(n)},\ell)\leqslant\varepsilon_2\big). \tag{14}$$

Considérons un réel strictement positif ε fixé et prenons $\varepsilon_1=\varepsilon/2$ dans (13) et $\varepsilon_2=\varepsilon/2$ dans (14). Pour tout $n\in\mathbb{N}$ avec $n\geqslant N$ où $N=\max\{N_1,N_2\}$, on a $h(n)\geqslant n\geqslant N$ car h est une extractrice. Par ailleurs, en utilisant la première inégalité triangulaire et les assertions (13) et (14), on obtient :

$$d(x_n,\ell)\leqslant \underbrace{d(x_n,x_{h(n)})}_{\leqslant\varepsilon/2}+\underbrace{d(x_{h(n)},\ell)}_{\leqslant\varepsilon/2}\leqslant\varepsilon.$$

On a donc montré que

$$\forall\varepsilon\in\mathbb{R}_+^* \quad \exists N\in\mathbb{N} \quad \forall n\in\mathbb{N} \quad \big(n\geqslant N\implies d(x_n,\ell)\leqslant\varepsilon\big)$$

---

[100] Voir la définition 7.17 p. 392.

autrement dit, que la suite $(x_n)_n$ converge vers ℓ dans (E,d).

Solution de l'exercice 7.14

Commençons par montrer que la suite $(f_n)_n$ est une suite de Cauchy dans $(\mathcal{C}^0([0,1],\mathbb{R}),\|\ \|_2)$. Soient ε un réel strictement positif fixé et m,n deux entiers naturels tels que $n \geqslant m$. On a

$$\begin{aligned} d_2(f_n,f_m)^2 &= \int_0^1 (f_n(x)-f_m(x))^2\,\mathrm{d}x = \int_{\frac12}^{\frac12+\frac1m} (f_n(x)-f_m(x))^2\,\mathrm{d}x \\ &= \int_{\frac12}^{\frac12+\frac1m} f_m(x)^2\,\mathrm{d}x + \int_{\frac12}^{\frac12+\frac1n} f_n(x)^2\,\mathrm{d}x - 2\int_{\frac12}^{\frac12+\frac1n} f_m(x)\,f_n(x)\,\mathrm{d}x \\ &= m^2\int_{\frac12}^{\frac12+\frac1m}\left(x-\frac12+\frac1m\right)^2\mathrm{d}x + n^2\int_{\frac12}^{\frac12+\frac1n}\left(x-\frac12+\frac1n\right)^2\mathrm{d}x \\ &\quad - 2mn\int_{\frac12}^{\frac12+\frac1n}\left(x-\frac12+\frac1m\right)\left(x-\frac12+\frac1n\right)\mathrm{d}x \\ &= \frac{7}{3n}+\frac{7}{3m}-\frac{5m+9n}{6n^2} \leqslant \frac{14}{3m}. \end{aligned}$$

On en déduit que $d_2(f_n,f_m)\leqslant\varepsilon$ pour tout $(m,n)\in\mathbb{N}^2$ avec $n\geqslant m\geqslant N$ où $N=\mathrm{E}(\frac{14}{3\epsilon^2})+1$.

La suite $(f_n)_n$ converge simplement sur $[0,1]$ vers l'application f définie par $f(x)=1$ si $x\in[0,\frac12]$ et $f(x)=0$ si $x\in]\frac12,1]$. L'application f n'est pas continue sur $[0,1]$; toutefois, elle est Riemann-intégrable sur $[0,1]$ car elle est continue par morceaux. Pour tout $n\in\mathbb{N}$, on a :

$$\|f_n-f\|_2^2 = \int_0^1 (f_n(t)-f(t))^2\,\mathrm{d}t = n^2\int_{\frac12}^{\frac12+\frac1n}\left(t-\frac12+\frac1n\right)^2\mathrm{d}t = \frac{7}{3n}$$

ce qui implique que $\lim_{n\to+\infty}\|f_n-f\|_2=0$. On en déduit que la suite $(f_n)_n$ converge dans $(\mathcal{RI}([0,1],\mathbb{R}),\|\cdot\|_2)$ vers f. Comme $\mathcal{C}^0([0,1],\mathbb{R})$ est un sous-espace vectoriel de $\mathcal{RI}([0,1])$, si $(f_n)_n$ convergeait dans $(\mathcal{C}^0([0,1],\mathbb{R}),\|\cdot\|_2)$ ce serait nécessairement vers f. Puisque f n'est pas continue sur $[0,1]$, on en déduit que la suite $(f_n)_n$ ne converge pas dans $(\mathcal{C}^0([0,1],\mathbb{R}),\|\cdot\|_2)$ et par conséquent que $(\mathcal{C}^0([0,1],\mathbb{R}),\|\cdot\|_2)$ n'est pas un espace de Banach.

Solution de l'exercice 7.15

1 - Remarquons que l'on a $N(f)=\|f\|_\infty+\|f'\|_\infty$ pour tout $f\in E$. On en déduit que $N(f)\geqslant 0$ pour tout $f\in E$. Par ailleurs, si $N(f)=0$ alors $\|f\|_\infty=0$ et $\|f'\|_\infty=0$. Comme $\|\cdot\|_\infty$ est une norme sur $\mathcal{C}^1([0,1],\mathbb{R})$, il en résulte que $f=0$. De plus, le fait que $\|\cdot\|_\infty$ soit une norme implique que pour tout $(f,g)\in E^2$ et pour tout $\lambda\in\mathbb{R}$, on a

$$\begin{aligned} N(\lambda f) &= \|\lambda f\|_\infty+\|(\lambda f)'\|_\infty = |\lambda|\|f\|_\infty+|\lambda|\|f'\|_\infty = |\lambda|N(f); \\ N(f+g) &= \|f+g\|_\infty+\|(f+g)'\|_\infty = \|f+g\|_\infty+\|f'+g'\|_\infty \\ &\leqslant \|f\|_\infty+\|g\|_\infty+\|f'\|_\infty+\|g'\|_\infty = N(f)+N(g). \end{aligned}$$

On en conclut que N est bien une norme sur $E = \{f \in \mathcal{C}^1([0,1],\mathbb{R}) \mid f(0) = 0\}$.

2 - a) Considérons une suite de Cauchy $(f_n)_n$ dans (E,N) ; on a :

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N \in \mathbb{N} \quad \forall (m,n) \in \mathbb{N}^2 \quad \big(m \geqslant n \geqslant N \Longrightarrow N(f_n - f_m) \leqslant \varepsilon\big),$$

ce qui implique que

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N \in \mathbb{N} \quad \forall (m,n) \in \mathbb{N}^2 \quad \big(m \geqslant n \geqslant N \Longrightarrow \sup_{x\in[0,1]} |f_n(x) - f_m(x)| \leqslant \varepsilon\big) \tag{15}$$

et que

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N \in \mathbb{N} \quad \forall (m,n) \in \mathbb{N}^2 \quad \big(m \geqslant n \geqslant N \Longrightarrow \sup_{x\in[0,1]} |f_n'(x) - f_m'(x)| \leqslant \varepsilon\big). \tag{16}$$

On déduit de l'assertion (15) que la suite $(f_n)_n$ est une suite de Cauchy dans $(\mathcal{C}^0([0,1],\mathbb{R}), \|\cdot\|_\infty)$ qui est complet donc que la suite $(f_n)_n$ converge vers un élément $f \in \mathcal{C}^0([0,1],\mathbb{R})$ pour la norme $\|\cdot\|_\infty$.

b) On déduit par ailleurs de l'assertion (16) que la suite $(f_n')_n$ est une suite de Cauchy dans $(\mathcal{C}^0([0,1],\mathbb{R}), \|\cdot\|_\infty)$ qui est complet donc que la suite $(f_n')_n$ converge vers un élément $g \in \mathcal{C}^0([0,1],\mathbb{R})$ pour la norme $\|\cdot\|_\infty$.

Pour montrer que (E,N) est un espace vectoriel normé complet, il faut montrer que la suite $(f_n)_n$ converge pour la norme N vers un élément de E. Si on vérifie que $f' = g$ et que $f(0) = 0$ alors on pourra d'une part conclure à l'appartenance de f à E et d'autre part affirmer que la suite $(f_n)_n$ converge pour la norme N car on a

$$N(f_n - f) = \|f_n - f\|_\infty + \|f_n' - f'\|_\infty.$$

Tout d'abord, la suite $(f_n)_n$ converge uniformément vers f sur $[0,1]$, donc en particulier la suite converge simplement en 0 vers $f(0)$. Puisque pour tout $n \in \mathbb{N}$ on a $f_n(0) = 0$, on en déduit que $f(0) = 0$. Par ailleurs, les fonctions f_n sont de classe $\mathcal{C}^1$ sur $[0,1]$, la suite $(f_n)_n$ converge simplement en 0 et la suite $(f_n')_n$ converge uniformément sur $[0,1]$. Il résulte du théorème 2.5 p. 92 que la fonction f est de classe $\mathcal{C}^1$ sur $[0,1]$ et que $g = f'$.

Solution de l'exercice 7.16

1 - a) Montrons que l'application $N_2 : f \in \mathcal{C}^1([a,b],\mathbb{R}) \mapsto \left(\int_a^b |f(t)|^2\,dt\right)^{\frac{1}{2}}$ est une norme sur $\mathcal{C}^1([a,b],\mathbb{R})$. Cette application est clairement à valeurs positives.

i) Si f est l'application nulle sur $[a,b]$, on a $N_2(f) = \left(\int_a^b 0^2\,dt\right)^{\frac{1}{2}} = 0$. Inversement, supposons que $f \in \mathcal{C}^1([a,b],\mathbb{R})$ soit telle que $N_2(f) = 0$; on a alors $\int_a^b |f(t)|^2\ \mathrm{d}t = 0$. Comme f est continue sur $[a,b]$, l'application $g : x \in [a,b] \mapsto |f(x)|^2$ est continue et à valeurs positives; la relation $\int_a^b |f(t)|^2\ \mathrm{d}t = 0$ implique[101] alors que g est l'application nulle. On en déduit que f est elle-même l'application nulle.

ii) Il résulte des propriétés de l'intégrale de Riemann que pour tout $f \in \mathcal{C}^1([a,b],\mathbb{R})$ et pour tout $\lambda \in \mathbb{R}$ on a,

$$N_2(\lambda f) = \left(\int_a^b |\lambda f(t)|^2\ \mathrm{d}t\right)^{\frac{1}{2}} = \left(\lambda^2 \int_a^b |f(t)|^2\ \mathrm{d}t\right)^{\frac{1}{2}} = |\lambda|\, N_2(f).$$

iii) Pour tout $(f,g) \in \mathcal{C}^1([a,b],\mathbb{R}) \times \mathcal{C}^1([a,b],\mathbb{R})$ on a

$$N_2(f+g) = \left(\int_a^b |f(t)+g(t)|^2\ \mathrm{d}t\right)^{\frac{1}{2}} \leqslant \left(\int_a^b |f(t)|^2\ \mathrm{d}t + \int_a^b |g(t)|^2\ \mathrm{d}t\right)^{\frac{1}{2}}$$

En utilisant la relation $\sqrt{\alpha+\beta} \leqslant \sqrt{\alpha} + \sqrt{\beta}$ valable pour tout $(\alpha,\beta) \in \mathbb{R}_+ \times \mathbb{R}_+$, on en déduit que

$$N_2(f+g) \leqslant \left(\int_a^b |f(t)|^2\ \mathrm{d}t\right)^{\frac{1}{2}} + \left(\int_a^b |g(t)|^2\ \mathrm{d}t\right)^{\frac{1}{2}} = N_2(f) + N_2(g).$$

b) Comme N_2 est une norme sur $\mathcal{C}^1([a,b],\mathbb{R})$ et que E est un sous-espace vectoriel de $\mathcal{C}^1([a,b],\mathbb{R})$, N_2 est aussi une norme sur E. On peut remarquer qu'en ce qui concerne l'application N_2', les conditions ii) et iii) ont une justification en tout point analogue. Intéressons-nous à la condition i).

Si f est l'application nulle sur $[a,b]$, on a $N_2'(f) = \left(\int_a^b 0^2\,dt\right)^{\frac{1}{2}} = 0$. Inversement, supposons que $f \in \mathcal{C}^1([a,b],\mathbb{R})$ soit telle que $N_2'(f) = 0$; on a alors $\int_a^b |f'(t)|^2\ \mathrm{d}t = 0$. Comme f' est continue sur $[a,b]$, l'application $g : x \in [a,b] \mapsto |f'(x)|^2$ est continue et à valeurs positives; la relation $\int_a^b g(t)\ \mathrm{d}t = \int_a^b |f'(t)|^2\ \mathrm{d}t = 0$ implique[101] alors que g est l'application nulle. On en déduit que f' est elle-même l'application nulle ce qui permet d'affirmer que f est une application constante sur $[a,b]$, mais pas nécessairement l'application nulle.

[101] Voir la proposition 18.5 p. 844 du *Cours de première année*.

2 - a) Pour montrer que E est un espace vectoriel, il suffit[102] d'établir que E est un sous-espace vectoriel de l'espace vectoriel $\mathcal{C}^0([a,b],\mathbb{R})$. Pour cela il faut montrer[103] que E est non vide et que pour tout $(\lambda,\mu)\in\mathbb{R}^2$ et pour tout $(f,g)\in E^2$, la fonction $\lambda f+\mu g$ appartient à E.
L'ensemble E est non vide puisque la fonction nulle appartient clairement à E. De plus, si $(f,g)\in E^2$ alors $f(a)=g(a)=0$ et par conséquent pour tout $(\lambda,\mu)\in\mathbb{R}^2$ la fonction $\lambda f+\mu g\in\mathcal{C}^0([a,b],\mathbb{R})$ vérifie $\lambda f(a)+\mu g(a)=0$ ce qui permet d'affirmer que $\lambda f+\mu g$ appartient à E.

b) Pour montrer que N_2' est une norme sur E, reprenons notre raisonnement de la question 1-b. On a vu que la condition $N_2'(f)=0$ impliquait que f était une application constante sur $[a,b]$. Si f appartient à E alors $f(a)=0$ ce qui implique que la constante est 0 et l'application f est l'application nulle. On en conclut que N_2' est bien une norme sur E.

3 - a) Soit $f\in E$; pour tout $x\in[a,b]$ on a

$$f(x)=f(a)+\int_a^x f'(t)\,\mathrm{d}t=\int_a^x f'(t)\,\mathrm{d}t.$$

On en déduit que

$$|f(x)|=\left|\int_a^x f'(t)\,\mathrm{d}t\right|\leqslant\int_a^x|f'(t)|\,\mathrm{d}t$$

puis en utilisant l'inégalité de Cauchy-Schwarz[104] que

$$\begin{aligned}\int_a^x|f'(t)|\,\mathrm{d}t &\leqslant \left(\int_a^x 1^2\,\mathrm{d}t\right)^{\frac12}\times\left(\int_a^x|f'(t)|^2\,\mathrm{d}t\right)^{\frac12}\\ &= \sqrt{x-a}\left(\int_a^x|f'(t)|^2\,\mathrm{d}t\right)^{\frac12}\leqslant\sqrt{x-a}\left(\int_a^b|f'(t)|^2\,\mathrm{d}t\right)^{\frac12}\end{aligned}$$

On a donc $|f(x)|\leqslant\sqrt{x-a}\;N_2'(f)$. On en déduit alors que

$$\begin{aligned}N_2(f)^2 &= \int_a^b|f(x)|^2\,\mathrm{d}x\leqslant\int_a^b(x-a)\;N_2'(f)^2\,\mathrm{d}x\\ &= N_2'(f)^2\int_a^b(x-a)\,\mathrm{d}x=\frac{(b-a)^2}{2}\;N_2'(f)^2.\end{aligned}$$

On en conclut que $N_2(f)\leqslant\dfrac{b-a}{\sqrt{2}}\;N_2'(f)$.

---

[102] Voir le *Cours de première année* p. 307.
[103] Voir la définition 8.5 p. 306 du *Cours de première année*.
[104] Voir la proposition 18.6 p. 846 du *Cours de première année*.

b) Pour montrer que les deux normes ne sont pas équivalentes, montrons que l'assertion « $\exists C \in \mathbb{R}_+^* \quad \forall f \in E \quad N_2'(f) \leqslant CN_2(f)$ » est fausse. Il s'agit donc de montrer que $\forall C \in \mathbb{R}_+^* \quad \exists f \in E \quad N_2'(f) > CN_2(f)$.
Considérons les fonctions puissances de la forme $f : x \in]a,b] \mapsto (x-a)^\alpha$ pour $\alpha \in \mathbb{R}^+$ prolongées par continuité en $x = a$ en posant $f(a) = 0$. Ces applications sont de classe $\mathcal{C}^\infty$ sur $[a,b]$ et s'annulent en $x = a$, donc appartiennent à E. On a

$$\begin{aligned} N_2(f)^2 &= \int_a^b (t-a)^{2\alpha}\,\mathrm{d}t = \frac{1}{2\alpha+1}\big[(t-a)^{2\alpha+1}\big]_a^b = \frac{(b-a)^{2\alpha+1}}{2\alpha+1}, \\ N_2'(f)^2 &= \int_a^b \big(\alpha(t-a)^{\alpha-1}\big)^2\,\mathrm{d}t = \alpha^2 \int_a^b (t-a)^{2\alpha-2}\,\mathrm{d}t \\ &= \alpha^2 \frac{1}{2\alpha-1}\big[(t-a)^{2\alpha-1}\big]_a^b = \frac{\alpha^2(b-a)^{2\alpha-1}}{2\alpha-1}. \end{aligned}$$

Pour $C \in \mathbb{R}_+^*$ fixé, on cherche à déterminer α tel que $N_2'(f) > CN_2(f)$, c'est-à-dire tel que $N_2'(f)^2 > C^2 N_2(f)^2$. Remarquons que

$$N_2'(f)^2 > C^2 N_2(f)^2 \iff \frac{\alpha^2}{2\alpha-1} > C^2 \frac{(b-a)^2}{2\alpha+1}.$$

Posons $\alpha = C(b-a)$; comme $2\alpha - 1 < 2\alpha + 1$ on a

$$\frac{\alpha^2}{2\alpha-1} > \frac{\alpha^2}{2\alpha+1} = \frac{C^2(b-a)^2}{2\alpha-1}.$$

On en conclut que pour tout $C \in \mathbb{R}_+^*$, l'application $f : x \in]a,b] \mapsto (x-a)^\alpha$ où $\alpha = C(b-a)$ est telle que $N_2'(f) > CN_2(f)$. Les deux normes ne sont donc pas équivalentes.

Solution de l'exercice 7.17

1 - a) Comme la fonction puissance 3/2 est continue et dérivable sur $[0,+\infty[$, l'application f est continue et dérivable sur $[0,+\infty[$. Pour tout $x \in [0,+\infty[$, on a $f'(x) = \sqrt{x} - a$. On en déduit le tableau de variation suivant pour f :

x	0		a^2		$+\infty$
$f'(x)$	$-a$	$-$	0	$+$	$+\infty$
$f(x)$	$\frac{a^3}{3}$	↘	0	↗	$+\infty$

b) On déduit de la question précédente que pour tout $b \in \mathbb{R}^+$, on a $f(b) \geqslant 0$. Or,

$$f(b) \geqslant 0 \iff \frac{1}{3}a^3 + \frac{2}{3}b^{\frac{3}{2}} - ab \geqslant 0 \iff ab \leqslant \frac{1}{3}a^3 + \frac{2}{3}b^{\frac{3}{2}}.$$

2 - a) Soient $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ tels que $N_3(\mathbf{u}) = 1$ et $N_{\frac{3}{2}}(\mathbf{v}) = 1$; en utilisant l'inégalité obtenue à la question précédente, on obtient

$$\begin{aligned}\sum_{k=1}^{n} |u_k\, v_k| &= \sum_{k=1}^{n} |u_k|\,|v_k| \leqslant \sum_{k=1}^{n} \left(\frac{1}{3}|u_k|^3 + \frac{2}{3}|v_k|^{\frac{3}{2}}\right) \\ &= \frac{1}{3}\sum_{k=1}^{n} |u_k|^3 + \frac{2}{3}\sum_{k=1}^{n} |v_k|^{\frac{3}{2}} = \frac{1}{3}N_3(\mathbf{u})^3 + \frac{2}{3}N_{\frac{3}{2}}(\mathbf{v})^{\frac{2}{3}} = \frac{1}{3} + \frac{2}{3} = 1.\end{aligned}$$

b) Soient $\mathbf{U}$ et $\mathbf{V}$ deux vecteurs de $\mathbb{R}^n$ supposés non nuls et

$$\mathbf{u} = \frac{1}{N_3(\mathbf{U})}\,\mathbf{U} \quad \text{et} \quad \mathbf{v} = \frac{1}{N_{\frac{3}{2}}(\mathbf{V})}\,\mathbf{V}.$$

On a

$$\begin{aligned}N_3(\mathbf{u}) &= N_3\left(\frac{1}{N_3(\mathbf{U})}\,\mathbf{U}\right) = \left(\sum_{k=1}^{n} \frac{1}{N_3(\mathbf{U})^3}\,|U_k|^3\right)^{\frac{1}{3}} \\ &= \frac{1}{N_3(\mathbf{U})}\left(\sum_{k=1}^{n} |U_k|^3\right)^{\frac{1}{3}} = \frac{1}{N_3(\mathbf{U})}\,N_3(\mathbf{U}) = 1.\end{aligned}$$

De même, on vérifie que $N_{\frac{3}{2}}(\mathbf{v}) = 1$. On peut donc appliquer le résultat établi à la question précédente ; on a

$$\sum_{k=1}^{n} |u_k\, v_k| = \sum_{k=1}^{n} \left|\frac{1}{N_3(\mathbf{U})}\,U_k\,\frac{1}{N_{\frac{3}{2}}(\mathbf{V})}V_k\right| = \frac{1}{N_3(\mathbf{U})}\,\frac{1}{N_{\frac{3}{2}}(\mathbf{V})}\sum_{k=1}^{n} |U_k\, V_k\,|.$$

De la relation $|\sum_{k=1}^{n} u_k\, v_k| \leqslant 1$ on déduit que :

$$\sum_{k=1}^{n} |U_k\, V_k| \leqslant N_3(\mathbf{U})\; N_{\frac{3}{2}}(\mathbf{V}).$$

Si l'un des vecteurs $\mathbf{U}$ ou $\mathbf{V}$ est nul, la relation est encore vraie car on a dans ce cas : $\sum_{k=1}^{n} |U_k\, V_k| = 0$ et $N_3(\mathbf{U})\; N_{\frac{3}{2}}(\mathbf{V}) = 0$.

3 - a) Soient $\mathbf{x} = (x_1, \ldots, x_n)$ et $\mathbf{y} = (y_1, \ldots, y_n)$ deux vecteurs de $\mathbb{R}^n$; on a $\mathbf{x} + \mathbf{y} = (x_1 + y_1, \ldots, x_n + y_n)$ et

$$N_3(\mathbf{x} + \mathbf{y})^3 = \sum_{k=1}^{n} |x_k + y_k|^3 = \sum_{k=1}^{n} |x_k + y_k|\,|x_k + y_k|^2.$$

Or, $|x_k + y_k| \leqslant |x_k| + |y_k|$ donc

$$N_3(\mathbf{x} + \mathbf{y})^3 \leqslant \sum_{k=1}^{n} |x_k|\,|x_k + y_k|^2 + \sum_{k=1}^{n} |y_k|\,|x_k + y_k|^2.$$

b) En désignant par $\mathbf{z}$ le vecteur de composantes $((x_1+y_1)^2, \ldots, (x_n+y_n)^2)$, la relation précédente s'écrit

$$N_3(\mathbf{x}+\mathbf{y})^3 \leqslant \sum_{k=1}^{n} |x_k|\,|z_k| + \sum_{k=1}^{n} |y_k|\,|z_k|.$$

On en déduit en utilisant la relation établie à la question 2 que :

$$N_3(\mathbf{x}+\mathbf{y})^3 \leqslant N_3(\mathbf{x})\, N_{\frac{3}{2}}(\mathbf{z}) + N_3(\mathbf{y})\, N_{\frac{3}{2}}(\mathbf{z}).$$

Or, $N_{\frac{3}{2}}(\mathbf{z}) = \left(\sum_{k=1}^{n} |(x_k+y_k)^2|^{\frac{3}{2}}\right)^{\frac{2}{3}} = \left(\sum_{k=1}^{n} |x_k+y_k|^3\right)^{\frac{2}{3}} = N_3(\mathbf{x}+\mathbf{y})^2$, donc la relation précédente se réécrit

$$N_3(\mathbf{x}+\mathbf{y})^3 \leqslant N_3(\mathbf{x}+\mathbf{y})^2\ (N_3(\mathbf{x}) + N_3(\mathbf{y})).$$

Si $\mathbf{x}+\mathbf{y}$ n'est pas le vecteur nul, on en déduit en divisant par $N_3(\mathbf{x}+\mathbf{y})^2$ que

$$N_3(\mathbf{x}+\mathbf{y}) \leqslant N_3(\mathbf{x}) + N_3(\mathbf{y}).$$

Par ailleurs, la relation est encore vraie si $\mathbf{x}+\mathbf{y} = \mathbf{0}$ car on a alors $N_3(\mathbf{x}+\mathbf{y}) = 0$ et $N_3(\mathbf{x}) + N_3(\mathbf{y}) \geqslant 0$.

4 - Montrons que l'application $N_3 : \mathbf{x} \in \mathbb{R}^n \mapsto \left(\sum_{k=1}^{n} |x_k|^3\right)^{\frac{1}{3}}$ est une norme sur $\mathbb{R}^n$. L'application N_3 est clairement à valeur positive et $N_3(\mathbf{x}) = 0$ si et seulement si pour tout $k \in \{1, \ldots, n\}$ on a $|x_k| = 0$ (on a une somme de termes positifs) autrement dit, si et seulement si $\mathbf{x}$ est le vecteur nul. Par ailleurs, pour tout $\lambda \in \mathbb{R}$ et pour tout $\mathbf{x} \in \mathbb{R}^n$, on a

$$\begin{aligned} N_3(\lambda\mathbf{x}) &= \left(\sum_{k=1}^{n} |\lambda x_k|^3\right)^{\frac{1}{3}} = \left(\sum_{k=1}^{n} |\lambda|^3\, |x_k|^3\right)^{\frac{1}{3}} = \left(|\lambda|^3\right)^{\frac{1}{3}} \left(\sum_{k=1}^{n} |x_k|^3\right)^{\frac{1}{3}} \\ &= |\lambda|\, N_3(\mathbf{x}). \end{aligned}$$

Enfin, le calcul effectué à la question précédente établit l'inégalité triangulaire : pour tous $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ on a $N_3(\mathbf{x}+\mathbf{y}) \leqslant N_3(\mathbf{x}) + N_3(\mathbf{y})$.

CHAPITRE 8

Algèbre bilinéaire

La plupart des lois physiques, des modèles économiques, des processus biologiques sont décrits par des équations non linéaires. Une des façons les plus simples d'étudier de tels phénomènes est de les « linéariser » (c'est-à-dire de les approcher par des équations linéaires). La formule de Taylor[1] indique que l'on peut approcher une fonction de $\mathbb{R}$ dans $\mathbb{R}$ qui est régulière, au premier ordre par une fonction linéaire et au second ordre par une fonction « quadratique ». L'interprétation géométrique de ces approximations est la suivante : une courbe est approchée au premier ordre par sa tangente ; au second ordre, elle est approchée par une courbe qui s'appelle une « conique ». Les notions relevant de l'algèbre linéaire ont été introduits dans le chap. 9 du *Cours de première année*. Ce chapitre est consacré à l'algèbre bilinéaire. Outre les formes quadratiques, différentes notions relevant de l'algèbre bilinéaire seront présentées dans ce chapitre, dont celle de produit scalaire. L'algèbre bilinéaire est donc une porte naturelle vers les notions d'orthogonalité et d'espace euclidien. Sauf mention explicite du contraire, tous les espaces vectoriels considérés ici sont des espaces vectoriels réels, c'est-à-dire que le corps de base est $\mathbb{R}$.

8.1 Formes bilinéaires et formes quadratiques

8.1.1 Formes bilinéaires

Rappelons[2] qu'une application f entre deux $\mathbb{R}$-espaces vectoriels E et F est qualifiée d'application linéaire lorsqu'elle vérifie l'assertion suivante :

$$\forall(\mathbf{x},\mathbf{y}) \in E^2 \quad \forall(\lambda,\mu) \in \mathbb{R}^2 \quad f(\lambda\mathbf{x}+\mu\mathbf{y}) = \lambda f(\mathbf{x}) + \mu f(\mathbf{y}).$$

Sur le même principe, une application f « à deux variables » de $E \times E$ vers F est appelée application bilinéaire si elle est linéaire par rapport à chacune de ces deux variables. Plus précisément, on a la définition suivante.

---

[1] Voir le théorème 16.4 p. 750 du *Cours de première année*.

[2] Voir le *Cours de première année* p. 355.

DÉFINITION 8.1 *Soient E et F deux $\mathbb{R}$-espaces vectoriels. L'application*

$$\begin{aligned} \varphi \ : \ E \times E &\longrightarrow F \\ (\mathbf{x}, \mathbf{y}) &\longmapsto \varphi(\mathbf{x}, \mathbf{y}) \end{aligned}$$

*est qualifiée d'*application bilinéaire *si*

$$\forall (\mathbf{x}, \mathbf{y}, \mathbf{z}) \in E^3 \quad \forall (\lambda, \mu) \in \mathbb{R}^2 \quad \big(\varphi(\lambda \mathbf{x} + \mu \mathbf{y}, \mathbf{z}) = \lambda \varphi(\mathbf{x}, \mathbf{z}) + \mu \varphi(\mathbf{y}, \mathbf{z}) \quad \textit{et} \quad \varphi(\mathbf{x}, \lambda \mathbf{y} + \mu \mathbf{z}) = \lambda \varphi(\mathbf{x}, \mathbf{y}) + \mu \varphi(\mathbf{x}, \mathbf{z}) \big).$$

Il découle de cette définition que si les vecteurs $\mathbf{x}$ et $\mathbf{y}$ s'expriment comme des combinaisons linéaires à coefficients réels d'éléments de E, *i.e.* si $\mathbf{x} = \sum_{i=1}^{p} \lambda_i \mathbf{x}_i$ et $\mathbf{y} = \sum_{j=1}^{q} \mu_j \mathbf{y}_j$ avec λ_i et μ_j des réels et $\mathbf{x}_i$, $\mathbf{y}_i$ des éléments de E, alors on a

$$\begin{aligned} \varphi(\mathbf{x}, \mathbf{y}) &= \varphi\Big(\sum_{i=1}^{p} \lambda_i \mathbf{x}_i, \mathbf{y}\Big) = \sum_{i=1}^{p} \lambda_i \varphi(\mathbf{x}_i, \mathbf{y}) \\ &= \sum_{i=1}^{p} \lambda_i \varphi\Big(\mathbf{x}_i, \sum_{j=1}^{q} \mu_j \mathbf{y}_j\Big) = \sum_{i=1}^{p} \sum_{j=1}^{q} \lambda_i \mu_j \varphi(\mathbf{x}_i, \mathbf{y}_j). \end{aligned} \tag{1}$$

D'autre part, il résulte de la définition 8.1 que l'on a pour tout $\mathbf{x} \in E$,

$$\varphi(\mathbf{x}, \mathbf{0}_E) = \mathbf{0}_F \quad \text{et} \quad \varphi(\mathbf{0}_E, \mathbf{x}) = \mathbf{0}_F.$$

Structure d'espace vectoriel

On note $\mathcal{L}_2(E \times E, F)$ l'ensemble des applications bilinéaires de $E \times E$ dans F. L'ensemble $\mathcal{L}_2(E \times E, F)$ possède une structure de $\mathbb{R}$-espace vectoriel[3] : pour tout $(f, g) \in \mathcal{L}_2(E \times E, F)^2$ et pour tout $(\alpha, \beta) \in \mathbb{R}^2$ on a

$$\alpha \cdot f + \beta \cdot g \in \mathcal{L}_2(E \times E, F).$$

Remarques

1. On appelle[4] « forme » toute application définie sur un espace vectoriel qui est à valeurs dans le corps de base de cet espace vectoriel. Ainsi, une *forme bilinéaire* sur le $\mathbb{R}$-espace vectoriel E est une application bilinéaire f de $E \times E$ à valeurs dans $\mathbb{R}$.

2. Dans ce chapitre, le corps de base des espaces vectoriels est le corps des nombres réels $\mathbb{R}$. Signalons toutefois qu'une théorie similaire à celle que nous

[3] Rappelons que l'ensemble des applications de $E \times E$ vers F est un $\mathbb{R}$-espace vectoriel, voir le *Cours de première année* p. 300, dont l'ensemble des applications bilinéaires de $E \times E$ vers F constitue un sous-espace vectoriel. Il possède ainsi lui-même une structure de $\mathbb{R}$-espace vectoriel.

[4] Voir le *Cours de première année* p. 355 et p. 465.

présentons dans ce chapitre existe pour des espaces vectoriels sur des corps différents, comme par exemple le corps des nombres complexes. Dans le cas d'un $\mathbb{C}$-espace vectoriel, l'analogue des formes bilinéaires de la définition 8.1 sont les formes « sesquilinéaires ». Une forme *sesquilinéaire* ψ sur un $\mathbb{C}$-espace vectoriel E est une application de $E \times E$ dans $\mathbb{C}$ qui vérifie

$$\forall(\mathbf{x},\mathbf{y},\mathbf{z}) \in E^3 \quad \forall(\lambda,\mu) \in \mathbb{C}^2 \quad \big(\varphi(\lambda\mathbf{x}+\mu\mathbf{y},\mathbf{z}) = \overline{\lambda}\,\varphi(\mathbf{x},\mathbf{z}) + \overline{\mu}\,\varphi(\mathbf{y},\mathbf{z}) \quad \text{et} \quad \varphi(\mathbf{x},\lambda\mathbf{y}+\mu\mathbf{z}) = \lambda\,\varphi(\mathbf{x},\mathbf{y}) + \mu\,\varphi(\mathbf{x},\mathbf{z}) \big)$$

où $\overline{\lambda}$ et $\overline{\mu}$ désignent les complexes conjugués de λ et μ respectivement. ◇

Exemples

1. On vérifiera sans peine que l'application $\varphi : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ définie pour $\mathbf{x} = (x_1, x_2, x_3)$ et $\mathbf{y} = (y_1, y_2, y_3)$ par

$$\varphi(\mathbf{x},\mathbf{y}) = x_1\ y_1 + x_2\ y_2 + x_3\ y_3$$

est une forme bilinéaire sur $E = \mathbb{R}^3$.

2. Dans l'espace vectoriel $\mathbb{R}^4$, l'application $\varphi : \mathbb{R}^4 \times \mathbb{R}^4 \to \mathbb{R}$ définie pour $\mathbf{x} = (x_1, x_2, x_3, x_4)$ et $\mathbf{y} = (y_1, y_2, y_3, y_4)$ par

$$\varphi(\mathbf{x},\mathbf{y}) = x_1\ y_1 + x_2\ y_2 + x_3\ y_3 - c^2 x_4\ y_4$$

où c désigne un paramètre réel, est une forme bilinéaire. Cette forme bilinéaire qui intervient dans la théorie de la relativité est nommée forme de Lorentz[5].

3. Sur l'espace vectoriel $\mathcal{C}([0,1],\mathbb{R})$ des applications continues de $[0,1]$ dans $\mathbb{R}$, l'application $\varphi : \mathcal{C}([0,1],\mathbb{R}) \times \mathcal{C}([0,1],\mathbb{R}) \to \mathbb{R}$ définie par

$$\varphi(f,g) = \int_0^1 f(t)g(t)\ \mathrm{d}t$$

est une forme bilinéaire. Ce résultat découle des propriétés de linéarité de l'intégrale[6].

4. Sur l'espace vectoriel $\mathcal{M}_n(\mathbb{R})$ des matrices carrées de taille n à coefficients réels, l'application φ définie par

$$\varphi : (\mathrm{A},\mathrm{B}) \in \mathcal{M}_n(\mathbb{R}) \times \mathcal{M}_n(\mathbb{R}) \longmapsto \mathrm{Tr}({}^t\mathrm{AB}) \in \mathbb{R}$$

est une forme bilinéaire. Le résultat découle de la linéarité de l'application Trace[7] qui à une matrice de $\mathcal{M}_n(\mathbb{R})$ associe la somme de ses éléments diagonaux.

---

[5] Lorentz, Hendrik Antoon (1853, Arnhem - 1928, Haarlem). Physicien néerlandais qui reçut le prix Nobel de physique en 1902.

[6] Voir le *Cours de première année*, chap. 18.

[7] Voir *Cours de première année* p. 448.

8.1.2 Cas d'un espace vectoriel en dimension finie

Rappel sur le lien entre les matrices et les applications linéaires

Soient E un $\mathbb{R}$-espace vectoriel de dimension p muni d'une base $\mathcal{B}$ et F un $\mathbb{R}$-espace vectoriel de dimension n muni d'une base $\mathcal{B}'$. À une application **linéaire** f de E dans F, on peut[8] associer une matrice relativement aux bases $\mathcal{B}$ et $\mathcal{B}'$. C'est une matrice de taille (n,p) à coefficients dans $\mathbb{R}$ dont la j–ième colonne est constituée par les coordonnées de l'image du j–ième vecteur de la base de départ $\mathcal{B}$ par rapport à la base d'arrivée $\mathcal{B}'$. On la note

$$\mathrm{Mat}_{\mathcal{B},\mathcal{B}'}(f)$$

et on dit que la matrice $\mathrm{Mat}_{\mathcal{B},\mathcal{B}'}(f)$ représente f dans les bases $\mathcal{B}$ et $\mathcal{B}'$.

En pratique, en notant $\mathcal{B} = (\mathbf{e}_1, ..., \mathbf{e}_p)$ et $\mathcal{B}' = (\mathbf{e}'_1, ..., \mathbf{e}'_n)$ les deux bases, on a de façon symbolique

$$\mathrm{Mat}_{\mathcal{B},\mathcal{B}'}(f) = \begin{array}{c} \begin{array}{ccc} f(\boldsymbol{e}_1) & f(\boldsymbol{e}_2) & f(\boldsymbol{e}_p\ ) \end{array} \\ \begin{pmatrix} a_{11} & a_{12} & a_{1p} \\ a_{21} & a_{22} & a_{2p} \\ \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & a_{np} \end{pmatrix} \end{array} \begin{array}{c} \boldsymbol{e}'_1 \\ \boldsymbol{e}'_2 \\ \vdots \\ \boldsymbol{e}'_n \end{array}.$$

Cette identification entre les applications linéaires et les matrices permet d'utiliser toute une classe de résultats issue de l'algèbre matriciel.

Formes bilinéaires et matrices

Comme pour les applications linéaires, il est possible de représenter les formes bilinéaires par des matrices. Insistons sur le fait qu'une telle représentation n'est possible que pour des formes bilinéaires et non pour des applications bilinéaires à valeurs dans un espace vectoriel quelconque.

Soit E un $\mathbb{R}$-espace vectoriel de dimension n muni d'une base $\mathcal{B} = (\mathbf{e}_1, ..., \mathbf{e}_n)$. Considérons un couple $(\mathbf{x}, \mathbf{y}) \in E^2$ donné dans la base $\mathcal{B}$ par

$$\mathbf{x} = \sum_{i=1}^{n} x_i\ \mathbf{e}_i \quad \text{et} \quad \mathbf{y} = \sum_{j=1}^{n} y_j\ \mathbf{e}_j.$$

La bilinéarité de φ, voir la relation (1), entraîne que $\varphi(\mathbf{x}, \mathbf{y})$ peut s'écrire sous la forme

$$\varphi(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i\ y_j\ \varphi(\mathbf{e}_i, \mathbf{e}_j). \tag{2}$$

Cela signifie que la connaissance des valeurs de $\varphi(\mathbf{e}_i, \mathbf{e}_j)$ pour tout couple $(i,j) \in \{1, ...n\}$ permet d'évaluer $\varphi(\mathbf{x}, \mathbf{y})$ dès que l'on sait décomposer $\mathbf{x}$ et $\mathbf{y}$

[8] Voir le *Cours de première année* p. 407.

dans la base $\mathcal{B}$: une forme bilinéaire est déterminée par les valeurs $a_{ij} = \varphi(\mathbf{e}_i, \mathbf{e}_j)$ qu'elle prend sur une base de E.

DÉFINITION 8.2 *Soient E un $\mathbb{R}$-espace vectoriel de dimension n muni d'une base $\mathcal{B} = (\mathbf{e}_1, ..., \mathbf{e}_n)$ et φ une forme bilinéaire sur E. On appelle* matrice de la forme bilinéaire φ par rapport à la base $\mathcal{B}$ *la matrice de $\mathcal{M}_n(\mathbb{R})$ dont les coefficients sont $\varphi(\mathbf{e}_i, \mathbf{e}_j)$ pour $(i, j) \in \{1, ..., n\}^2$. On la note*

$$\mathrm{Mat}_{\mathcal{B}}(\varphi) = \begin{pmatrix} \varphi(\mathbf{e}_1, \mathbf{e}_1) & \varphi(\mathbf{e}_1, \mathbf{e}_2) & \varphi(\mathbf{e}_1, \mathbf{e}_n) \\ \varphi(\mathbf{e}_2, \mathbf{e}_1) & \varphi(\mathbf{e}_2, \mathbf{e}_2) & \varphi(\mathbf{e}_2, \mathbf{e}_n) \\ \vdots & \vdots & \vdots \\ \varphi(\mathbf{e}_n, \mathbf{e}_1) & \varphi(\mathbf{e}_n, \mathbf{e}_2) & \varphi(\mathbf{e}_n, \mathbf{e}_n) \end{pmatrix}$$

et on dit que la matrice $\mathrm{Mat}_{\mathcal{B}}(\varphi)$ *représente φ dans la base $\mathcal{B}$.*

La relation (2) s'écrit matriciellement

$$\varphi(\mathbf{x}, \mathbf{y}) = {}^t\mathrm{X}\mathrm{A}\mathrm{Y} \tag{3}$$

où $\mathrm{A} = \mathrm{Mat}_{\mathcal{B}}(\varphi)$ et où X et Y sont les matrices colonnes représentant $\mathbf{x}$ et $\mathbf{y}$ dans la base $\mathcal{B}$:

$$\mathrm{X} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{et} \quad \mathrm{Y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Remarques

1. Inversement, toute matrice de $\mathcal{M}_n(\mathbb{R})$ définit une forme bilinéaire sur E *via* la formule (3).

2. L'application qui à une forme bilinéaire φ associe sa matrice est un isomorphisme d'espaces vectoriels[(9)] entre l'espace vectoriel $\mathcal{L}_2(E \times E, \mathbb{R})$ et l'espace vectoriel $\mathcal{M}_n(\mathbb{R})$. L'existence de cet isomorphisme permet d'affirmer que les dimensions du $\mathbb{R}$-espace vectoriel $\mathcal{L}_2(E \times E, \mathbb{R})$ et du $\mathbb{R}$-espace vectoriel $\mathcal{M}_n(\mathbb{R})$ sont les mêmes. Sachant que l'espace vectoriel $\mathcal{M}_n(\mathbb{R})$ est de dimension n^2, on en déduit que $\dim_{\mathbb{R}} \mathcal{L}_2(E \times E, \mathbb{R}) = n^2$. ◇

Exemples

1. Pour déterminer la matrice de la forme bilinéaire définie sur $\mathbb{R}^3$ par

$$\forall (\mathbf{x}, \mathbf{y}) \in \mathbb{R}^3 \quad \varphi(\mathbf{x}, \mathbf{y}) = x_1\, y_1 + x_2\, y_2 + x_3\, y_3$$

(9) Pour la définition d'isomorphisme d'espaces vectoriels, voir le *Cours de première année* p. 358.

dans la base canonique[10] de $\mathbb{R}^3$, on calcule les valeurs de $\varphi(\mathbf{e}_i, \mathbf{e}_j)$ pour tous les couples $(i,j) \in \{1,2,3\}^2$. On a par exemple $\varphi(\mathbf{e}_1, \mathbf{e}_1) = 1$ et $\varphi(\mathbf{e}_1, \mathbf{e}_2) = 0$. On obtient

$$\mathrm{Mat}_{\mathcal{B}_{\mathrm{can}}}(\varphi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

2. La forme de Lorentz définie sur $\mathbb{R}^4$ par

$$\varphi(\mathbf{x}, \mathbf{y}) = x_1\ y_1 + x_2\ y_2 + x_3\ y_3 - c^2 x_4\ y_4 \qquad \text{avec } c \in \mathbb{R}$$

admet pour matrice dans la base canonique de $\mathbb{R}^4$

$$\mathrm{Mat}_{\mathcal{B}_{\mathrm{can}}}(\varphi) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -c^2 \end{pmatrix}.$$

Exercice 8.1 Déterminer la matrice de l'application φ définie sur l'espace vectoriel $\mathcal{M}_2(\mathbb{R}) \times \mathcal{M}_2(\mathbb{R})$ par

$$\varphi : (\mathrm{A}, \mathrm{B}) \in \mathcal{M}_2(\mathbb{R}) \times \mathcal{M}_2(\mathbb{R}) \longmapsto \mathrm{Tr}({}^t\mathrm{AB})$$

par rapport à base canonique[11] de $\mathcal{M}_2(\mathbb{R})$.

Une matrice carré de taille n à coefficients réels peut représenter aussi bien une application linéaire entre deux espaces vectoriels de dimension n qu'une forme bilinéaire sur un espace de dimension n. On sera attentif à ne pas confondre ces deux notions (applications linéaires de E vers F et formes bilinéaires sur E) même si toutes les deux peuvent être représentées par le même objet mathématique : une matrice.

8.1.3 Changement de base

Bien entendu, comme pour les applications linéaires, la notion de matrice représentant une forme bilinéaire dépend de la base choisie sur l'espace vectoriel E. Il est donc important de savoir comment est modifiée cette représentation lorsque l'on change de base.

---

[10] Rappelons que la base canonique de $\mathbb{R}^n$, $n \in \mathbb{N}^*$ est formée des n vecteurs de composantes : $(1,0,0,\ldots,0), (0,1,0,\ldots,0), \ldots, (0,\ldots,0,0,1)$.

[11] La base canonique de l'espace vectoriel $\mathcal{M}_2(\mathbb{R})$ est constituée des matrices carrées, notées E^{ij}, ne comportant que des coefficients nuls sauf celui en position (i,j) qui vaut 1.

Changement de base pour représenter une application linéaire

Commençons par effectuer quelques rappels concernant les applications linéaires. Considérons une application **linéaire** f entre deux $\mathbb{R}$-espaces vectoriels E et F. On munit E des bases $\mathcal{B}$ et $\mathcal{B}'$ et F des bases $\mathcal{C}$ et $\mathcal{C}'$. On sait[12] que les deux matrices $\mathrm{M} = \mathrm{Mat}_{\mathcal{B},\mathcal{C}}(f)$ et $\mathrm{M}' = \mathrm{Mat}_{\mathcal{B}',\mathcal{C}'}(f)$ vérifient :

$$\mathrm{M}' = \mathrm{Q}^{-1}\mathrm{MP} \tag{4}$$

où P est la matrice de passage[13] de $\mathcal{B}$ à $\mathcal{B}'$ etQ est la matrice de passage de $\mathcal{C}$ à $\mathcal{C}'$. Deux matrices M et M′ pour lesquelles il existe deux matrices inversibles P et Q satisfaisant l'égalité (4) sont dites *équivalentes.* Deux matrices sont donc équivalentes si elles représentent la même application linéaire dans des bases différentes.

Considérons maintenant un endomorphisme f de E (c'est-à-dire une application **linéaire** de E dans E) et deux bases $\mathcal{B}$ et $\mathcal{B}'$ de l'espace vectoriel E. Les deux matrices $\mathrm{M} = \mathrm{Mat}_{\mathcal{B}}(f)$ et $\mathrm{M}' = \mathrm{Mat}_{\mathcal{B}'}(f)$ vérifient :

$$\mathrm{M}' = \mathrm{P}^{-1}\mathrm{MP} \tag{5}$$

où P désigne la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$. Deux matrices M et M′ pour lesquels il existe une matrice inversible P satisfaisant l'égalité (5) sont dites *semblables.* Deux matrices carrées sont donc semblables si elles représentent le même endomorphisme dans deux bases différentes.

Rappelons que l'on note $\mathrm{GL}_n(\mathbb{R})$ le sous-espace vectoriel de $\mathcal{M}_n(\mathbb{R})$ formé par les matrices inversibles.

Changement de base pour représenter une application bilinéaire

Intéressons-nous à présent au changement de base dans la représentation matricielle d'une application bilinéaire.

Proposition 8.1 *Soient $\mathcal{B}$ et $\mathcal{B}'$ deux bases d'un espace vectoriel E de dimension finie et* P *la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$. Si φ est une* **forme bilinéaire** *sur E, de matrice* M *dans la base $\mathcal{B}$ et de matrice* M′ *dans la base $\mathcal{B}'$ alors*

$$\mathrm{M}' = {}^t\mathrm{PMP}.$$

Démonstration Soient $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ et $\mathcal{B}' = (\mathbf{e}'_1, \mathbf{e}'_2, .., \mathbf{e}'_n)$ deux bases d'un espace vectoriel E de dimension finie n, et φ une forme bilinéaire sur E. Par définition de la matrice d'une **forme bilinéaire** dans une base donnée[14] les

[12] Voir le *Cours de première année* p. 441.
[13] Voir le *Cours de première année* p. 434.
[14] Voir la définition 8.2.

matrices $M = \mathrm{Mat}_{\mathcal{B}}(\varphi)$ et $M' = \mathrm{Mat}_{\mathcal{B}'}(\varphi)$ ont respectivement pour coefficients $(m_{ij})_{(i,j)\in\{1,\dots,n\}^2}$ et $(m'_{ij})_{(i,j)\in\{1,\dots,n\}^2}$ où

$$\forall (i,j) \in \{1,\dots,n\}^2 \qquad m_{ij} = \varphi(\mathbf{e}_i, \mathbf{e}_j) \quad \text{et} \quad m'_{ij} = \varphi(\mathbf{e}'_i, \mathbf{e}'_j).$$

Pour établir le lien entre les coefficients m_{ij} et m'_{ij}, on part de celui existant entre les deux bases $\mathcal{B}$ et $\mathcal{B}'$. En notant $P = (p_{ij})_{ij}$ la matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{B}'$, on a[15]

$$\forall i \in \{1,\dots,n\} \qquad \mathbf{e}'_i = \sum_{k=1}^{n} p_{ki}\, \mathbf{e}_k.$$

En utilisant la bilinéarité de l'application φ, On obtient pour tout $(i,j) \in \{1,\dots,n\}^2$,

$$m'_{ij} = \varphi\Big(\sum_{k=1}^{n} p_{ki}\, \mathbf{e}_k\,,\, \sum_{\ell=1}^{n} p_{\ell j}\, \mathbf{e}_\ell\Big) = \sum_{k=1}^{n}\sum_{\ell=1}^{n} p_{ki}\, p_{\ell j}\, \varphi(\mathbf{e}_k, \mathbf{e}_\ell).$$

Cette égalité se réécrit en faisant intervenir les coefficients $m_{ij} = \varphi(\mathbf{e}_i, \mathbf{e}_j)$ de la manière suivante :

$$m'_{ij} = \sum_{k=1}^{n}\sum_{\ell=1}^{n} p_{ki}\, p_{\ell j}\, m_{k\ell} = \sum_{k=1}^{n} p_{ki} \Big(\sum_{\ell=1}^{n} m_{k\ell}\, p_{\ell j}\Big) = \sum_{k=1}^{n} p_{ki}\, q_{kj}$$

où q_{kj} désigne le coefficient (k,j) de la matrice MP. Le membre de droite de l'égalité correspond donc à l'expression du terme (i,j) du produit matriciel ${}^t\mathrm{P}(\mathrm{MP})$, i.e.

$$m'_{ij} = ({}^t\mathrm{PMP})_{ij}.$$

Cette relation étant vraie pour tout $(i,j) \in \{1,\dots,n\}^2$, on a bien $M' = {}^t\mathrm{PMP}$. □

Matrices congrues

Définition 8.3 *Une matrice* M *de* $\mathcal{M}_n(\mathbb{R})$ *est dite* congrue *à une matrice* M′ *de* $\mathcal{M}_n(\mathbb{R})$ *si*

$$\exists\, \mathrm{P} \in \mathrm{GL}_n(\mathbb{R}) \quad M' = {}^t\mathrm{PMP}.$$

Remarque La définition 8.3 indique que deux matrices carrées de même taille sont congrues si elles représentent la même application bilinéaire dans deux bases différentes. ◇

Proposition 8.2 *La relation de congruence entre matrices de* $\mathcal{M}_n(\mathbb{R})$ *définit une relation d'équivalence sur* $\mathcal{M}_n(\mathbb{R})$.

Démonstration Pour vérifier que la relation de congruence est une relation d'équivalence, il faut montrer[16] que cette relation est réflexive, symétrique et

[15] Voir la *Cours de première année*, p. 434.

[16] Voir la définition 2.28 p. 55 du *Cours de première année*.

transitive.

– Toute matrice M de $\mathcal{M}_n(\mathbb{R})$ est congrue à elle-même (propriété de réflexivité) puisque

$$M = {}^tI_n M I_n \quad \text{et} \quad I_n \in GL_n(\mathbb{R}).$$

– Soient M et M′ deux matrices de $\mathcal{M}_n(\mathbb{R})$. Montrons que si M est congrue à M′ alors M′ est congrue à M (propriété de symétrie). La matrice M est congrue à M′ si

$$\exists P \in GL_n(\mathbb{R}) \quad M' = {}^tPMP.$$

En multipliant à gauche par $({}^tP)^{-1}$ et à droite par P^{-1}, on en déduit la relation $({}^tP)^{-1}M'P^{-1} = M$, puis la relation[17] $M = {}^t(P^{-1})M'P^{-1}$ qui permet de conclure que M′ est congrue à M.

– Soient M, M′ et M″ trois matrices de $\mathcal{M}_n(\mathbb{R})$. Montrons que si M est congrue à M′ et si M′ est congrue à M″ alors M est congrue à M″ (propriété de transitivité). D'une part, si M est congrue à M′ alors

$$\exists P \in GL_n(\mathbb{R}) \quad M' = {}^tPMP. \tag{6}$$

D'autre part, si M′ est congrue à M″ alors

$$\exists Q \in GL_n(\mathbb{R}) \quad M'' = {}^tQM'Q. \tag{7}$$

En combinant la relation (6) avec la relation (7), on obtient $M'' = {}^tQ{}^tPMPQ$. Comme ${}^tQ{}^tP = {}^t(PQ)$ et que le produit de 2 matrices inversibles est une matrice inversible, cette relation s'écrit sous la forme $M'' = {}^t(PQ)M(PQ)$. On en conclut que M″ est congrue à M. □

Si deux matrices sont congrues alors elles sont équivalentes. Par ailleurs, le fait que deux matrices équivalentes aient toujours le même rang[18] justifie la définition suivante.

Définition 8.4 *Soit φ une application bilinéaire sur un $\mathbb{R}$-espace vectoriel de dimension finie. On appelle* rang de l'application bilinéaire φ *le rang de la matrice de φ dans une base quelconque de E.*

Exemple Calculons le rang de la forme de Lorentz dont la matrice dans la base canonique de $\mathbb{R}^4$ s'écrit, voir p. 441,

$$\mathrm{Mat}_{\mathcal{B}_{can}}(\varphi) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -c^2 \end{pmatrix}.$$

[17] On utilise ici le fait que l'inverse de ${}^tP$ est aussi la transposée de P^{-1}, c'est-à-dire que pour toute matrice $P \in GL_n(\mathbb{R})$ on a ${}^t(P^{-1}) = ({}^tP)^{-1}$.

[18] Voir le *Cours de première année*, corollaire 10.2 p. 447.

Les trois premiers vecteurs lignes de cette matrice forment une famille de trois vecteurs libres dans $\mathbb{R}^4$ (ce sont en effet les trois premiers vecteurs de la base canonique de $\mathbb{R}^4$). Si $c = 0$ alors le quatrième vecteur ligne de la matrice $\mathrm{Mat}_{\mathcal{B}_{\mathrm{can}}}(\varphi)$ est le vecteur nul, il est par conséquent lié aux trois premiers et l'espace vectoriel engendré par les quatre vecteurs lignes de la matrice $\mathrm{Mat}_{\mathcal{B}_{\mathrm{can}}}(\varphi)$ est un espace vectoriel de dimension 3. Dans ce cas, le rang de cette matrice et donc le rang de la forme de Lorentz valent 3. Par contre, si $c \neq 0$ alors les quatre vecteurs lignes de la matrice $\mathrm{Mat}_{\mathcal{B}_{\mathrm{can}}}(\varphi)$ forment une famille libre de $\mathbb{R}^4$ et engendrent un espace vectoriel de dimension 4. Dans ce cas, le rang de la forme de Lorentz vaut 4.

8.1.4 Formes bilinéaires symétriques

DÉFINITION 8.5 *Soit φ une forme bilinéaire sur un $\mathbb{R}$-espace vectoriel E.*
✗ *φ est dite* symétrique *si pour tout $(\mathbf{x},\mathbf{y}) \in E^2$ on a $\varphi(\mathbf{x},\mathbf{y}) = \varphi(\mathbf{y},\mathbf{x})$.*
✗ *φ est dite* antisymétrique *si pour tout $(\mathbf{x},\mathbf{y}) \in E^2$ on a $\varphi(\mathbf{x},\mathbf{y}) = -\varphi(\mathbf{y},\mathbf{x})$.*

Par la suite, on notera $\mathcal{L}_2^{\mathrm{s}}(E \times E, \mathbb{R})$ l'ensemble des formes bilinéaires symétriques sur E et $\mathcal{L}_2^{\mathrm{a}}(E \times E, \mathbb{R})$ celui des formes bilinéaires antisymétriques sur E. Ces deux ensembles sont des sous-espaces vectoriels de l'espace vectoriel $\mathcal{L}_2(E \times E, \mathbb{R})$ des formes bilinéaires sur E.

Exemples

1. On vérifie facilement que la forme bilinéaire sur $\mathbb{R}^3$ définie par

$$\forall (\mathbf{x},\mathbf{y}) \in \mathbb{R}^3 \quad \varphi(\mathbf{x},\mathbf{y}) = x_1\, y_1 + x_2\, y_2 + x_3\, y_3$$

est une forme bilinéaire symétrique.

2. La forme bilinéaire sur $\mathcal{M}_n(\mathbb{R})$ définie par

$$\varphi : (\mathrm{A},\mathrm{B}) \in \mathcal{M}_n(\mathbb{R}) \times \mathcal{M}_n(\mathbb{R}) \longmapsto \mathrm{Tr}({}^t\mathrm{AB})$$

est une forme bilinéaire symétrique. La trace d'une matrice étant égale à la somme des éléments diagonaux et lors de la transposition les éléments diagonaux restant inchangés, pour tout $(\mathrm{A},\mathrm{B}) \in \mathcal{M}_n(\mathbb{R})^2$, on a

$$\varphi(\mathrm{A},\mathrm{B}) = \mathrm{Tr}({}^t\mathrm{AB}) = \mathrm{Tr}({}^t({}^t\mathrm{AB})).$$

Comme[19] ${}^t({}^t\mathrm{AB}) = {}^t\mathrm{BA}$, on en déduit que $\varphi(\mathrm{A},\mathrm{B}) = \mathrm{Tr}({}^t\mathrm{BA}) = \varphi(\mathrm{B},\mathrm{A})$.

La proposition suivante résulte de la définition 8.5 ainsi que de la représentation matricielle des formes bilinéaires en dimension finie introduite dans la partie précédente.

[19] On utilise ici le fait que d'une part pour tout couple $(\mathrm{C},\mathrm{D}) \in \mathcal{M}_n(\mathbb{R})^2$ on a ${}^t(\mathrm{CD}) = {}^t\mathrm{D}\,{}^t\mathrm{C}$ et d'autre part, pour toute matrice $\mathrm{C} \in \mathcal{M}_n(\mathbb{R})$ on a ${}^t({}^t\mathrm{C}) = \mathrm{C}$.

Proposition 8.3 *Soient E un $\mathbb{R}$-espace vectoriel de dimension finie, $\mathcal{B}$ une base de E et φ une forme bilinéaire sur E. La forme bilinéaire φ est symétrique (resp. antisymétrique) si et seulement si la matrice $\mathrm{Mat}_{\mathcal{B}}(\varphi)$ est symétrique (resp. antisymétrique).*

Rappelons[20] qu'une matrice A de $\mathcal{M}_n(\mathbb{R})$ est dite *symétrique* si ${}^t\mathrm{A} = \mathrm{A}$ et qu'elle est dite *antisymétrique* si ${}^t\mathrm{A} = -\mathrm{A}$. L'ensemble des matrices symétriques de taille n à coefficients réels est noté $\mathcal{S}_n(\mathbb{R})$ et l'ensemble des matrices antisymétriques de taille n à coefficients réels est noté $\mathcal{A}_n(\mathbb{R})$. On peut montrer[21] que les deux sous-espaces vectoriels $\mathcal{S}_n(\mathbb{R})$ et $\mathcal{A}_n(\mathbb{R})$ sont en somme directe[22] dans $\mathcal{M}_n(\mathbb{R})$. Plus précisément, on a la proposition suivante, dont la preuve qui a été effectuée dans le *Cours de première année* p. 402, est basée sur la décomposition de toute matrice M de $\mathcal{M}_n(\mathbb{R})$ comme la somme d'une matrice symétrique S et d'une matrice antisymétrique A :

$$\mathrm{M} = \underbrace{\tfrac{1}{2}\left(\mathrm{M} + {}^t\mathrm{M}\right)}_{=\mathrm{S}} + \underbrace{\tfrac{1}{2}\left(\mathrm{M} - {}^t\mathrm{M}\right)}_{=\mathrm{A}}. \tag{8}$$

Proposition 8.4 *Les deux sous-espaces $\mathcal{S}_n(\mathbb{R})$ et $\mathcal{A}_n(\mathbb{R})$ sont supplémentaires dans $\mathcal{M}_n(\mathbb{R})$:*

$$\mathcal{S}_n(\mathbb{R}) \oplus \mathcal{A}_n(\mathbb{R}) = \mathcal{M}_n(\mathbb{R}).$$

De plus, les dimensions des $\mathbb{R}$-espaces vectoriels $\mathcal{S}_n(\mathbb{R})$ et $\mathcal{A}_n(\mathbb{R})$ sont :

$$\dim_{\mathbb{R}}\left(\mathcal{S}_n(\mathbb{R})\right) = \tfrac{1}{2}n(n+1) \quad \textit{et} \quad \dim_{\mathbb{R}}\left(\mathcal{A}_n(\mathbb{R})\right) = \tfrac{1}{2}n(n-1).$$

Remarque Tous les coefficients diagonaux d'une matrice antisymétrique sont nuls. ◇

En mécanique des milieux continus, le gradient des vitesses d'un fluide est représenté en chaque point par une matrice $\nabla\mathbf{v} \in \mathcal{M}_3(\mathbb{R})$. La partie symétrique de cette matrice, c'est-à-dire la matrice $\mathrm{D} = \frac{1}{2}(\nabla\mathbf{v} + {}^t\nabla\mathbf{v})$, est appelée la matrice des déformations et la partie antisymétrique $\mathrm{W} = \frac{1}{2}(\nabla\mathbf{v} - {}^t\nabla\mathbf{v})$ est appelée la matrice de vorticité. Ces deux matrices jouent des rôles très importants dans la descriptions des écoulements d'un fluide.

En vertu de la proposition 8.3 toutes les propriétés des matrices symétriques ou antisymétriques réelles pourront être interprétées comme propriétés des

[20] Voir le *Cours de première année*, p. 401.

[21] Voir par exemple le *Cours de première année* p. 402.

[22] Voir la définition 8.16 p. 339 du *Cours de première année*.

formes bilinéaires symétriques ou antisymétriques. La proposition suivante en est l'illustration.

PROPOSITION 8.5 *Soit E un $\mathbb{R}$-espace vectoriel. Les deux sous-espaces $\mathcal{L}_2^s(E\times E,\mathbb{R})$ et $\mathcal{L}_2^a(E\times E,\mathbb{R})$ sont supplémentaires dans $\mathcal{L}_2(E\times E,\mathbb{R})$:*

$$\mathcal{L}_2^s(E\times E,\mathbb{R})\oplus\mathcal{L}_2^a(E\times E,\mathbb{R})=\mathcal{L}_2(E\times E,\mathbb{R}).$$

Dans le cas où E est un $\mathbb{R}$-espace vectoriel de dimension finie $n\in\mathbb{N}^*$, on a

$$\dim_{\mathbb{R}}\left(\mathcal{L}_2^s(E\times E,\mathbb{R})\right)=\tfrac{1}{2}n(n+1)\quad\text{et}\quad\dim_{\mathbb{R}}\left(\mathcal{L}_2^a(E\times E,\mathbb{R})\right)=\tfrac{1}{2}n(n-1).$$

La proposition 8.5 n'est pas un corollaire de la proposition 8.4. Elle ne peut en effet pas s'en déduire car elle couvre aussi le cas où E n'est pas de dimension finie. Par contre, la démonstration reste identique à celle qui est généralement utilisée pour prouver la proposition 8.4 (voir par exemple le *Cours de première année* p. 402). En particulier, on montre que toute forme bilinéaire φ sur un $\mathbb{R}$-espace vectoriel E se décompose de la façon suivante (on pourra comparer avec la relation (8)) :

$$\forall(\mathbf{x},\mathbf{y})\in E^2\quad\varphi(\mathbf{x},\mathbf{y})=\underbrace{\tfrac{1}{2}\big(\varphi(\mathbf{x},\mathbf{y})+\varphi(\mathbf{y},\mathbf{x})\big)}_{=\varphi_s(\mathbf{x},\mathbf{y})}+\underbrace{\tfrac{1}{2}\big(\varphi(\mathbf{x},\mathbf{y})-\varphi(\mathbf{y},\mathbf{x})\big)}_{=\varphi_a(\mathbf{x},\mathbf{y})}\qquad(9)$$

où l'application $\varphi_s:(\mathbf{x},\mathbf{y})\in E^2\mapsto\frac{1}{2}\big(\varphi(\mathbf{x},\mathbf{y})+\varphi(\mathbf{y},\mathbf{x})\big)$ est une forme bilinéaire symétrique et où $\varphi_a:(\mathbf{x},\mathbf{y})\in E^2\mapsto\frac{1}{2}\big(\varphi(\mathbf{x},\mathbf{y})-\varphi(\mathbf{y},\mathbf{x})\big)$ est une forme bilinéaire antisymétrique.

Remarque Une forme bilinéaire φ_a antisymétrique sur E vérifie

$$\forall\mathbf{x}\in E\quad\varphi_a(\mathbf{x},\mathbf{x})=0$$

car pour tout $\mathbf{x}\in E$ on a $\varphi_a(\mathbf{x},\mathbf{x})=-\varphi_a(\mathbf{x},\mathbf{x})$. ◇

8.1.5 Formes quadratiques

DÉFINITION 8.6 (Forme quadratique)
Une application q d'un espace vectoriel E dans $\mathbb{R}$ est appelée forme quadratique *s'il existe une forme bilinéaire φ sur E telle que*

$$\forall\mathbf{x}\in E\quad q(\mathbf{x})=\varphi(\mathbf{x},\mathbf{x}).$$

Remarque Soient H un sous-espace vectoriel d'un $\mathbb{R}$-espace vectoriel E et q une forme quadratique sur E. L'application $q|_H:x\in H\mapsto q(x)$ est une forme quadratique sur H. On dira que c'est la *restriction de la forme quadratique q* au sous-espace vectoriel H. ◇

Par définition, toute forme quadratique q est issue d'une forme bilinéaire φ. Cette forme bilinéaire n'a aucune raison d'être unique ni d'être symétrique. On a toutefois le résultat suivant.

PROPOSITION 8.6 (Formule de polarisation)
Soit E un $\mathbb{R}$-espace vectoriel. Pour toute forme quadratique q sur E, il existe une unique forme bilinéaire symétrique φ_0 pour laquelle on a

$$\forall \mathbf{x} \in E \quad q(\mathbf{x}) = \varphi_0(\mathbf{x}, \mathbf{x}).$$

Cette forme bilinéaire symétrique φ_0 s'appelle la forme polaire *de q. Elle est donnée par la relation*

$$\forall (\mathbf{x}, \mathbf{y}) \in E^2 \quad \varphi_0(\mathbf{x}, \mathbf{y}) = \tfrac{1}{2}\big(q(\mathbf{x}+\mathbf{y}) - q(\mathbf{x}) - q(\mathbf{y})\big) = \tfrac{1}{4}\big(q(\mathbf{x}+\mathbf{y}) - q(\mathbf{x}-\mathbf{y})\big),$$

appelée formule de polarisation.

Démonstration Considérons une forme quadratique q définie sur un $\mathbb{R}$-espace vectoriel E. Par définition, il existe une forme bilinéaire $\varphi : E \times E \to \mathbb{R}$ telle que

$$\forall \mathbf{x} \in E \quad q(\mathbf{x}) = \varphi(\mathbf{x}, \mathbf{x}). \tag{10}$$

$\trianglerighteq$ Dans un premier temps, on va montrer que l'on peut trouver une forme bilinéaire symétrique φ_0 vérifiant (10). D'après la proposition 8.5, on sait que la forme bilinéaire φ se décompose comme la somme d'une forme bilinéaire symétrique φ_s et d'un forme bilinéaire antisymétrique φ_a : $\varphi = \varphi_s + \varphi_a$. D'après la remarque de la p. 448, on sait que pour tout $\mathbf{x} \in E$ on a $\varphi_a(\mathbf{x}, \mathbf{x}) = 0$. Ainsi, la forme quadratique q vérifie

$$\forall \mathbf{x} \in E \quad q(\mathbf{x}) = \varphi_s(\mathbf{x}, \mathbf{x}).$$

L'application φ_s est une forme bilinéaire symétrique : on pose $\varphi_0 = \varphi_s$.

$\trianglerighteq$ On va maintenant prouver que toute forme bilinéaire symétrique satisfaisant l'égalité (10) est donnée par la formule de polarisation ; cela permettra de prouver que φ_0 est unique. Considérons une forme bilinéaire symétrique φ satisfaisant la relation (10). En utilisant la bilinéarité de φ, pour $(\mathbf{x}, \mathbf{y}) \in E^2$, on a

$$q(\mathbf{x}+\mathbf{y}) = \varphi(\mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y}) = \varphi(\mathbf{x}, \mathbf{x}) + \varphi(\mathbf{x}, \mathbf{y}) + \varphi(\mathbf{y}, \mathbf{x}) + \varphi(\mathbf{y}, \mathbf{y}).$$

Comme φ est supposée symétrique, cette relation s'écrit encore

$$q(\mathbf{x}+\mathbf{y}) = q(\mathbf{x}) + 2\varphi(\mathbf{x}, \mathbf{y}) + q(\mathbf{y}). \tag{11}$$

On en déduit que $\varphi(\mathbf{x}, \mathbf{y}) = \frac{1}{2}\big(q(\mathbf{x}+\mathbf{y}) - q(\mathbf{x}) - q(\mathbf{y})\big)$, ce qui correspond à la première formule de polarisation. On peut donc affirmer que φ est unique.

$\trianglerighteq$ Pour établir la seconde formule de polarisation, on utilise la relation (11) puis cette même relation en substituant $-\mathbf{y}$ à $\mathbf{y}$. Puisque $q(-\mathbf{y}) = \varphi(-\mathbf{y}, -\mathbf{y}) = \varphi(\mathbf{y}, \mathbf{y}) = q(\mathbf{y})$, on en déduit que :

$$\begin{aligned} & q(\mathbf{x}+\mathbf{y}) = q(\mathbf{x}) + 2\varphi(\mathbf{x},\mathbf{y}) + q(\mathbf{y}), \\ \text{et} \quad & q(\mathbf{x}-\mathbf{y}) = q(\mathbf{x}) - 2\varphi(\mathbf{x},\mathbf{y}) + q(\mathbf{y}). \end{aligned}$$

En ajoutant membre à membre ces deux relations, on a la seconde formule de polarisation $\varphi(\mathbf{x},\mathbf{y}) = \frac{1}{4}\big(q(\mathbf{x}+\mathbf{y}) - q(\mathbf{x}-\mathbf{y})\big)$. □

La façon la plus immédiate d'obtenir une forme quadratique est d'utiliser une forme bilinéaire φ et de construire l'application $\mathbf{x} \in E \mapsto \varphi(\mathbf{x},\mathbf{x})$.

Exemples

1. L'application $q : \mathbb{R}^3 \to \mathbb{R}$, définie par

$$q(\mathbf{x}) = x_1^2 + x_2^2 + x_3^2$$

est une forme quadratique sur $E = \mathbb{R}^3$. Elle est construite à partir de la forme bilinéaire sur $\mathbb{R}^3$ définie par

$$\varphi(\mathbf{x},\mathbf{y}) = x_1\ y_1 + x_2\ y_2 + x_3\ y_3.$$

Puisque cette forme bilinéaire est symétrique, on déduit de la proposition 8.6 que φ est la forme polaire de q.

2. La forme bilinéaire $\varphi : (f,g) \in \mathcal{C}([0,1],\mathbb{R})^2 \mapsto \int_0^1 f(t)g(t)\ \mathrm{d}t$ permet de construire la forme quadratique sur $\mathcal{C}([0,1],\mathbb{R})$ définie par

$$q(f) = \int_0^1 f(t)^2\ \mathrm{d}t.$$

3. Sur l'espace vectoriel $\mathcal{M}_n(\mathbb{R})$, l'application $q : \mathcal{M}_n(\mathbb{R}) \to \mathbb{R}$ définie par

$$q(\mathrm{A}) = \mathrm{Tr}({}^t\mathrm{AA})$$

est une forme quadratique. Sa forme polaire est l'application bilinéaire[23]

$$\varphi : (\mathrm{A},\mathrm{B}) \in \mathcal{M}_n(\mathbb{R}) \times \mathcal{M}_n(\mathbb{R}) \longmapsto \mathrm{Tr}({}^t\mathrm{AB}) \in \mathbb{R}.$$

[23] En effet, il a été prouvé dans l'exemple p. 446 que cette application bilinéaire était symétrique.

Remarques

1. Un fois que l'on connaît la forme polaire φ_0 associée à une forme quadratique q, il est intéressant de réécrire la formule de polarisation sous la forme suivante :

$$\forall (\mathbf{x}, \mathbf{y}) \in E^2 \quad q(\mathbf{x}+\mathbf{y}) = q(\mathbf{x}) + 2\varphi_0(\mathbf{x}, \mathbf{y}) + q(\mathbf{y}).$$

Cela donne le comportement d'une forme quadratique vis-à-vis de l'addition dans E (en particulier cela prouve qu'une forme quadratique n'est en général pas linéaire).

2. Soient q une forme quadratique sur un $\mathbb{R}$-espace vectoriel E et φ la forme bilinéaire la définissant (voir la définition 8.6, φ peut par exemple être la forme polaire de q). Pour tout $\lambda \in \mathbb{R}$ et tout $\mathbf{x} \in E$, on a

$$q(\lambda \mathbf{x}) = \varphi(\lambda \mathbf{x}, \lambda \mathbf{x}) = \lambda^2 \varphi(\mathbf{x}, \mathbf{x}) = \lambda^2 q(\mathbf{x}).$$

Cela donne le comportement d'une forme quadratique vis-à-vis de la multiplication par un scalaire. ◇

Pour montrer qu'une application q définie sur un $\mathbb{R}$ espace vectoriel E et à valeurs dans $\mathbb{R}$ est une forme quadratique, on peut utiliser la caractérisation suivante.

PROPOSITION 8.7 *Une application q d'un espace vectoriel E dans $\mathbb{R}$ est une forme quadratique si et seulement si*

1. $\forall \mathbf{x} \in E \quad \forall \lambda \in \mathbb{R} \quad q(\lambda \mathbf{x}) = \lambda^2 q(\mathbf{x})$,
2. *l'application* $(\mathbf{x}, \mathbf{y}) \in E^2 \mapsto \frac{1}{2}\big(q(\mathbf{x}+\mathbf{y}) - q(\mathbf{x}) - q(\mathbf{y})\big) \in \mathbb{R}$ *est une forme bilinéaire.*

Démonstration $\trianglerighteq$ Supposons que l'application q de E dans $\mathbb{R}$ soit une forme quadratique. D'après la remarque qui précède, elle satisfait la première condition. L'application donnée par la condition *2* correspond à la forme polaire de la forme quadratique q, c'est donc une forme bilinéaire.

$\trianglerighteq$ Réciproquement, supposons que l'application q vérifie les conditions *1* et *2* et considérons l'application $\varphi : (\mathbf{x}, \mathbf{y}) \in E^2 \mapsto \frac{1}{2}\big(q(\mathbf{x}+\mathbf{y}) - q(\mathbf{x}) - q(\mathbf{y})\big) \in \mathbb{R}$. D'après la condition *2*, l'application φ est une forme bilinéaire. De plus, d'après la condition *1*, on a

$$\forall \mathbf{x} \in E \quad \varphi(\mathbf{x}, \mathbf{x}) = \tfrac{1}{2}\big(q(2\mathbf{x}) - q(\mathbf{x}) - q(\mathbf{x})\big) = \tfrac{1}{2}\big(4q(\mathbf{x}) - 2q(\mathbf{x})\big) = q(\mathbf{x}).$$

Cela implique, d'après la définition 8.6, que q est une forme quadratique. □

Exemple Montrons en utilisant la proposition 8.7 que l'application q définie sur le $\mathbb{R}$-espace vectoriel $\mathbb{R}_2[X]$ des polynômes de degré 2 à coefficients réels par $q : P \in \mathbb{R}_2[X] \mapsto P(0)P(1)$ est une forme quadratique.

D'une part, pour tout $P \in \mathbb{R}_2[X]$ et tout $\lambda \in \mathbb{R}$ on a

$$q(\lambda P) = \lambda P(0)\lambda P(1) = \lambda^2 q(P).$$

La condition *1* de la proposition 8.7 est satisfaite.

Montrons d'autre part que l'application

$$\varphi : (P, Q) \in \mathbb{R}_2[X]^2 \longmapsto \tfrac{1}{2}\big(q(P+Q) - q(P) - q(Q)\big)$$

est une application bilinéaire sur $\mathbb{R}_2[X]$. Pour tout $(P, Q) \in \mathbb{R}_2[X]^2$ on a $\varphi(P, Q) = \frac{1}{2}\big(P(0)Q(1) + P(1)Q(0)\big)$. Remarquons tout d'abord que pour tout $(P, Q) \in \mathbb{R}_2[X]^2$ on a $\varphi(P, Q) = \varphi(Q, P)$, autrement dit, l'application φ est symétrique. Ainsi, pour prouver que φ est bilinéaire, il suffit de prouver qu'elle est linéaire par rapport à sa première variable[(24)]. Soient $(P, Q, R) \in \mathbb{R}_2[X]^3$ et $(\lambda, \mu) \in \mathbb{R}^2$. On a

$$\begin{aligned}
\varphi(\lambda P + \mu Q, R) &= \tfrac{1}{2}\big((\lambda P + \mu Q)(0)\ R(1) + (\lambda P + \mu Q)(1)\ R(0)\big) \\
&= \tfrac{1}{2}\big(\lambda P(0)\, R(1) + \mu Q(0)\, R(1) + \lambda P(1)\, R(0) + \mu Q(1)\, R(0)\big) \\
&= \tfrac{1}{2}\lambda\,\big(P(0)\, R(1) + P(1)\, R(0)\big) + \tfrac{1}{2}\mu\,\big(Q(0)\, R(1) + Q(1)\, R(0)\big) \\
&= \lambda\varphi(P, R) + \mu\varphi(Q, R),
\end{aligned}$$

ce qui montre la linéarité de φ par rapport à la première variable.

DÉFINITION 8.7 *Soient E un $\mathbb{R}$-espace vectoriel de dimension finie, $\mathcal{B}$ une base de E et q une forme quadratique sur E.*

✗ *On appelle matrice de q dans la base $\mathcal{B}$ la matrice de la forme polaire de q dans la base $\mathcal{B}$.*

✗ *Le* rang de la forme quadratique q *est le rang de la matrice de q.*

En désignant par X la matrice colonne dont les éléments sont les composantes d'un vecteur $\mathbf{x} \in E$ dans une base donnée $\mathcal{B}$ de l'espace vectoriel E, par A la matrice d'une forme quadratique q sur E dans cette même base $\mathcal{B}$ et par φ_0 la forme polaire de q, l'égalité $q(\mathbf{x}) = \varphi_0(\mathbf{x}, \mathbf{x})$ s'écrit matriciellement

$$q(\mathbf{x}) = {}^tXAX.$$

(24) Cet argument est souvent utilisé pour montrer qu'une application de deux variables qui est symétrique est une application bilinéaire. En effet, si $\varphi : E \times E \mapsto \mathbb{R}$ est symétrique et linéaire par rapport à sa première variable alors elle sera linéaire par rapport à sa deuxième variable et donc bilinéaire.

Remarques

1. Soit E un $\mathbb{R}$-espace vectoriel de dimension n muni d'une base $\mathcal{B} = (\mathbf{e}_1, ..., \mathbf{e}_n)$. Considérons une forme quadratique q sur E, dont la forme polaire est notée φ_0 et la matrice associée A. Si $\mathbf{x} \in E$ est donné dans la base $\mathcal{B}$ par $\mathbf{x} = \sum_{i=1}^{n} x_i\, \mathbf{e}_i$ alors on sait que[25]

$$q(\mathbf{x}) = \varphi_0(\mathbf{x}, \mathbf{x}) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}\, x_i\, x_j.$$

Dans la somme sur l'indice noté « j », on peut distinguer les cas $j = i$, $j > i$ et $j < i$. On remarque que puisque φ_0 est la forme polaire de q, sa matrice A est symétrique et les coefficients a_{ij} et a_{ji} sont égaux. On obtient :

$$q(\mathbf{x}) = \sum_{i=1}^{n} a_{ii}\, x_i^2 + 2 \sum_{i=1}^{n} \sum_{j=i+1}^{n} a_{ij}\, x_i\, x_j.$$

2. Inversement, si $\mathrm{A} \in \mathcal{S}_n(\mathbb{R})$ est une matrice symétrique, l'expression

$$q(\mathbf{x}) = \sum_{i=1}^{n} a_{ii}\, x_i^2 + 2 \sum_{i=1}^{n} \sum_{j=i+1}^{n} a_{ij}\, x_i\, x_j$$

définit la forme quadratique $q : \mathbf{x} \mapsto q(\mathbf{x})$ associée à la forme bilinéaire symétrique donnée par

$$\varphi_0(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} a_{ii}\, x_i\, y_i + \sum_{i=1}^{n} \sum_{j=i+1}^{n} a_{ij}\, (x_i\, y_j + x_j\, y_i).$$

Matriciellement, cela permet d'associer à toute matrice symétrique $\mathrm{A} \in \mathcal{S}_n(\mathbb{R})$ une forme quadratique q sur $\mathbb{R}^n$ telle que $q(\mathbf{x}) = {}^t\mathrm{XAX}$, le vecteur X correspondant à la matrice colonne dont les éléments sont les composantes du vecteur $\mathbf{x}$ dans la base canonique de $\mathbb{R}^n$. La forme polaire φ_0 associée à q est déterminée matriciellement par $\varphi_0(\mathbf{x}, \mathbf{y}) = {}^t\mathrm{XAY}$. ◇

Exemple L'application $q : \mathbb{R}^3 \to \mathbb{R}$ définie par

$$q(\mathbf{x}) = 3x_1^2 + x_2^2 + 2x_1x_2 - 3x_1x_3$$

est une forme quadratique sur $E = \mathbb{R}^3$. La matrice de q dans la base canonique de $\mathbb{R}^3$ est

$$M = \begin{pmatrix} 3 & 1 & -3/2 \\ 1 & 1 & 0 \\ -3/2 & 0 & 0 \end{pmatrix}.$$

[25] Voir l'expression d'une forme bilinéaire dans une base, donnée par la relation (2) p. 440.

8.1.6 Orthogonalité

Si deux droites $\mathcal{D}_1$ et $\mathcal{D}_2$ sont dessinées sur une feuille, on connaît des critères permettant de dire si elles sont ou non perpendiculaires. Par exemple, en considérant les vecteurs directeurs

$$\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \quad \text{et} \quad \mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

des droites $\mathcal{D}_1$ et $\mathcal{D}_2$, il suffit d'évaluer la quantité $u_1v_1 + u_2v_2$ appelée « produit scalaire » des vecteurs $\mathbf{u}$ et $\mathbf{v}$. Si cette valeur est nulle alors les droites sont perpendiculaires, sinon elles ne sont pas perpendiculaires. Comme on va le voir, cette notion familière d'orthogonalité est liée à un choix d'application bilinéaire sur le plan et peut être généralisée à des espaces de dimension quelconque.

DÉFINITION 8.8 *Soient E un $\mathbb{R}$-espace vectoriel, q une forme quadratique sur E et φ_0 la forme polaire de q. Deux vecteurs $\mathbf{x}$ et $\mathbf{y}$ de E sont dits* orthogonaux *selon q (ou selon φ_0) si $\varphi_0(\mathbf{x}, \mathbf{y}) = 0$.*

Remarque Une forme polaire étant symétrique, si x et y sont *orthogonaux* selon q alors y et x sont aussi *orthogonaux* selon q. ◇

Exemples

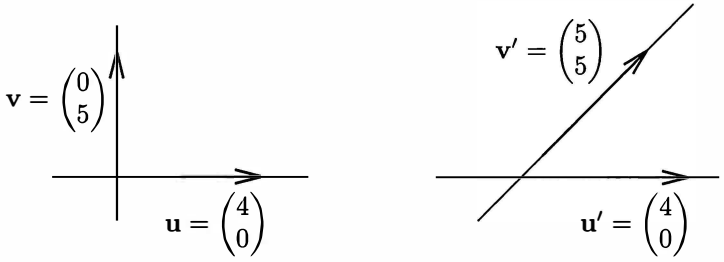
1. Dans l'espace vectoriel $\mathbb{R}^2$, la forme quadratique $q : \mathbf{x} \in \mathbb{R}^2 \mapsto x_1^2 + x_2^2$ fournit la notion d'orthogonalité usuelle dans le plan. La forme polaire de q est l'application $\varphi_0 : (\mathbf{x}, \mathbf{y}) \in (\mathbb{R}^2)^2 \mapsto x_1\, y_1 + x_2\, y_2$ de sorte que deux vecteurs $\mathbf{u}$ et $\mathbf{v}$ de composantes (u_1, u_2) et (v_1, v_2) sont orthogonaux selon cette forme quadratique q si et seulement si $u_1v_1 + v_1v_2 = 0$. Ainsi, selon la forme quadratique q, les vecteurs

$$\mathbf{u}' = \begin{pmatrix} 4 \\ 0 \end{pmatrix} \quad \text{et} \quad \mathbf{v}' = \begin{pmatrix} 5 \\ 5 \end{pmatrix}$$

ne sont pas orthogonaux car $\varphi_0(\mathbf{u}', \mathbf{v}') = 20 \neq 0$ mais les vecteurs

$$\mathbf{u} = \begin{pmatrix} 4 \\ 0 \end{pmatrix} \quad \text{et} \quad \mathbf{v} = \begin{pmatrix} 0 \\ 5 \end{pmatrix}$$

le sont car $\varphi_0(\mathbf{u}, \mathbf{v}) = 0$.



2. Dans l'espace vectoriel $\mathbb{R}^2$, la forme quadratique $q : \mathbf{x} \in \mathbb{R}^2 \mapsto x_1^2 - 2x_1 x_2$ a pour forme polaire $\varphi_0 : (\mathbf{x}, \mathbf{y}) \in (\mathbb{R}^2)^2 \mapsto x_1 y_1 - x_1 y_2 - x_2 y_1$. Selon cette forme quadratique q, les vecteurs

$$\mathbf{u}' = \begin{pmatrix} 4 \\ 0 \end{pmatrix} \quad \text{et} \quad \mathbf{v}' = \begin{pmatrix} 5 \\ 5 \end{pmatrix}$$

sont orthogonaux car $\varphi_0(\mathbf{u}', \mathbf{v}') = 0$, mais les vecteurs

$$\mathbf{u} = \begin{pmatrix} 4 \\ 0 \end{pmatrix} \quad \text{et} \quad \mathbf{v} = \begin{pmatrix} 0 \\ 5 \end{pmatrix}$$

ne le sont pas car $\varphi_0(\mathbf{u}, \mathbf{v}) = -20 \neq 0$.

Définition 8.9 (Orthogonal)
Soient A et B deux sous-ensembles d'un $\mathbb{R}$-espace vectoriel E, q une forme quadratique sur E et φ_0 la forme polaire associée à q.

✗ *On appelle* orthogonal *de A selon q (ou selon φ_0) l'ensemble*

$$A^\perp = \{\mathbf{x} \in E \ ; \ \forall \mathbf{y} \in A \quad \varphi_0(\mathbf{x}, \mathbf{y}) = 0\}.$$

✗ *Les deux ensembles A et B sont dit orthogonaux selon q (ou selon φ_0) et on note $A \perp B$ si*

$$\forall \mathbf{x} \in A \quad \forall \mathbf{y} \in B \quad \varphi_0(\mathbf{x}, \mathbf{y}) = 0.$$

On notera que $\{\mathbf{0}\}^\perp = E$ car pour tout $\mathbf{x} \in E$ on a $\varphi_0(\mathbf{x}, \mathbf{0}) = 0$. Pour la même raison, l'ensemble $\{\mathbf{0}\}$ est orthogonal à tous les autres sous-ensembles de E.

De nombreuses formes quadratiques pouvant être définies sur un espace vectoriel E, il s'ensuit qu'il existe plusieurs relations d'orthogonalité sur un même espace. La notation $\perp$ ne précise pas la forme quadratique pour laquelle il y a orthogonalité. Lorsqu'une ambiguïté est possible, on notera $\perp_{\varphi_0}$ ou $\perp_q$ pour faire référence à la forme quadratique ou forme bilinéaire symétrique considérée.

Proposition 8.8 *Soient E un $\mathbb{R}$-espace vectoriel, q une forme quadratique sur E et A un sous-ensemble de E.*

✗ *L'ensemble $A^\perp$ est un sous-espace vectoriel de E.*

✗ *On a $A \subset A^{\perp\perp}$.*

Démonstration ⊵ On a $A^\perp \subset E$; pour montrer que $A^\perp$ est un sous-espace vectoriel de E, il faut[(26)] montrer que $\mathbf{0} \in A^\perp$ (ce qui est immédiat selon la définition de l'orthogonal $A^\perp$), et que

$$\forall (\mathbf{x}, \mathbf{y}) \in (A^\perp)^2 \quad \forall (\lambda, \mu) \in \mathbb{R}^2 \quad \lambda \mathbf{x} + \mu \mathbf{y} \in A^\perp.$$

[(26)] Voir le *Cours de première année* p. 306.

Soit $(\mathbf{x},\mathbf{y}) \in (A^{\perp})^2$ et $(\lambda,\mu) \in \mathbb{R}^2$; afin de prouver que $\lambda\mathbf{x}+\mu\mathbf{y} \in A^{\perp}$, considérons un élément $\mathbf{z} \in A$ et évaluons la quantité $\varphi_0(\lambda\mathbf{x}+\mu\mathbf{y},\mathbf{z})$. En vertu de la bilinéarité de φ_0, on a

$$\varphi_0(\lambda\mathbf{x}+\mu\mathbf{y},\mathbf{z}) = \lambda\varphi_0(\mathbf{x},\mathbf{z}) + \mu\varphi_0(\mathbf{y},\mathbf{z}).$$

Les éléments $\mathbf{x}$ et $\mathbf{y}$ étant dans $A^{\perp}$, ils satisfont les relations $\varphi_0(\mathbf{x},\mathbf{z}) = 0$ et $\varphi_0(\mathbf{y},\mathbf{z}) = 0$ pour tout $\mathbf{z} \in A$. On en déduit que $\varphi_0(\lambda\mathbf{x}+\mu\mathbf{y},\mathbf{z}) = 0$ pour tout $z \in A$, ce qui prouve que $\lambda\mathbf{x}+\mu\mathbf{y} \in A^{\perp}$.

$\trianglerighteq$ Dire qu'un élément $\mathbf{z} \in E$ appartient à $A^{\perp\perp}$ signifie que

$$\forall \mathbf{y} \in A^{\perp} \quad \varphi_0(\mathbf{z},\mathbf{y}) = 0.$$

Soit $\mathbf{x} \in A$; pour tout $\mathbf{y} \in A^{\perp}$ par définition de $A^{\perp}$ on a $\varphi_0(\mathbf{x},\mathbf{y}) = 0$. On en déduit que $\mathbf{x} \in A^{\perp\perp}$ ce qui prouve que $A \subset A^{\perp\perp}$. □

Exemple Soient $E = \mathbb{R}_2[X]$ et $q : P \in \mathbb{R}_2[X] \mapsto P(0)P(1) \in \mathbb{R}$. On a vu p. 451 que q est une forme quadratique sur $\mathbb{R}_2[X]$ dont la forme polaire $\varphi_0 : \mathbb{R}_2[X] \times \mathbb{R}_2[X] \to \mathbb{R}$ est définie par

$$\varphi_0(P,Q) = \tfrac{1}{2}\big(P(0)Q(1) + P(1)Q(0)\big).$$

Considérons comme sous-ensemble de $\mathbb{R}_2[X]$ le singleton $A = \{P\}$ où P désigne le polynôme $P = X^2+X+1$. Par définition, un polynôme $Q \in \mathbb{R}_2[X]$ appartient à $A^{\perp}$ si et seulement si $\varphi_0(P,Q) = 0$. On a donc

$$Q \in A^{\perp} \iff \tfrac{1}{2}\big(P(0)Q(1) + P(1)Q(0)\big) = 0.$$

Compte tenu de l'expression du polynôme P, on obtient

$$Q \in A^{\perp} \iff Q(1) + 3Q(0) = 0.$$

En écrivant Q dans la base canonique de $\mathbb{R}_2[X]$ sous la forme $Q = aX^2+bX+c$ avec a, b et c trois nombres réels, on a

$$Q \in A^{\perp} \iff a+b+4c = 0.$$

L'ensemble $A^{\perp}$ est par conséquent constitué des polynômes Q de la forme $Q = (-b-4c)X^2+bX+c$; c'est-à-dire de la forme $Q = b(X-X^2)+c(1-4X^2)$, b et c étant deux réels quelconques. On en conclut que

$$A^{\perp} = \mathrm{Vect}\{X - X^2, 1 - 4X^2\}.$$

De la même manière, un polynôme $R = \alpha X^2 + \beta X + \gamma \in \mathbb{R}_2[X]$ appartiendra à $A^{\perp\perp}$ si et seulement si[27]

$$\begin{cases} \varphi_0(R, X - X^2) = 0 \\ \varphi_0(R, 1 - 4X^2) = 0 \end{cases}$$

[27] D'après la linéarité de l'application φ_0 par rapport à la seconde variable, on a $\varphi_0(R,S) = 0$ pour tous les polynômes S de l'espace vectoriel engendré par deux polynômes S_1 et S_2 si et seulement si $\varphi_0(R,S_1) = 0$ et $\varphi_0(R,S_2) = 0$.

autrement dit, si et seulement si $\alpha + \beta - 2\gamma = 0$. On en déduit que $R \in A^{\perp\perp}$ si et seulement si il est de la forme $R = \alpha(X^2 + \frac{1}{2}) + \beta(X + \frac{1}{2})$, α et β étant deux réels quelconques. On a donc[28]

$$A^{\perp\perp} = \text{Vect}\{X^2 + \tfrac{1}{2}, X + \tfrac{1}{2}\} = \text{Vect}\{X^2 - X, X^2 + X + 1\}.$$

Avec cet exemple, on peut noter qu'en général l'inclusion $A \subset A^{\perp\perp}$ est stricte.

DÉFINITION 8.10 (Noyau et cône isotrope)
Soient E un $\mathbb{R}$-espace vectoriel et q une forme quadratique sur E.

✗ *On appelle* noyau *de la forme quadratique q, l'ensemble*

$$N(q) = E^{\perp}.$$

✗ *On appelle* cône isotrope *de q, l'ensemble*

$$C(q) = \{\mathbf{x} \in E \ ; \ q(\mathbf{x}) = 0\}.$$

Les éléments du cône isotrope sont appelés des vecteurs isotropes.

Remarques

1. On a toujours $\mathbf{0} \in N(q)$ et $\mathbf{0} \in C(q)$.

2. Il ne faut pas confondre $N(q)$ et $C(q)$. L'ensemble $N(q)$ est l'ensemble des vecteurs qui sont orthogonaux à tous les vecteurs de E alors que $C(q)$ est l'ensemble des vecteurs qui sont orthogonaux à eux-mêmes. Ainsi, on a l'inclusion $N(q) \subset C(q)$ et cette inclusion est en général stricte :$N(q) \neq C(q)$. ◇

Exemple Considérons la forme quadratique définie sur $\mathbb{R}^2$ par

$$q : \mathbf{x} \in \mathbb{R}^2 \mapsto x_1^2 - x_2^2.$$

La forme polaire de q est l'application $\varphi_0 : (\mathbf{x}, \mathbf{y}) \in (\mathbb{R}^2)^2 \mapsto x_1y_1 - x_2y_2$ et le noyau $N(q)$ de la forme quadratique q correspond à l'ensemble des éléments $\mathbf{x} \in \mathbb{R}^2$ tels que $x_1y_1 - x_2y_2 = 0$ pour tout $\mathbf{y} \in \mathbb{R}^2$. En choisissant successivement $\mathbf{y} = (1, 0)$ et $\mathbf{y} = (0, 1)$, on établit que le noyau de q est $N(q) = \{\mathbf{0}\}$.

Le cône isotrope de cette forme quadratique q n'est pas égal à $\{\mathbf{0}\}$ puisque l'élément $\mathbf{x} = (1, 1)$ appartient à $C(q)$ (on vérifiera que $q(1, 1) = 0$). Plus précisément, le cône isotrope $C(q)$ est

$$C(q) = \{\mathbf{x} \in \mathbb{R}^2 \quad ; \ x_1 = x_2 \ \text{ ou } \ x_1 = -x_2\},$$

c'est-à-dire $C(q)$ est la réunion de deux droites du plan qui sont la première et la seconde bissectrice.

[28] La dernière égalité provient du fait que l'espace vectoriel engendré par deux vecteurs $\mathbf{x}$ et $\mathbf{y}$ est le même que l'espace vectoriel engendré par $\mathbf{x} + \mathbf{y}$ et $\mathbf{x} - \mathbf{y}$

Intéressons-nous au cas où E est un espace vectoriel de dimension finie n. Étant donné une base $\mathcal{B} = (\mathbf{e}_1, ..., \mathbf{e}_n)$ de E, à toute forme quadratique q sur E est associée une matrice $\mathrm{M} \in \mathcal{S}_n(\mathbb{R})$. Dire qu'un élément $\mathbf{x} \in E$, représenté par la matrice colonne $\mathrm{X} \in \mathbb{R}^n$ dans la base $\mathcal{B}$, appartient au noyau de la forme quadratique q signifie que

$$\forall \mathrm{Y} \in \mathbb{R}^n \quad {}^t\mathrm{YMX} = 0.$$

En choisissant pour Y le vecteur $\mathrm{Y} = \mathrm{MX}$, on en déduit[29] que $\mathbf{x} \in N(q)$ si et seulement si $\mathrm{MX} = 0$. Ainsi, en dimension finie, le noyau d'une forme quadratique correspond exactement au noyau de l'endomorphisme ayant la même matrice que la forme quadratique dans une même base donnée. En particulier, le théorème du rang[30] s'écrit dans le cadre de l'algèbre bilinéaire sous la forme suivante : « La dimension de E est la somme du rang de q et de la dimension du noyau de q ». Autrement dit, en dimension finie on a

$$\boxed{\dim(E) = \mathrm{rg}(q) + \dim(N(q)).}$$

8.1.7 Bases orthogonales

Commençons par quelques rappels. Étant donnés un espace vectoriel E et un ensemble I, on appelle[31] *famille de vecteurs* indexée par I toute application

$$i \in I \longmapsto \mathbf{v}_i \in E.$$

On la note $(\mathbf{v}_i)_{i \in I}$. L'ensemble de départ I est appelé *ensemble des indices* de la famille. Lorsque I est un ensemble fini (resp. infini), on dit que $(\mathbf{v}_i)_{i \in I}$ est une *famille finie* (resp. *infinie*).

Définition 8.11 *Soit $\mathcal{F}$ une famille de vecteurs[32] d'un $\mathbb{R}$-espace vectoriel E et q une forme quadratique sur E de forme polaire associée φ_0.*

✗ *La famille $\mathcal{F}$ est dite q-orthogonale si*

$$\forall (\mathbf{f}, \mathbf{f}') \in \mathcal{F}^2 \quad \big(\mathbf{f} \neq \mathbf{f}' \implies \varphi_0(\mathbf{f}, \mathbf{f}') = 0\big).$$

✗ *La famille $\mathcal{F}$ est dite q-orthonormale si*

la famille $\mathcal{F}$ est q-orthogonale et $\forall \mathbf{f} \in \mathcal{F} \quad \varphi_0(\mathbf{f}, \mathbf{f}) = 1.$

---

[29] On remarquera que l'égalité ${}^t\mathrm{YY} = 0$ est équivalente à $y_1^2 + \cdots + y_n^2 = 0$ et donc équivalente à $\mathrm{Y} = 0$.

[30] Voir le *Cours de première année* p. 376.

[31] Voir le *Cours de première année* p. 303.

[32] Famille finie ou infinie.

Remarque Une famille à un seul élément non nul sera considérée comme q-orthogonale pour toute forme quadratique q. ◇

Exemples

1. Sur l'espace vectoriel $\mathcal{C}([0,1],\mathbb{R})$ des applications continues de $[0,1]$ dans $\mathbb{R}$, l'application $\varphi : \mathcal{C}([0,1],\mathbb{R}) \times \mathcal{C}([0,1],\mathbb{R}) \to \mathbb{R}$ définie par

$$\varphi(f,g) = \int_0^1 f(t)g(t)\ \mathrm{d}t$$

est une forme bilinéaire dont la forme quadratique associée est l'application $q : f \in \mathcal{C}([0,1],\mathbb{R}) \mapsto \int_0^1 f(t)^2\ \mathrm{d}t$, voir p. 439. Considérons la famille $(f_n)_{n\in\mathbb{N}^*}$ de $\mathcal{C}([0,1],\mathbb{R})$ indexée par l'ensemble des entiers naturels non nuls et définie par

$$\forall n \in \mathbb{N}^* \quad f_n : t \in [0,1] \longmapsto \cos(2\pi nt).$$

Pour tout $(n,m) \in (\mathbb{N}^*)^2$ on a

$$\begin{aligned}\varphi(f_n,f_m) &= \int_0^1 \cos(2\pi nt)\cos(2\pi mt)\ \mathrm{d}t \\ &= \frac{1}{2}\left(\int_0^1 \cos(2\pi(n+m)t)\ \mathrm{d}t + \int_0^1 \cos(2\pi(n-m)t)\ \mathrm{d}t\right).\end{aligned}$$

Puisque pour tout réel $\lambda \neq 0$ on a $\int_0^1 \cos(2\pi\lambda t)\ \mathrm{d}t = 0$, on en déduit que lorsque $n \neq m$ on a $\varphi(f_n, f_m) = 0$, ce qui montre que la famille $(f_n)_{n\in\mathbb{N}^*}$ est une famille orthogonale.

Cette famille n'est pas orthonormale puisque pour tout $n \in \mathbb{N}^*$ on a

$$\varphi(f_n,f_n) = \frac{1}{2}\left(\int_0^1 \cos(4\pi nt)\ \mathrm{d}t + \int_0^1 1\ \mathrm{d}t\right) = \frac{1}{2} \neq 1.$$

Par contre, en considérant la famille $(g_n)_{n\in\mathbb{N}^*}$ définie par $g_n = \sqrt{2}\, f_n$ alors on a d'une part, pour tout couple $(n,m) \in (\mathbb{N}^*)^2$ tel que $n \neq m$ la relation $\varphi(g_n,g_m) = 2\varphi(f_n,f_m) = 0$ et d'autre part, pour tout $n \in \mathbb{N}^*$, la relation $\varphi(g_n,g_n) = 2\varphi(f_n,f_n) = 1$. Ainsi, la famille $(g_n)_{n\in\mathbb{N}^*}$ est orthonormale.

2. Dans l'espace vectoriel $\mathbb{R}^4$, la forme de Lorentz, voir p. 439, est la forme bilinéaire définie par

$$\varphi : (\mathbf{x},\mathbf{y}) \in \mathbb{R}^4 \times \mathbb{R}^4 \longmapsto x_1\ y_1 + x_2\ y_2 + x_3\ y_3 - c^2 x_4\ y_4$$

où c désigne un paramètre réel. La base canonique de $\mathbb{R}^4$, c'est-à-dire la famille finie $\{\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3,\mathbf{e}_4\}$, est une famille orthogonale pour cette forme bilinéaire puisque pour tout $(i,j) \in \{1,2,3,4\}^2$ avec $i \neq j$ on a $\varphi(\mathbf{e}_i,\mathbf{e}_j) = 0$. Par contre, la famille $\{\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3,\mathbf{e}_4\}$ n'est pas orthonormale pour la forme de Lorentz puisque on a $\varphi(\mathbf{e}_4,\mathbf{e}_4) = -c^2 \neq 1$.

Rappelons[33] qu'une famille finie $(\mathbf{v}_i)_{i\in\{1,..,p\}}$ $(p \in \mathbb{N}^*)$ d'un espace vectoriel est dite *liée* s'il existe des réels α_1, ..., α_p, dont un au moins est non nul, tels que

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + ... + \alpha_p\mathbf{v}_p = 0.$$

Si une famille finie n'est pas liée, on dit que c'est une *famille libre* (ou que les vecteurs de la familles sont linéairement indépendants).

Ces définitions de familles liées et libres se généralisent aux familles infinies de la manière suivante : une famille infinie est *libre* si toutes ses sous-familles[34] finies sont libres, et une famille infinie est *liée* si elle n'est pas libre.

Concernant les familles q-orthonormales, on a la proposition suivante.

PROPOSITION 8.9 *Soient E un $\mathbb{R}$-espace vectoriel et q une forme quadratique sur E. Toute famille de vecteurs de E qui est q-orthonormale est une famille libre.*

Démonstration Soit $(\mathbf{v}_i)_{i\in I}$ une famille de vecteurs de E qui est q-orthogonale. Si cette famille est une famille infinie, pour montrer que c'est une famille libre, il faut montrer que toutes ses sous-familles finies sont libres. Ainsi, si on montre que toute famille *finie* de vecteurs de E qui est q-orthonormale est une famille libre alors la proposition sera démontrée aussi bien pour les familles finies que pour les familles infinies.

Considérons une famille finie $\mathcal{F} = \{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ $(p \in \mathbb{N}^*)$ q-orthonormale. Supposons qu'il existe des réels $\lambda_1, \ldots, \lambda_p$ tels que $\lambda_1\mathbf{v}_1 + \cdots + \lambda_p\mathbf{v}_p = 0$. Pour tout $k \in \{1, \ldots, p\}$, on a

$$\varphi_0(\lambda_1\mathbf{v}_1 + \cdots + \lambda_p\mathbf{v}_p, \mathbf{v}_k) = 0.$$

La bilinéarité de l'application φ_0 implique que

$$\lambda_1\varphi_0(\mathbf{v}_1, \mathbf{v}_k) + \cdots + \lambda_p\varphi_0(\mathbf{v}_p, \mathbf{v}_k) = 0.$$

Puisque la famille $\mathcal{F}$ est orthonormale, tous les termes $\varphi_0(\mathbf{v}_i, \mathbf{v}_k)$ sont nuls, excepté celui correspondant à $i = k$ qui vaut 1. On en déduit que $\lambda_k = 0$ pour tout $k \in \{1, \ldots, p\}$. Cela montre que tous les λ_k sont nuls : la famille $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ est libre. □

Une famille peut être q-orthogonale sans pour autant être libre. Par exemple la famille $\{(1,1),(2,2)\}$ de $\mathbb{R}^2$ est q-orthogonale pour la forme quadratique $q(\mathbf{x}) = x_1^2 - x_2^2$ mais cette famille n'est manifestement pas libre puisque l'un des vecteurs est le double de l'autre.

[33] Voir le *Cours de première année* p. 317.

[34] Une sous-famille d'une famille $(\mathbf{v}_i)_{i\in I}$ est une famille de la forme $(\mathbf{v}_i)_{i\in J}$ où $J \subset I$.

Cas de la dimension finie

Dans le cas où l'on considère un $\mathbb{R}$-espace vectoriel de dimension finie, certaines familles de vecteurs jouent un rôle particulier : ce sont les bases.

DÉFINITION 8.12 *Soient E un $\mathbb{R}$-espace vectoriel de dimension finie et q une forme quadratique sur E. On dit qu'une* base $\mathcal{B}$ de E est q-orthogonale *(resp. q-orthonormale) si la famille $\mathcal{B}$ est q-orthogonale (resp. q-orthonormale).*

Si on dispose d'une base $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ de E qui est q-orthonormale alors on peut facilement déterminer la matrice de la forme quadratique q dans cette base : par définition, cette matrice est celle de la forme polaire φ_0 associée à la forme quadratique q. Les coefficients de cette matrice sont donc les réels $\varphi_0(\mathbf{e}_i, \mathbf{e}_j)$. Puisque la famille $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ est q-orthonormale, tous les termes $\varphi_0(\mathbf{e}_i, \mathbf{e}_j)$ sont nuls sauf ceux correspondant à $i = j$ qui valent 1. La matrice de la forme quadratique est donc la matrice identité. On a démontré la proposition suivante.

PROPOSITION 8.10 *Soient E un $\mathbb{R}$-espace vectoriel de dimension finie et q une forme quadratique sur E. La matrice de q dans une base q-orthonormale est la matrice identité.*

Remarque La matrice de q dans une base q-orthogonale est une matrice diagonale. ◇

En dimension finie, si $\mathcal{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ est une base de E alors d'après la relation (2) p. 440, on a

$$\forall \mathbf{x} \in E \quad q(\mathbf{x}) = \varphi_0(\mathbf{x}, \mathbf{x}) = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i \, x_j \, \varphi_0(\mathbf{e}_i, \mathbf{e}_j),$$

où $x_1, \ldots, x_n$ désignent les composantes de $\mathbf{x}$ dans la base $\mathcal{B}$. Si la base considérée est q-orthogonale alors tous les coefficients $\varphi_0(\mathbf{e}_i, \mathbf{e}_j)$ sont nuls sauf lorsque $i = j$ où ils valent $q(\mathbf{e}_i)$. On a donc

$$\forall \mathbf{x} \in E \quad q(\mathbf{x}) = \sum_{i=1}^{n} x_i^2 \, q(\mathbf{e}_i).$$

Lorsque la base considérée est q-orthonormale, les coefficients $q(\mathbf{e}_i)$ sont égaux à 1 et on a

$$\forall \mathbf{x} \in E \quad q(\mathbf{x}) = \sum_{i=1}^{n} x_i^2.$$

On voit donc quel est l'intérêt de disposer d'une base q-orthonormale (ou q-orthogonale) : la donnée des images d'une telle base par q permet de connaître l'image de n'importe quel élément de E.

Étant donné une forme quadratique, une question que l'on peut se poser est de savoir s'il existe toujours une base orthogonale pour cette forme quadratique. Le théorème suivant apporte la réponse.

THÉORÈME 8.1 *Soient E un $\mathbb{R}$-espace vectoriel de dimension finie et q une forme quadratique sur E. L'espace vectoriel E admet une base q-orthogonale.*

Démonstration On procède par récurrence sur la dimension de E. Pour tout $n \in \mathbb{N}^*$, on va prouver l'assertion $(\mathcal{P}_n)$ suivante :

> « Pour tout $\mathbb{R}$-espace vectoriel E de dimension n et toute forme quadratique q sur E, il existe une base q-orthogonale de E. »

$\trianglerighteq$ Si $n = 1$ alors une base d'un $\mathbb{R}$-espace vectoriel de dimension $n = 1$ est formé par n'importe quel élément non nul de E. Pour toute forme quadratique, une telle base est q-orthogonale : la propriété $(\mathcal{P}_1)$ est donc vraie.

$\trianglerighteq$ Supposons la propriété $(\mathcal{P}_n)$ vraie pour un entier $n \in \mathbb{N}^*$ et montrons que cela entraîne que la propriété $(\mathcal{P}_{n+1})$ est vraie. Soient E un $\mathbb{R}$-espace vectoriel de dimension $n+1$ et q une forme quadratique sur E. Si q est l'application nulle alors toute base de E est q-orthogonale. Intéressons-nous au cas où q n'est pas l'application nulle : il existe $\mathbf{v} \in E$ tel que $q(\mathbf{v}) \neq 0$. La forme polaire φ_0 associée à q étant bilinéaire, l'application $f : \mathbf{x} \in E \mapsto \varphi_0(\mathbf{v}, \mathbf{x})$ est une forme linéaire. Ce n'est pas l'application nulle car $f(\mathbf{v}) = \varphi_0(\mathbf{v}, \mathbf{v}) = q(\mathbf{v})$ et $q(\mathbf{v}) \neq 0$. Le noyau H de cette forme linéaire[35] est un sous-espace vectoriel de E de dimension[36] n. D'après l'hypothèse de récurrence $(\mathcal{P}_n)$, il existe une base $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ de H orthogonale pour la forme quadratique[37] $q|_H$. Ainsi, la famille $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n, \mathbf{v}\}$ est une base q-orthogonale de E : la propriété $(\mathcal{P}_{n+1})$ est vraie. □

Jusqu'à présent nous avons utilisé les propriétés matricielles (produit d'une matrice par un vecteur, changement de bases, etc.) afin d'obtenir des résultats sur les formes bilinéaires ou quadratiques. Ici, c'est le contraire : la formulation du théorème 8.1 sous forme matricielle implique le résultat suivant.

PROPOSITION 8.11 *Toute matrice symétrique réelle est congrue à une matrice diagonale. En d'autres termes, pour toute matrice* $\mathrm{A} \in \mathcal{S}_n(\mathbb{R})$ *il existe une matrice* P *telle que* ${}^t\mathrm{PAP}$ *soit diagonale. De plus, la matrice* P *est inversible.*

---

[35] Rappelons que le noyau d'une application linéaire $f : E \to F$ est le sous-espace vectoriel de E défini par $\mathrm{Ker}(f) = \{\mathbf{x} \in E \ ; \ f(\mathbf{x}) = \mathbf{0}\}$.

[36] En effet, le noyau d'une forme linéaire non nulle $f : E \to \mathbb{R}$ est un hyperplan, c'est-à-dire un sous-espace vectoriel de dimension $\dim(E) - 1$. Ce résultat se déduit du théorème du rang (voir le *Cours de première année* p. 376) : on a $\dim(E) = \dim(\mathrm{Ker}(f)) + \dim(\mathrm{Im}(f))$ avec, pour une forme linéaire non nulle $\mathrm{Im}(f) = \mathbb{R}$.

[37] La restriction d'une forme quadratique sur E à un sous-espace vectoriel de E est encore une forme quadratique, voir p. 448.

Démonstration À toute matrice $A \in \mathcal{S}_n(\mathbb{R})$, on peut associer la forme quadratique q sur $\mathbb{R}^n$ définie par

$$\forall \mathbf{x} \in \mathbb{R}^n \quad q(\mathbf{x}) = {}^t\mathrm{XAX},$$

le vecteur X correspondant à la matrice colonne dont les éléments sont les composantes du vecteur $\mathbf{x}$ dans la base canonique de $\mathbb{R}^n$. D'après le théorème 8.1, il existe une base $\mathcal{B}$ de $\mathbb{R}^n$ qui est q-orthogonale. La matrice D représentant la forme quadratique q dans cette base q-orthogonale est diagonale[38]. Si P désigne la matrice de passage de la base canonique de $\mathbb{R}^n$ à la base $\mathcal{B}$ alors P est une matrice inversible[39] et on a $D = {}^t\mathrm{PAP}$. La matrice ${}^t\mathrm{PAP}$ est donc diagonale. □

8.1.8 Décomposition en somme de carrés de formes linéaires

Quelques notions sur l'ensemble des formes linéaires

L'ensemble $\mathcal{L}(E, \mathbb{R})$ des formes linéaires sur un $\mathbb{R}$-espace vectoriel E possède lui-même une structure de $\mathbb{R}$-espace vectoriel. Cet espace vectoriel $\mathcal{L}(E, \mathbb{R})$ est généralement noté E^* et est appelé le *dual* algébrique[40] de E.

Intéressons-nous au cas où l'espace vectoriel E est un espace vectoriel de dimension finie n. Le dual E^* est un espace vectoriel de dimension finie et on a[41] $\dim E = \dim E^*$. Considérons une base $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ de E et définissons n formes linéaires $\mathbf{e}_j^*$, $j \in \{1, \ldots, n\}$ par

$$\forall j \in \{1, \ldots, n\} \quad \mathbf{e}_j^* : \mathbf{x} \longmapsto x_j \in \mathbb{R},$$

où les réels $x_1, \ldots, x_n$ désignent les composantes de $\mathbf{x}$ dans la base $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$. Les formes linéaires $\mathbf{e}_j^*$ sont caractérisées par les valeurs qu'elles prennent sur une base de E. On a[42]

$$\forall (i, j) \in \{1, \ldots, n\}^2 \quad \mathbf{e}_j^*(\mathbf{e}_i) = \delta_{ij}.$$

Il est aisé de vérifier que les n formes linéaires $\mathbf{e}_1^*, \ldots, \mathbf{e}_n^*$ sont linéairement indépendantes. Puisque l'espace vectoriel E^* des formes linéaires sur E est un espace vectoriel de dimension n, on en déduit que la famille $(\mathbf{e}_1^*, \ldots, \mathbf{e}_n^*)$ est une base de E^*.

---

[38] Voir la remarque de la p. 461.

[39] Voir le *Cours de première année* p. 435.

[40] Dans cet ouvrage, nous n'évoquerons que ce « dual » mais il existe d'autres notions de dual. Ainsi, on appelle dual topologique (noté E') le sous-espace vectoriel de E^* composé des formes linéaires continues (pour la notion de continuité, il faut néanmoins que l'espace vectoriel E soit muni par exemple d'une distance), voir la section 9.3 du chap. 9.

[41] Plus généralement, si E et F sont deux espaces vectoriels de dimension finie alors l'espace vectoriel $\mathcal{L}(E, F)$ des applications linéaires de E dans F est aussi de dimension finie et on a $\dim \mathcal{L}(E, F) = \dim E \times \dim F$, voir le *Cours de première année* p. 416.

[42] Le symbole δ_{ij} vaut 1 si $i = j$ et 0 si $i \neq j$. Ce symbole est généralement appelé symbole de Kronecker.

PROPOSITION 8.12 *Soient E un espace vectoriel de dimension finie et $\mathcal{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ une base de E. La famille de formes linéaires $\{\mathbf{e}_1^*, \ldots, \mathbf{e}_n^*\}$ définie par*

$$\forall (i,j) \in \{1, \ldots, n\}^2 \quad \mathbf{e}_j^*(\mathbf{e}_i) = \delta_{ij}.$$

est une base de E^*. Cette base est appelée la base duale *de $\mathcal{B}$.*

Le principe des méthodes dite « par dualité » est d'obtenir des informations sur un vecteur $\mathbf{x} \in E$ à partir des images de ce vecteur par des formes linéaires sur E. La proposition suivante en est l'illustration.

PROPOSITION 8.13 *Soit $\mathbf{x}$ un élément d'un espace vectoriel E de dimension finie. On a l'équivalence*

$$(\forall f \in E^* \quad f(\mathbf{x}) = 0) \iff \mathbf{x} = \mathbf{0}.$$

Démonstration Considérons une base $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ de E et désignons par x_1, x_2, ..., x_n les composantes de $\mathbf{x}$ dans cette base. Par définition des formes linéaires $\mathbf{e}_i^*$, on a $\mathbf{e}_i^*(\mathbf{x}) = x_i$ pour tout $i \in \{1, \ldots, n\}$. Si pour toute forme linéaire $f \in E^*$ on a $f(\mathbf{x}) = 0$ alors en considérant le cas particulier des formes linéaires $\mathbf{e}_i^*$ on obtient $\mathbf{e}_i^*(\mathbf{x}) = 0$ pour tout $i \in \{1, \ldots, n\}$. Comme $\mathbf{e}_i^*(x) = x_i$, toutes les composantes de $\mathbf{x}$ sont nulles et par conséquent $\mathbf{x} = \mathbf{0}$. La réciproque est évidente[43] : l'image du vecteur nul par une forme linéaire est le réel 0. □

Nous avons construit, pour toute base de E, une base de E^* appelée base duale. La proposition suivante montre qu'inversement, à partir d'une base $\widetilde{\mathcal{B}}$ de E^*, il est possible de trouver une base de E dont $\widetilde{\mathcal{B}}$ est la base duale.

PROPOSITION 8.14 *Soient E un espace vectoriel de dimension finie et $\widetilde{\mathcal{B}} = (f_1, \ldots, f_n)$ qui est une base de E^*. Il existe une base $\mathcal{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ de E telle que $\widetilde{\mathcal{B}}$ soit la base duale de $\mathcal{B}$, c'est-à-dire telle que*

$$\forall i \in \{1, \ldots, n\} \quad f_i = \mathbf{e}_i^*.$$

Démonstration Puisque $\widetilde{\mathcal{B}}$ est une base de l'espace vectoriel E^*, en utilisant la proposition 8.12, on peut construire $(f_1^*, \ldots, f_n^*)$ la base duale de $\widetilde{\mathcal{B}}$ qui est une base de l'espace vectoriel $(E^*)^*$. Montrons que l'application

$$\begin{aligned} \Phi : E &\longrightarrow (E^*)^* \\ \mathbf{x} &\longmapsto \varphi_{\mathbf{x}} \quad \text{où} \quad \varphi_{\mathbf{x}} : f \in E^* \mapsto f(\mathbf{x}) \in \mathbb{R} \end{aligned}$$

est un isomorphisme, c'est-à-dire une application linéaire bijective. Cela établira que les deux espaces vectoriels E et $(E^*)^*$ sont isomorphes.

[43] Voir le *Cours de première année* p. 355.

$\trianglerighteq$ Pour montrer que l'application Φ est un isomorphisme, commençons par remarquer que l'application Φ est linéaire : pour tous $(\lambda, \mu) \in \mathbb{R}^2$ et $(\mathbf{x}, \mathbf{y}) \in E^2$, $\Phi(\lambda\mathbf{x} + \mu\mathbf{y})$ est l'application qui à une forme linéaire $f \in E^*$ associe le réel $f(\lambda\mathbf{x} + \mu\mathbf{y})$. Ainsi pour toute application linéaire f, on a

$$\begin{aligned}\Phi(\lambda\mathbf{x} + \mu\mathbf{y})(f) &= f(\lambda\mathbf{x} + \mu\mathbf{y}) = \lambda f(\mathbf{x}) + \mu f(\mathbf{y}) \\ &= \lambda\Phi(\mathbf{x})(f) + \mu\Phi(\mathbf{y})(f) = (\lambda\Phi(\mathbf{x}) + \mu\Phi(\mathbf{y}))\,(f),\end{aligned}$$

autrement dit, $\Phi(\lambda\mathbf{x} + \mu\mathbf{y}) = \lambda\Phi(\mathbf{x}) + \mu\Phi(\mathbf{y})$.

Montrons ensuite que l'application Φ est injective[44] : si $\Phi(\mathbf{x}) = 0$ alors pour toute application linéaire $f \in E^*$, on a $f(\mathbf{x}) = 0$. D'après la proposition 8.13, cela implique que $\mathbf{x} = \mathbf{0}$: l'application Φ est donc bien injective.

Les dimensions des espaces vectoriels E et $(E^*)^*$ étant identiques, on en déduit que Φ est une application linéaire bijective, et par conséquent un isomorphisme de E sur $(E^*)^*$.

$\trianglerighteq$ La famille $(f_1^*, \ldots, f_n^*)$ est une base de l'espace vectoriel $(E^*)^*$, en utilisant l'isomorphisme Φ, la famille $\{\Phi^{-1}(f_1^*), \ldots, \Phi^{-1}(f_n^*)\}$ est[45] une base de l'espace vectoriel E. Il ne reste plus qu'à montrer que cette base vérifie la condition donnée dans la proposition 8.14, *i.e.* à montrer que

$$\forall i \in \{1, \ldots, n\} \quad \left(\Phi^{-1}(f_i^*)\right)^* = f_i.$$

Soit $i \in \{1, \ldots, n\}$ un entier fixé. Pour montrer que les deux formes linéaires $\left(\Phi^{-1}(f_i^*)\right)^*$ et f_i sont égales, il suffit de prouver qu'elles agissent de la même façon sur une base donnée de E. Considérant la base $(\Phi^{-1}(f_1^*), \ldots, \Phi^{-1}(f_n^*))$ de E, il suffit donc de montrer que :

$$\forall j \in \{1, \ldots, n\} \quad \left(\Phi^{-1}(f_i^*)\right)^* (\Phi^{-1}(f_j^*)) = f_i(\Phi^{-1}(f_j^*)). \tag{12}$$

Tout d'abord, par définition d'une base duale, on a

$$\forall j \in \{1, \ldots, n\} \quad \left(\Phi^{-1}(f_i^*)\right)^* (\Phi^{-1}(f_j^*)) = \delta_{ij}. \tag{13}$$

D'autre part, par définition de l'application Φ, on a $\Phi(\mathbf{x})(f) = f(\mathbf{x})$ pour tous $\mathbf{x} \in E$ et $f \in E^*$. En particulier, pour $\mathbf{x} = \Phi^{-1}(f_j^*)$ et pour $f = f_i$ on a $\Phi(\Phi^{-1}(f_j^*))(f_i) = f_i(\Phi^{-1}(f_j^*))$. On obtient donc

$$\forall j \in \{1, \ldots, n\} \quad f_i(\Phi^{-1}(f_j^*)) = f_j^*(f_i) = \delta_{ij}. \tag{14}$$

Les relations (13) et (14) établissent la relation (12) et termine la démonstration. □

---

[44] Rappelons que pour établir qu'une application L linéaire entre deux espaces vectoriels E et F est injective, il faut montrer que le seul vecteur $\mathbf{x} \in E$ tel que $L(\mathbf{x}) = \mathbf{0}_F$ est le vecteur nul de E, voir la proposition 9.7 p. 367 du *Cours de première année*.

[45] Voir le *Cours de première année* p. 372.

PROPOSITION 8.15 *Soient E un espace vectoriel de dimension finie $n \in \mathbb{N}^*$ et $\{f_1, \dots, f_r\}$ une famille de r formes linéaires indépendantes. On a*

$$\dim\left(\bigcap_{i=1}^{r} \mathrm{Ker}(f_i)\right) = n - r.$$

Démonstration D'après le théorème de la base incomplète[46], la famille libre $\{f_1, \dots, f_r\}$ de l'espace vectoriel E^* peut être complétée en une base $\widetilde{\mathcal{B}} = (f_1, f_2, \dots, f_n)$ de E^*. D'après la proposition 8.14, il existe une base $\mathcal{B} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ de E dont la base duale est $\widetilde{\mathcal{B}}$.

Soit $\mathbf{x}$ un élément de E dont la décomposition dans la base $\mathcal{B}$ est $\mathbf{x} = \sum_{j=1}^{n} x_j \mathbf{e}_j$ avec $(x_1, \dots, x_n) \in \mathbb{R}^n$. On a les équivalences suivantes :

$$\begin{aligned}
\mathbf{x} \in \bigcap_{i=1}^{r} \mathrm{Ker}(f_i) &\iff \forall i \in \{1, \dots, r\} \quad f_i(\mathbf{x}) = 0 \\
&\iff \forall i \in \{1, \dots, r\} \quad f_i\Big(\sum_{j=1}^{n} x_j \mathbf{e}_j\Big) = 0 \\
&\iff \forall i \in \{1, \dots, r\} \quad \sum_{j=1}^{n} x_j f_i(\mathbf{e}_j) = 0
\end{aligned}$$

Puisque $\widetilde{\mathcal{B}}$ est la base duale de la base $\mathcal{B}$, d'après la proposition 8.12, on a pour tout $(i, j) \in \{1, \dots, n\}^2$ l'égalité $f_i(\mathbf{e}_j) = \delta_{ji}$, de sorte que

$$\begin{aligned}
\mathbf{x} \in \bigcap_{i=1}^{r} \mathrm{Ker}(f_i) &\iff \forall i \in \{1, \dots, r\} \quad x_i = 0 \\
&\iff \mathbf{x} \in \mathrm{Vect}\{\mathbf{e}_{r+1}, \dots, \mathbf{e}_n\}.
\end{aligned}$$

Autrement dit on a montré que $\bigcap_{i=1}^{r} \mathrm{Ker}(f_i) = \mathrm{Vect}\{\mathbf{e}_{r+1}, \dots, \mathbf{e}_n\}$. Les $n - r$ vecteurs $\mathbf{e}_{r+1}, \dots, \mathbf{e}_n$ étant des vecteurs linéairement indépendants[47], l'espace vectoriel $\mathrm{Vect}\{\mathbf{e}_{r+1}, \dots, \mathbf{e}_n\}$ est de dimension $n - r$ ce qui termine la preuve. □

Remarques

1. Un sous-espace vectoriel de dimension $n-1$ d'un espace vectoriel E de dimension n est appelé un *hyperplan* de E. Par exemple, les hyperplans de $\mathbb{R}^3$ sont les plans vectoriels alors que les hyperplans de $\mathbb{R}^2$ correspondent aux droites vectorielles. La proposition 8.15 affirme (en prenant $r = 1$) que le noyau d'une forme linéaire non nulle est un hyperplan.

---

[46] Voir par exemple le *Cours de première année* p. 327.

[47] Rappelons qu'une sous-famille d'une famille libre est libre.

2. Il existe un résultat plus général que celui énoncé dans la proposition 8.15 (mais qui ne sera pas utilisé dans ce cours) affirmant que tout sous-espace vectoriel de dimension $n-r$ d'un espace vectoriel de dimension n est l'intersection de r noyaux de formes linéaires indépendantes. Par exemple, tout hyperplan est le noyau d'une forme linéaire non nulle. ◇

Exemples

1. L'application $(x_1, x_2) \in \mathbb{R}^2 \mapsto a_1x_1 + a_2x_2 \in \mathbb{R}$ où $(a_1, a_2) \in \mathbb{R}^2$ est une forme linéaire sur $\mathbb{R}^2$. Elle est non nulle si $(a_1, a_2) \neq (0, 0)$. Dans ce cas, son noyau, qui est l'ensemble $\{(x_1, x_2) \in \mathbb{R}^2 \ ; \ a_1x_1 + a_2x_2 = 0\}$ est un espace vectoriel de dimension 1.

2. Étant donnés deux vecteurs $\mathbf{a} = (a_1, a_2, a_3)$ et $\mathbf{b} = (b_1, b_2, b_3)$ dans $\mathbb{R}^3$, les deux applications $(x_1, x_2, x_3) \in \mathbb{R}^3 \mapsto a_1x_1 + a_2x_2 + a_3x_3 \in \mathbb{R}$ et $(x_1, x_2, x_3) \in \mathbb{R}^3 \mapsto b_1x_1 + b_2x_2 + b_3x_3 \in \mathbb{R}$ sont deux formes linéaires sur $\mathbb{R}^3$. Elles sont indépendantes si les 2 vecteurs $\mathbf{a}$ et $\mathbf{b}$ sont indépendants. La proposition 8.15 indique que l'ensemble

$$\{(x_1, x_2, x_3) \in \mathbb{R}^3 \ ; \ a_1x_1 + a_2x_2 + a_3x_3 = 0 \quad \text{et} \quad b_1x_1 + b_2x_2 + b_3x_3 = 0\}$$

est un espace vectoriel de dimension 1.

Le théorème de Gauss

Le premier exemple de forme quadratique en dimension finie que l'on a vu p. 450 est l'application $\mathbf{x} \in \mathbb{R}^3 \mapsto x_1^2 + x_2^2 + x_3^2$. Sur $\mathbb{R}^4$, on a introduit la forme quadratique de Lorentz définie par $\mathbf{x} \in \mathbb{R}^4 \mapsto x_1^2 + x_2^2 + x_3^2 - c^2x_4^2$. Ces deux formes quadratiques sont des sommes de formes quadratiques « simples » du type $\mathbf{x} \in \mathbb{R}^n \mapsto x_i^2$. La forme quadratique sur $\mathbb{R}^2$ définie par $q : \mathbf{x} \in \mathbb{R}^2 \mapsto 4x_1x_2$ n'a pas, semble-t-il, cette structure. Néanmoins, pour tout $(x_1, x_2) \in \mathbb{R}^2$ en écrivant $4x_1x_2 = (x_1+x_2)^2-(x_1-x_2)^2$ on en déduit que la forme quadratique q peut aussi s'écrire comme la somme de « carrés ». Ce résultat se généralise à toutes les formes quadratiques sur un espace vectoriel de dimension finie et la méthode pour déterminer la décomposition est appelée la méthode de Gauss. Nous allons maintenant décrire cette méthode.

Nous avons établi à la p. 453 que toute forme quadratique sur un $\mathbb{R}$-espace vectoriel E de dimension n peut s'écrire sous la forme[(48)]

$$q(\mathbf{x}) = \sum_{i=1}^{n} a_{ii}\, x_i^2 + \sum_{i=1}^{n} \sum_{j=i+1}^{n} a_{ij}\, x_i\, x_j,$$

chaque a_{ij}, $(i, j) \in \{1, \ldots, n\}^2$ désignant un nombre réel et les x_i, $i \in \{1, \ldots, n\}$ désignant les coefficients de $\mathbf{x} \in E$ dans une base fixée de E.

[(48)] Par rapport à la décomposition de la p. 453 , on a adopté ici le changement de notation a_{ij} au lieu de $2a_{ij}$ pour $i < j$.

GAUSS, Carl-Friedrich (1777, Brunswick - 1855, Göttingen).



Carl-Friedrich Gauss a soutenu en 1799 une thèse sur le théorème fondamental de l'algèbre : « tout polynôme sur le corps des nombres complexes a au moins un zéro » dont il apporta une preuve rigoureuse. Il s'intéressa aussi à la géométrie non euclidienne et étudia les surfaces. Le nom de Gauss est également attaché au domaine des probabilités avec la loi de Gauss ou loi normale. En tant que physicien, ses travaux les plus remarquables concernent la mesure des grandeurs magnétiques et en particulier l'induction magnétique.

Afin de faire apparaître uniquement des « carrés », on va utiliser une méthode connue sous le nom de *méthode de Gauss*[49]. Selon les coefficients a_{ij}, deux cas sont à envisager.

$\trianglerighteq$ Premier cas : il existe au moins un entier i tel que $a_{ii} \neq 0$; supposons qu'il s'agisse de a_{11}. En notant $\mathbf{x}' = (x_2, \ldots, x_n)$, on écrit dans un premier temps q sous la forme :

$$q(\mathbf{x}) = a_{11}x_1^2 + x_1 B(\mathbf{x}') + C(\mathbf{x}'),$$

où B est une forme linéaire sur un espace vectoriel de dimension $n-1$ et C une forme quadratique sur un espace vectoriel de dimension $n-1$. La forme quadratique q peut ensuite s'écrire comme la somme du carré d'une forme linéaire et d'une forme quadratique en $\mathbf{x}'$ de la manière suivante :

$$q(\mathbf{x}) = a_{11}\left(x_1 + \frac{B(\mathbf{x}')}{2a_{11}}\right)^2 + \left(C(\mathbf{x}') - \frac{B(\mathbf{x}')^2}{4a_{11}}\right).$$

$\trianglerighteq$ Deuxième cas : tous les termes a_{ii} sont nuls et il existe un coefficient a_{ij} , $i \neq j$, non nul[50]. Supposons qu'il s'agisse de a_{12}. En notant $\mathbf{x}'' = (x_3, \ldots, x_n)$, on écrit dans un premier temps q sous la forme :

$$q(\mathbf{x}) = a_{12}x_1x_2 + x_1 B(\mathbf{x}'') + x_2 C(\mathbf{x}'') + D(\mathbf{x}''),$$

où B et C sont des formes linéaires et D une forme quadratique. On écrit ensuite q comme la somme de deux carrés de forme linéaire et d'une forme

[49] A ne pas confondre avec la méthode du pivot de Gauss introduite pour résoudre des systèmes linéaires, voir le *Cours de première année*.

[50] On n'a pas traité le cas où tous les coefficients a_{ij} sont nuls ; ce cas correspond à $q = 0$ qui est déjà écrite sous la forme de sommes de carrés.

quadratique en $\mathbf{x}''$ de la façon suivante :

$$\begin{aligned} q(\mathbf{x}) &= a_{12}\left(x_1 + \frac{C(\mathbf{x}'')}{a_{12}}\right)\left(x_2 + \frac{B(\mathbf{x}'')}{a_{12}}\right) + \left(D(\mathbf{x}'') - \frac{B(\mathbf{x}'')C(\mathbf{x}'')}{a_{12}}\right) \\ &= \frac{a_{12}}{4}\left(\left(x_1 + x_2 + \frac{B(\mathbf{x}'') + C(\mathbf{x}'')}{a_{12}}\right)^2 - \left(x_1 - x_2 + \frac{C(\mathbf{x}'') - B(\mathbf{x}'')}{a_{12}}\right)^2\right) \\ &\quad + \left(D(\mathbf{x}'') - \frac{B(\mathbf{x}'')C(\mathbf{x}'')}{a_{12}}\right). \end{aligned}$$

On peut itérer alors ce processus pour décomposer une forme quadratique comme la somme de carrés de formes linéaires sur l'espace vectoriel E et d'une forme quadratique sur un espace vectoriel de dimension $m < n$. Au terme du processus, on a démontré le théorème suivant.

THÉORÈME 8.2 (Théorème de Gauss)
Soit q une forme quadratique sur un $\mathbb{R}$-espace vectoriel E de dimension n. Les assertions suivantes sont équivalentes :

- *le rang de q est égal à r ;*
- *il existe r formes linéaires indépendantes*[51] *$f_1, f_2, \ldots, f_r$ sur E et il existe r réels non nuls $\gamma_1, \gamma_2, \ldots, \gamma_r$ tels que*

$$q = \sum_{i=1}^{r} \gamma_i f_i^2.$$

Démonstration La construction qui a été effectuée en préambule indique que la première assertion implique la seconde. Montrons la réciproque ; supposons qu'il existe r formes linéaires indépendantes $f_1, f_2, \ldots, f_r$ et qu'il existe r réels non nuls $\gamma_1, \gamma_2, \ldots, \gamma_r$ tels que $q = \sum_{i=1}^{r} \gamma_i f_i^2$. D'après le théorème de la base incomplète[52], la famille $\{f_1, \ldots, f_r\}$ de l'espace vectoriel E^*, qui est une famille libre, peut être complétée en une base $\widetilde{\mathcal{B}} = (f_1, f_2, \ldots, f_n)$ de E^*. D'après la proposition 8.14, il existe une base $\mathcal{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ de E dont la base duale est $\widetilde{\mathcal{B}}$. Déterminons la matrice de q dans cette base en évaluant $\varphi_0(\mathbf{e}_i, \mathbf{e}_j)$ pour tout $(i, j) \in \{1, \ldots, n\}^2$ où φ_0 est la forme polaire de q définie par

$$\forall (\mathbf{x}, \mathbf{y}) \in E^2 \quad \varphi_0(\mathbf{x}, \mathbf{y}) = \sum_{k=1}^{r} \gamma_k f_k(\mathbf{x}) f_k(\mathbf{y}).$$

[51] Une forme linéaire sur un $\mathbb{R}$-espace vectoriel E est une application linéaire de E sur $\mathbb{R}$. L'ensemble des formes linéaires sur E constitue un $\mathbb{R}$-espace vectoriel. Dire que r formes linéaires sont indépendantes signifie que la famille formée par ces formes linéaires est une famille libre.

[52] Voir par exemple le *Cours de première année* p. 327.

Par définition d'une base duale[53] on a $f_i(\mathbf{e}_j) = \delta_{ij}$ pour tout $(i,j) \in \{1,\dots,n\}^2$ et par conséquent

$$\forall (i,j) \in \{1,\dots,n\}^2 \quad \varphi_0(\mathbf{e}_i,\mathbf{e}_j) = \sum_{k=1}^{r} \gamma_k \delta_{ik}\delta_{jk}.$$

On a $\delta_{ik} \neq 0$ si et seulement si $k = i$ et $\delta_{jk} \neq 0$ si et seulement si $k = j$ donc $\varphi_0(\mathbf{e}_i,\mathbf{e}_j) \neq 0$ si et seulement si $i = j$. La matrice de q dans la base $\mathcal{B}$ est donc diagonale. De plus, pour tout $i \in \{1,\dots,r\}$ on a $\varphi_0(\mathbf{e}_i,\mathbf{e}_i) = \gamma_i$ et pour tout $i \in \{r+1,\dots,n\}$ on a $\varphi_0(\mathbf{e}_i,\mathbf{e}_i) = 0$. La matrice est donc composée sur la diagonale des réels non nuls γ_1, γ_2,..., γ_r, complétée par des 0. Le rang d'une telle matrice est r. □

Exemples

1. Décomposons en utilisant la méthode de Gauss la forme quadratique q définie sur $\mathbb{R}^3$ par

$$q(\mathbf{x}) = x_1^2 + 2x_2^2 + 8x_3^2 - 2x_1x_2 + 4x_1x_3.$$

On a

$$\begin{aligned} q(\mathbf{x}) &= x_1^2 + x_1(-2x_2 + 4x_3) + 2x_2^2 + 8x_3^2 \\ &= \left((x_1^2 + 2x_1(2x_3 - x_2) + (2x_3 - x_2)^2\right) - (2x_3 - x_2)^2 + 2x_2^2 + 8x_3^2 \\ &= (x_1 + 2x_3 - x_2)^2 + 4x_3^2 + x_2^2 + 4x_2x_3 = (x_1 + 2x_3 - x_2)^2 + (x_2 + 2x_3)^2. \end{aligned}$$

On a donc $q(\mathbf{x}) = f_1(\mathbf{x})^2 + f_2(\mathbf{x})^2$ où $f_1(\mathbf{x}) = x_1 + 2x_3 - x_2$ et $f_1(\mathbf{x}) = x_2 + 2x_3$.

2. Décomposons en utilisant la méthode de Gauss la forme quadratique q définie sur $\mathbb{R}^4$ par

$$q(\mathbf{x}) = x_1x_2 + x_1x_3 + x_1x_4 + x_2x_3 + x_2x_4 + x_3x_4.$$

Aucun carré n'apparaissant, on regroupe tous les termes faisant apparaître x_1 et x_2 de la façon suivante :

$$\begin{aligned} q(\mathbf{x}) &= x_1x_2 + x_1(x_3 + x_4) + x_2(x_3 + x_4) + x_3x_4 \\ &= (x_1 + x_3 + x_4)(x_2 + x_3 + x_4) - (x_3 + x_4)^2 + x_3x_4 \\ &= \tfrac{1}{4}(x_1 + x_2 + 2(x_3 + x_4))^2 - \tfrac{1}{4}(x_1 - x_2)^2 - x_3^2 - x_4^2 - x_3x_4. \end{aligned}$$

Il reste ensuite à s'occuper des termes ne faisant pas apparaître x_1 et x_2, c'est-à-dire de la forme quadratique sur $\mathbb{R}^2$ définie par

$$\widetilde{q}(x_3,x_4) = x_3^2 + x_4^2 + x_3x_4.$$

On regroupe dans un même terme tous les termes où est présent x_3 :

$$\widetilde{q}(x_3,x_4) = \left(x_3 + \tfrac{1}{2}x_4\right)^2 - \tfrac{1}{4}x_4^2 + x_4^2 = \left(x_3 + \tfrac{1}{2}x_4\right)^2 + \tfrac{3}{4}x_4^2.$$

[53] Voir la proposition 8.12.

On obtient finalement :

$$q(\mathbf{x}) = f_1(\mathbf{x})^2 - f_2(\mathbf{x})^2 - f_3(\mathbf{x})^2 - 3f_4(\mathbf{x})^2$$

où les formes linéaires f_i sont données par $f_1(\mathbf{x}) = \frac{1}{2}(x_1 + x_2 + 2x_3 + 2x_4)$, $f_2(\mathbf{x}) = \frac{1}{2}(x_1 - x_2)$, $f_3(\mathbf{x}) = x_3 + \frac{1}{2}x_4$ et $f_4(\mathbf{x}) = \frac{1}{2}x_4$.

Remarque Il existe bien sûr d'autres décompositions d'une forme quadratique comme combinaison linéaire de carrés de formes linéaires que celle donnée par le théorème de Gauss. Par exemple, la forme quadratique q définie sur $\mathbb{R}^2$ par

$$q(\mathbf{x}) = (x_1 + x_2)^2 - x_1^2 + (x_1 - x_2)^2$$

est une combinaison linéaire de carrés de formes linéaires. On a $q(\mathbf{x}) = g_1(\mathbf{x})^2 - g_2(\mathbf{x})^2 + g_3(\mathbf{x})^2$ avec

$$g_1(\mathbf{x}) = x_1 + x_2, \quad g_2(\mathbf{x}) = x_1, \quad g_3(\mathbf{x}) = x_1 - x_2.$$

Par contre, les trois formes linéaires g_1, g_2 et g_3 ne sont pas linéairement indépendantes car $g_1 + g_3 = 2g_2$.

Ici, en développant l'écriture de la forme quadratique q, on a

$$q(\mathbf{x}) = x_1^2 + 2x_2^2 = f_1(\mathbf{x})^2 + 2f_2(\mathbf{x})^2,$$

et les deux formes linéaires $f_1 : \mathbf{x} \mapsto x_1$ et $f_2 : \mathbf{x} \mapsto x_2$ sont indépendantes.

Un des avantages de la méthode de Gauss est qu'elle assure l'indépendance des formes linéaires obtenues. ◇

Exercice 8.2 Pour $a \in \mathbb{R}$, on considère la forme quadratique q_a définie sur $\mathbb{R}^3$ par

$$q_a(\mathbf{x}) = x_1^2 + (1 + a)x_2^2 + (1 + a + a^2)x_3^2 + 2x_1x_2 - 2ax_2x_3.$$

Trouver une décomposition de q_a comme combinaison linéaire de carrés de formes linéaires indépendantes.

8.1.9 Classification des formes quadratiques

De nombreux travaux mathématiques ont pour but de classifier les objets qu'ils définissent. La question se pose de savoir si on peut classer les formes quadratiques par type, et selon quels critères. Le théorème de Gauss (théorème 8.2) fournit une première classification selon le rang de la forme quadratique : toute forme quadratique de rang r est une somme de r carrés de formes linéaires. On peut être plus précis au niveau de la classification en prenant en compte le signe devant les carrés dans la décomposition.

Soient q une forme quadratique sur un $\mathbb{R}$-espace vectoriel E de dimension n et $\mathcal{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ une base q-orthogonale de E. On note

$$\alpha(\mathcal{B}) = card\{i \in \mathbb{N} \ ; \ 1 \leqslant i \leqslant n \ \text{ et } \ q(\mathbf{e}_i) > 0\}$$
$$\text{et} \quad \beta(\mathcal{B}) = card\{i \in \mathbb{N} \ ; \ 1 \leqslant i \leqslant n \ \text{ et } \ q(\mathbf{e}_i) < 0\}.$$

On a vu p. 461 que la matrice de q dans la base q-orthogonale $\mathcal{B}$ est une matrice diagonale, les coefficients diagonaux de cette matrice étant les réels $q(\mathbf{e}_i)$. Les entiers $\alpha(\mathcal{B})$ et $\beta(\mathcal{B})$ correspondent respectivement aux nombres de valeurs propres strictement positives et strictement négatives de la matrice de q dans la base $\mathcal{B}$. On en déduit que le rang de q, que l'on sait être indépendant de la base choisie[54], vaut $\mathrm{rg}(q) = \alpha(\mathcal{B}) + \beta(\mathcal{B})$.

Le théorème suivant indique que non seulement $\alpha(\mathcal{B})+\beta(\mathcal{B})$ est indépendant de la base orthogonale $\mathcal{B}$ mais que les deux entiers $\alpha(\mathcal{B})$ et $\beta(\mathcal{B})$ sont eux-mêmes indépendants de ce choix.

THÉORÈME 8.3 (Loi d'inertie de Sylvester)
Soient q une forme quadratique sur un $\mathbb{R}$-espace vectoriel E de dimension n et $\mathcal{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ une base q-orthogonale de E. On note

$$\alpha(\mathcal{B}) = card\{i \in \mathbb{N} \ ; \ 1 \leqslant i \leqslant n \ \text{ et } \ q(\mathbf{e}_i) > 0\}$$
$$\text{et} \quad \beta(\mathcal{B}) = card\{i \in \mathbb{N} \ ; \ 1 \leqslant i \leqslant n \ \text{ et } \ q(\mathbf{e}_i) < 0\}.$$

Les entiers $\alpha(\mathcal{B})$ et $\beta(\mathcal{B})$ sont indépendants de la base orthogonale $\mathcal{B}$.

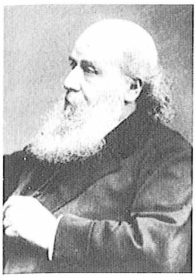
Démonstration Afin de montrer que les deux entiers $\alpha(\mathcal{B})$ et $\beta(\mathcal{B})$ ne dépendent pas de la base orthogonale choisie, nous allons considérer deux bases q-orthogonales $\mathcal{B} = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$ et $\mathcal{C} = (\mathbf{c}_1, \ldots, \mathbf{c}_n)$ et montrer que $\alpha(\mathcal{B}) = \alpha(\mathcal{C})$ puis que $\beta(\mathcal{B}) = \beta(\mathcal{C})$. Quitte à échanger l'ordre des vecteurs des deux bases $\mathcal{B}$ et $\mathcal{C}$, on peut supposer que :

$$\begin{aligned}
&q(\mathbf{b}_i) > 0 \quad \text{pour} \quad 1 \leqslant i \leqslant \alpha(\mathcal{B}),\\
&q(\mathbf{b}_i) < 0 \quad \text{pour} \quad \alpha(\mathcal{B}) + 1 \leqslant i \leqslant \alpha(\mathcal{B}) + \beta(\mathcal{B}) = \mathrm{rg}(q),\\
&q(\mathbf{c}_i) > 0 \quad \text{pour} \quad 1 \leqslant i \leqslant \alpha(\mathcal{C}),\\
&q(\mathbf{c}_i) < 0 \quad \text{pour} \quad \alpha(\mathcal{C}) + 1 \leqslant i \leqslant \alpha(\mathcal{C}) + \beta(\mathcal{C}) = \mathrm{rg}(q).
\end{aligned}$$

$\trianglerighteq$ Considérons les sous-espaces vectoriels $G_+ = \mathrm{Vect}\{\mathbf{b}_i \ ; \ 1 \leqslant i \leqslant \alpha(\mathcal{B})\}$ et $H_- = \mathrm{Vect}\{\mathbf{c}_i \ ; \ \alpha(\mathcal{C})+1 \leqslant i \leqslant n\}$. Nous allons dans un premier temps montrer que $G_+ \cap H_- = \{\mathbf{0}\}$. L'inclusion $\{\mathbf{0}\} \subset G_+ \cap H_-$ est immédiate car l'intersection de deux sous-espaces vectoriels est un sous-espace vectoriel et tout sous-espace vectoriel contient l'élément $\mathbf{0}$. Considérons $\mathbf{x} \in G_+ \cap H_-$ et montrons, en raisonnant par l'absurde, que nécessairement $\mathbf{x} = \mathbf{0}$. Supposons $\mathbf{x} \neq \mathbf{0}$; il se

[54] Voir la définition du rang d'une forme bilinéaire p. 445.

SYLVESTER, James-Joseph (1814, Londres - 1897, Londres).



Mathématicien et géomètre anglais, Sylvester a enseigné à l'université de Baltimore aux Etats-Unis, puis à l'université d'Oxford en Angleterre. Ses principaux travaux portent sur les formes quadratiques et sur la théorie des déterminants. On lui doit le terme de matrice qu'il introduit en 1850 pour désigner un tableau rectangulaire de nombres. En 1901, une distinction est créée en son honneur : la médaille Sylvester. Elle est décernée tous les trois ans par la Royal Society (Académie des Sciences britannique).

décompose de manière unique dans la base $\{\mathbf{b}_i \ ; \ 1 \leqslant i \leqslant \alpha(\mathcal{B})\}$ de G_+ de la manière suivante : $\mathbf{x} = \sum_{i=1}^{\alpha(\mathcal{B})} \lambda_i \mathbf{b}_i$ où les λ_i sont des réels non tous nuls. Puisque la base $\mathcal{B}$ est q-orthogonale, on en déduit que $q(\mathbf{x}) = \sum_{i=1}^{\alpha(\mathcal{B})} \lambda_i^2 q(\mathbf{b}_i)$. De plus, chaque $q(\mathbf{b}_i)$ $(i = 1, \ldots, \alpha(B))$ est strictement positif, ce qui implique $q(\mathbf{x}) > 0$. En raisonnant de manière analogue, on montre à partir de l'hypothèse $\mathbf{x} \in H_- \setminus \{\mathbf{0}\}$ que $q(\mathbf{x}) \leqslant 0$. Ainsi, si $\mathbf{x} \in G_+ \cap H_-$ et si $\mathbf{x} \neq \mathbf{0}$ alors on a à la fois $q(\mathbf{x}) > 0$ et $q(\mathbf{x}) \leqslant 0$ ce qui constitue la contradiction.

$\trianglerighteq$ L'intersection des espaces G_+ et H_- étant réduite à $\{\mathbf{0}\}$, on sait que[55] :

$$\dim G_+ + \dim H_- = \dim(G_+ \oplus H_-).$$

Par ailleurs, comme $G_+ \oplus H_- \subset E$, on a $\dim(G_+ \oplus H_-) \leqslant \dim E$. On en déduit que $\dim G_+ + \dim H_- \leqslant \dim E$, c'est-à-dire que $\alpha(\mathcal{B}) + n - \alpha(\mathcal{C}) \leqslant n$ ou encore que $\alpha(\mathcal{B}) \leqslant \alpha(\mathcal{C})$. On a donc prouvé que si $\mathcal{B}$ et $\mathcal{C}$ sont deux bases q-orthogonales alors $\alpha(\mathcal{B}) \leqslant \alpha(\mathcal{C})$. En effectuant un raisonnement analogue en échangeant le rôle de $\mathcal{B}$ et celui de $\mathcal{C}$, c'est-à-dire en introduisant les espaces $H_+ = \text{Vect}\{\mathbf{c}_i \ ; \ 1 \leqslant i \leqslant \alpha(\mathcal{C})\}$ et $G_- = \text{Vect}\{\mathbf{b}_i \ ; \ \alpha(\mathcal{B}) + 1 \leqslant i \leqslant n\}$, on montre que $\alpha(\mathcal{C}) \leqslant \alpha(\mathcal{B})$. Ceci permet de prouver que $\alpha(\mathcal{B}) = \alpha(\mathcal{C})$.

Pour conclure cette preuve remarquons que puisque le rang de la forme quadratique q ne dépend pas de la base[56] on a $\text{rg}(q) = \alpha(\mathcal{B}) + \beta(\mathcal{B}) = \alpha(\mathcal{C}) + \beta(\mathcal{C})$. L'égalité des entiers $\alpha(\mathcal{B})$ et $\alpha(\mathcal{C})$ implique celle des entiers $\beta(\mathcal{B})$ et $\beta(\mathcal{C})$. □

Les entiers $\alpha(\mathcal{B})$ et $\beta(\mathcal{B})$ ne dépendant pas de la base $\mathcal{B}$, ils peuvent donc être notés α et β sans faire référence à la base.

[55] Voir la formule de Grassmann dans le *Cours de première année* p. 343.

[56] Voir la définition du rang d'une forme bilinéaire, définition 8.4.

DÉFINITION 8.13 (Signature d'une forme quadratique)
Soient q une forme quadratique sur un espace vectoriel E de dimension n, $\mathcal{B} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ est une base q-orthogonale de E et α, β les entiers

$$\alpha = card\{i \in \mathbb{N} \ ; \ 1 \leqslant i \leqslant n \ \ et \ \ q(\mathbf{e}_i) > 0\}$$
$$et \quad \beta = card\{i \in \mathbb{N} \ ; \ 1 \leqslant i \leqslant n \ \ et \ \ q(\mathbf{e}_i) < 0\}.$$

On appelle signature *de la forme quadratique q le couple $\varepsilon(q) = (\alpha, \beta)$.*

Calculer la signature d'une forme quadratique en utilisant la définition des entiers α et β requiert une base orthogonale pour cette forme quadratique. Le théorème de Gauss permet de déterminer la signature d'une forme quadratique sans avoir recours à une telle base. Plus précisément, on a le théorème suivant.

THÉORÈME 8.4 *Soient E un $\mathbb{R}$-espace vectoriel de dimension finie et q une forme quadratique sur E dont une décomposition en carrés de formes linéaires indépendantes est $q = \sum_{i=1}^{r} \gamma_i f_i^2$. La signature de q est (α, β) avec*

$$\alpha = card\{i \in \mathbb{N} \ ; \ 1 \leqslant i \leqslant r \ \ et \ \ \gamma_i > 0\}$$
$$et \quad \beta = card\{i \in \mathbb{N} \ ; \ 1 \leqslant i \leqslant r \ \ et \ \ \gamma_i < 0\}.$$

Démonstration Le théorème de la base incomplète[57] permet de compléter la famille libre $\{f_1, \dots, f_r\}$ en une base $\widetilde{\mathcal{B}} = (f_1, \dots, f_n)$ de l'espace vectoriel E^*. D'après la proposition 8.14, il existe une base $\mathcal{B} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ de E dont la base duale est $\widetilde{\mathcal{B}}$. Dans cette base $\mathcal{B}$, la matrice de q est une matrice diagonale[58] qui s'écrit

$$\mathrm{Mat}_{\mathcal{B}}(q) = \begin{pmatrix} \gamma_1 & & & & & \\ & \ddots & & & & 0 \\ & & \gamma_r & & & \\ & & & 0 & & \\ & 0 & & & \ddots & \\ & & & & & 0 \end{pmatrix}.$$

En particulier, la base $\mathcal{B}$ est orthogonale selon q et on a

$$\alpha(\mathcal{B}) = card\{i \in \mathbb{N} \ ; \ 1 \leqslant i \leqslant r \ \text{ et } \ \gamma_i > 0\}$$
$$\text{et} \quad \beta(\mathcal{B}) = card\{i \in \mathbb{N} \ ; \ 1 \leqslant i \leqslant r \ \text{ et } \ \gamma_i < 0\}.$$

Le théorème est démontré. □

[57] Voir le *Cours de première année* p. 327.

[58] On pourra consulter la preuve du théorème de Gauss, à la p. 469, au cours de laquelle un calcul similaire est détaillé.

Exemples

1. La forme de Lorentz définie sur $\mathbb{R}^4$ par

$$\varphi : (\mathbf{x}, \mathbf{y}) \in \mathbb{R}^4 \times \mathbb{R}^4 \longmapsto x_1\, y_1 + x_2\, y_2 + x_3\, y_3 - c^2 x_4\, y_4,$$

où c désigne un paramètre réel, admet pour forme quadratique associée l'application

$$q : \mathbf{x} \in \mathbb{R}^4 \longmapsto x_1^2 + x_2^2 + x_3^2 - c^2 x_4^2.$$

Déterminons la signature de q en utilisant le théorème 8.4. Remarquons que l'expression $q(\mathbf{x}) = x_1^2 + x_2^2 + x_3^2 - c^2 x_4^2$ est une décomposition en somme de quatre carrées de formes linéaires et les quatre formes linéaires $\mathbf{x} \mapsto x_1$, $\mathbf{x} \mapsto x_2$, $\mathbf{x} \mapsto x_3$ et $\mathbf{x} \mapsto x_4$ sont indépendantes. Selon la valeur du coefficient c on a

$$\varepsilon(q) = (3,0) \text{ si } c = 0 \qquad \text{et} \qquad \varepsilon(q) = (3,1) \text{ si } c \neq 0\ .$$

2. La forme quadratique sur $\mathbb{R}^4$ définie par

$$q : \mathbf{x} \in \mathbb{R}^4 \longmapsto x_1x_2 + x_1x_3 + x_1x_4 + x_2x_3 + x_2x_4 + x_3x_4$$

a pour décomposition[59] en somme de carrés de formes linéaires indépendantes

$$q(\mathbf{x}) = f_1(\mathbf{x})^2 - f_2(\mathbf{x})^2 - f_3(\mathbf{x})^2 - 3f_4(\mathbf{x})^2$$

où les formes linéaires f_i sont données par $f_1(\mathbf{x}) = \frac{1}{2}(x_1 + x_2 + 2x_3 + 2x_4)$, $f_2(\mathbf{x}) = \frac{1}{2}(x_1 - x_2)$, $f_3(\mathbf{x}) = x_3 + \frac{1}{2}x_4$ et $f_4(\mathbf{x}) = \frac{1}{2}x_4$. D'après le théorème 8.2, le rang de q est 4 et d'après le théorème 8.4 sa signature vaut $(1,3)$.

Exercice 8.3 Déterminer en fonction du réel a le rang et la signature de la forme quadratique q_a définie par, voir l'exercice 8.2 p. 471,

$$q_a(\mathbf{x}) = x_1^2 + (1+a)x_2^2 + (1+a+a^2)x_3^2 + 2x_1x_2 - 2ax_2x_3.$$

Autre critère de classification

On dit que deux formes quadratiques q et q' sur un $\mathbb{R}$-espace vectoriel E de dimension finie sont *isométriques* s'il existe une application linéaire et bijective u de E sur E telle que :

$$\forall \mathbf{x} \in E \qquad q(\mathbf{x}) = q'(u(\mathbf{x})).$$

Notons φ (respectivement φ') la forme polaire de q (resp. de q'). En utilisant la relation de polarisation donnée à la proposition 8.6, on établit que les deux formes quadratiques q et q' sont isométriques si et seulement si

$$\forall (\mathbf{x}, \mathbf{y}) \in E^2 \qquad \varphi(\mathbf{x}, \mathbf{y}) = \varphi'(u(\mathbf{x}), u(\mathbf{y})). \tag{15}$$

[59] Voir l'exemple de la p. 470.

Considérons la matrice A (resp. B) représentant q (resp. q') dans une base $\mathcal{B}$ de E. Matriciellement, la relation (15) se traduit sous la forme

$$\forall (X, Y) \in (\mathbb{R}^n)^2 \qquad {}^tXAY = {}^t(PX)B(PY) = {}^tX{}^tPBPY, \tag{16}$$

où P est la matrice de l'application linéaire u dans la base $\mathcal{B}$ et où X et Y sont les vecteurs colonnes représentant les éléments $\mathbf{x}$ et $\mathbf{y}$ de E dans cette base. La relation (16) étant valable pour tout $(X, Y) \in (\mathbb{R}^n)^2$, on en déduit[60] que q et q' sont isométriques si et seulement si il existe une matrice $P \in GL_n(\mathbb{R})$ telle que

$$A = {}^tPBP. \tag{17}$$

La relation d'isométrie est une relation d'équivalence sur l'ensemble des formes quadratiques sur E. Classifier les formes quadratiques pour cette relation revient à trouver les classes d'équivalence pour cette relation d'équivalence. Ainsi, on dispose de deux façons pour classer les formes quadratiques : les ranger selon leur signature ou les ranger par classe d'isométrie. On peut montrer que ces deux classifications sont identiques.

PROPOSITION 8.16 *Deux formes quadratiques sur un $\mathbb{R}$-espace vectoriel de dimension finie sont isométriques si et seulement si elles ont même signature.*

8.2 Produit scalaire sur un espace vectoriel

8.2.1 Produit scalaire

DÉFINITION 8.14 *Soient E un $\mathbb{R}$-espace vectoriel et q une forme quadratique sur E.*

✗ *La forme quadratique q est dite* positive *si $q(\mathbf{x}) \geqslant 0$ pour tout $\mathbf{x} \in E$.*

✗ *La forme quadratique q est dite* définie *si $q(\mathbf{x}) \neq 0$ pour tout $\mathbf{x} \in E \setminus \{\mathbf{0}\}$.*

Remarque Pour montrer qu'une forme quadratique q sur un $\mathbb{R}$-espace vectoriel E est définie, il faut montrer que si $\mathbf{x} \neq \mathbf{0}$ alors $q(\mathbf{x}) \neq 0$. En pratique, il est souvent plus simple d'utiliser la contraposée : on montre que si $q(\mathbf{x}) = 0$ alors nécessairement $\mathbf{x} = \mathbf{0}$.

Inversement, pour montrer qu'une forme quadratique q sur un $\mathbb{R}$-espace vectoriel E n'est pas définie, il suffit de trouver un vecteur $\mathbf{x} \neq \mathbf{0}$ pour lequel $q(\mathbf{x}) = 0$. ◇

[60] Cette affirmation résulte du fait que si pour tout $X \in \mathbb{R}^n$ on a ${}^tXZ = 0$ alors, en choisissant $X = Z$ on montre que nécessairement $Z = 0$. De même, si pour tout $Y \in \mathbb{R}^n$ on a ${}^tZY = 0$ alors on montre que $Z = 0$.

Le terme « défini » n'est pas à prendre dans son sens usuel en mathématiques. Pour une forme quadratique q sur un espace vectoriel E, le réel $q(\mathbf{x})$ « existe » pour tout $\mathbf{x} \in E$.

Exemples

1. La forme quadratique $q : \mathbf{x} \in \mathbb{R}^n \mapsto \sum_{i=1}^n x_i^2$ est une forme quadratique définie et une forme quadratique positive. En effet, cette application est positive car une somme de carrés est toujours positive. Elle est définie car $q(\mathbf{x}) = 0$ implique que $x_i = 0$ pour tout $i \in \{1, \dots, n\}$, donc que $\mathbf{x} = \mathbf{0}$ (une somme de réels positifs est nulle si et seulement si tous les termes la composant sont nuls).

2. La forme quadratique $q : f \in \mathcal{C}([0,1], \mathbb{R}) \mapsto \int_0^1 f(t)^2 \, \mathrm{d}t$ est aussi une forme quadratique définie et une forme quadratique positive. La positivité provient de la positivité de l'intégrale d'une fonction positive[61]. C'est la continuité de f qui permet[62] d'affirmer que $\varphi(f, f) \neq 0$ lorsque $f \neq 0$: l'application q est donc une forme quadratique définie.

3. La forme quadratique associée à la forme de Lorentz donnée par : $q : \mathbf{x} \in \mathbb{R}^4 \mapsto x_1^2 + x_2^2 + x_3^2 - c^2 x_4^2$ avec $c > 0$ n'est ni positive (puisque $q((0,0,0,1)) = -c^2 < 0$), ni définie (puisque $q((c,0,0,1)) = 0$).

4. La forme quadratique $q : \mathrm{A} \in \mathcal{M}_n(\mathbb{R}) \mapsto \mathrm{Tr}({}^t\mathrm{AA})$ est une forme quadratique définie et une forme quadratique positive. En effet, si on note $(a_{ij})_{(i,j)\in\{1,\dots,n\}^2}$ les coefficients de la matrice A alors le coefficient (i,j) de la matrice ${}^t\mathrm{AA}$ vaut[63] $\sum_{k=1}^n a_{ki} a_{kj}$. La trace d'une matrice étant la somme de ces coefficients diagonaux, on obtient

$$q(\mathrm{A}) = \sum_{i=1}^n \sum_{k=1}^n a_{ki}^2.$$

On se retrouve dans le même cas que celui du premier des exemples décrits ici : pour tout $\mathrm{A} \in \mathcal{M}_n(\mathbb{R})$, $q(\mathrm{A})$ est positive car c'est une somme de carrés. Elle est définie car $q(\mathrm{A}) = 0$ implique que pour tout $(i,j) \in \{1, \dots, n\}^2$ on a $a_{ij} = 0$, autrement dit que $\mathrm{A} = 0$.

Remarques

1. On dira qu'une *forme quadratique* q sur un $\mathbb{R}$-espace vectoriel E est *définie positive* si elle est à la fois définie et positive. Autrement dit, une forme quadratique q est définie positive si $q(\mathbf{x}) > 0$ pour tout $\mathbf{x} \in E \setminus \{\mathbf{0}\}$.

---

[61] Voir la proposition 18.4 p. 844 du *Cours de première année*.

[62] Voir la proposition 18.5 p. 844 du *Cours de première année*.

[63] On rappelle, voir la définition 10.8 p. 398 du *Cours de première année*, que pour tout $(\mathrm{A}, \mathrm{B}) \in \mathcal{M}_n(\mathbb{R})^2$ le terme d'indice (i,j) du produit $\mathrm{A} \times \mathrm{B}$ vaut $\sum_{k=1}^n a_{ik} b_{kj}$. Par ailleurs, le terme d'indice (i,k) de la matrice ${}^t\mathrm{A}$ est a_{ki}.

2. On dira qu'une *forme bilinéaire* φ sur un $\mathbb{R}$-espace vectoriel E est *définie positive* si la forme quadratique qui lui est associée $\mathbf{x} \in E \mapsto \varphi(\mathbf{x}, \mathbf{x})$ est une forme quadratique définie positive. ◇

PROPOSITION 8.17 *Soient E un $\mathbb{R}$-espace vectoriel et q une forme quadratique sur E.*

✗ *La forme quadratique q est une forme quadratique définie si et seulement si son cône isotrope $C(q)$ est $\{\mathbf{0}\}$.*

✗ *Si la forme quadratique q est une forme quadratique positive alors son cône isotrope $C(q)$ et son noyau $N(q)$ coïncident.*

Démonstration $\trianglerighteq$ D'après la définition du cône isotrope d'une forme quadratique, il est immédiat qu'une forme quadratique est définie si et seulement si son cône isotrope est $\{\mathbf{0}\}$.

$\trianglelefteq$ Considérons une forme quadratique q sur un espace vectoriel E et notons $N(q)$ le noyau de cette forme quadratique et $C(q)$ son cône isotrope. L'inclusion $N(q) \subset C(q)$ étant toujours vraie[64], il s'agit de prouver que si la forme quadratique q est positive alors $C(q) \subset N(q)$.

Supposons donc que la forme quadratique q est positive et considérons sa décomposition en carrés de formes linéaires indépendantes (voir le théorème 8.4) $q = \sum_{i=1}^{r} \alpha_i f_i^2$ où $r \leqslant n$ et où chaque α_i est un réel non nul. On montre par l'absurde que tous les coefficients α_i sont strictement positifs : supposons que $\alpha_{i_0} < 0$ où $i_0 \in \{1, \ldots, r\}$. En vertu de la proposition 8.15, l'espace vectoriel $\cap_{i=1, i\neq i_0}^{r} \mathrm{Ker}(f_i)$ est un espace vectoriel de dimension $n-(r-1)$ alors que l'espace vectoriel $\cap_{i=1}^{r} \mathrm{Ker}(f_i)$ est un espace vectoriel de dimension $n-r$. Il existe donc un élément non nul $\mathbf{x}$ dans $\left(\cap_{i=1, i\neq i_0}^{r} \mathrm{Ker}(f_i)\right) \setminus \mathrm{Ker}(f_{i_0})$. Pour ce choix de $\mathbf{x} \neq \mathbf{0}$, on a $q(\mathbf{x}) = \sum_{i=1}^{r} \alpha_i f_i(\mathbf{x})^2 = \alpha_{i_0} f_{i_0}(\mathbf{x})^2 < 0$ ce qui contredit le fait que la forme quadratique q est positive.

Ainsi, si $\mathbf{x} \in C(q)$ alors on a $q(\mathbf{x}) = \sum_{i=1}^{r} \alpha_i f_i(\mathbf{x})^2 = 0$ où tous les coefficients α_i sont strictement positifs, ce qui implique que $f_i(\mathbf{x}) = 0$ pour chaque $i \in \{1, \ldots, r\}$. La forme polaire de q qui est $\varphi_0 : (\mathbf{x}, \mathbf{y}) \mapsto \sum_{i=1}^{r} \alpha_i f_i(\mathbf{x}) f_i(\mathbf{y})$ vérifie donc $\varphi_0(\mathbf{x}, \mathbf{y}) = 0$ pour tout $\mathbf{y} \in E$. Cela implique que $\mathbf{x} \in N(q)$. On a donc montré l'inclusion $C(q) \subset N(q)$. □

Remarque On déduit[65] de la proposition 8.17 que, si une forme bilinéaire φ sur un $\mathbb{R}$-espace vectoriel est une forme bilinéaire définie alors on a

$$\forall \mathbf{x} \in E \quad \varphi(\mathbf{x}, \mathbf{y}) = 0 \iff \mathbf{y} = \mathbf{0}.$$

Par linéarité, on en déduit que si φ est une forme bilinéaire définie alors

$$\forall \mathbf{x} \in E \quad \varphi(\mathbf{x}, \mathbf{a}) = \varphi(\mathbf{x}, \mathbf{b}) \iff \mathbf{a} = \mathbf{b}.$$ ◇

[64] Voir la remarque de la p. 457.

[65] Voir la définition du cône isotrope, définition 8.10.

DÉFINITION 8.15 (Produit scalaire)
On appelle produit scalaire *sur un $\mathbb{R}$-espace vectoriel E une forme bilinéaire symétrique définie positive.*

Pour établir qu'une forme bilinéaire φ est un produit scalaire, il faut montrer que :

- l'application φ est symétrique : $\forall (\mathbf{x}, \mathbf{y}) \in E^2 \quad \varphi(\mathbf{x}, \mathbf{y}) = \varphi(\mathbf{y}, \mathbf{x})$,
- la forme quadratique $q : \mathbf{x} \in E \mapsto \varphi(\mathbf{x}, \mathbf{x})$ est définie positive :
  - $\cdot \forall \mathbf{x} \in E \quad \varphi(\mathbf{x}, \mathbf{x}) \geqslant 0$,
  - $\cdot \forall \mathbf{x} \in E \quad \big(\varphi(\mathbf{x}, \mathbf{x}) = 0 \implies \mathbf{x} = 0\big)$.

Exemples

1. La forme bilinéaire $\varphi : (\mathbf{x}, \mathbf{y}) \in (\mathbb{R}^n)^2 \mapsto \sum_{i=1}^n x_i y_i$ est un produit scalaire sur $\mathbb{R}^n$, appelé *produit scalaire canonique*. Matriciellement, on a

$$\varphi(\mathbf{x}, \mathbf{y}) = {}^t\mathrm{X}\mathrm{Y},$$

X et Y étant les matrices colonnes ayant pour coordonnées les composantes de $\mathbf{x}$ et $\mathbf{y}$ dans la base canonique de $\mathbb{R}^n$.

2. La forme bilinéaire $\varphi : (f, g) \in \mathcal{C}([0,1], \mathbb{R})^2 \mapsto \int_0^1 f(t)g(t)\,\mathrm{d}t$ est un produit scalaire. Le fait que cette application est bilinéaire a été montré p. 439. Elle est clairement symétrique et le fait qu'elle soit définie positive a été établi p. 477.

3. Sur l'espace vectoriel $\mathcal{M}_n(\mathbb{R})$ des matrices carrées de taille n à coefficients réels, l'application φ définie par

$$\varphi : (\mathrm{A}, \mathrm{B}) \in \mathcal{M}_n(\mathbb{R}) \times \mathcal{M}_n(\mathbb{R}) \longmapsto \mathrm{Tr}({}^t\mathrm{AB}) \in \mathbb{R}$$

est un produit scalaire. On a établi p. 439 qu'il s'agit d'une forme bilinéaire et p. 477 que la forme quadratique qui lui est associée est définie positive.

Exercice 8.4 Soient $\mathcal{M}_n(\mathbb{R})$ l'espace vectoriel des matrices carrées réelles de taille n et $\mathcal{S}_n(\mathbb{R})$ le sous-espace vectoriel formé des matrices symétriques. On considère l'application φ définie sur $\mathcal{M}_n(\mathbb{R})$ par

$$\forall (\mathrm{A}, \mathrm{B}) \in \mathcal{M}_n(\mathbb{R})^2 \quad \varphi(\mathrm{A}, \mathrm{B}) = \mathrm{Tr}(\mathrm{AB})$$

où Tr désigne l'application trace[66].
1 - Montrer que φ est un produit scalaire sur $\mathcal{S}_n(\mathbb{R})$.
2 - Cette application définit-elle un produit scalaire sur $\mathcal{M}_n(\mathbb{R})$?

[66] Rappelons que la trace d'une matrice est la somme de ses éléments diagonaux, voir l'exercice 5 p. 448 du *Cours de première année*.

PROPOSITION 8.18 *Soient E un $\mathbb{R}$-espace vectoriel de dimension finie égale à n et q une forme quadratique sur E. La forme polaire de q est un produit scalaire si et seulement si sa signature est $(n, 0)$.*

Démonstration $\trianglerighteq$ Supposons que q soit un produit scalaire sur E. Puisque la forme quadratique q est positive, sa décomposition en carrés de formes linéaires indépendantes donnée par le théorème 8.4 s'écrit $q = \sum_{i=1}^{r} \alpha_i f_i^2$. De plus, on peut montrer comme dans la preuve de la proposition 8.17 que si la forme quadratique q est positive alors tous les coefficients α_i sont strictement positifs. La signature de q est donc $(r, 0)$.

Montrons en raisonnant par l'absurde que $r = n$. Supposons que $r < n$. Puisque les r formes linéaires $f_1, \ldots, f_r$ sont indépendantes, on sait d'après la proposition 8.15 que l'espace vectoriel $\cap_{i=1}^{r} \mathrm{Ker}(f_i)$ est de dimension $n - r > 0$. Par conséquent, il existe $\mathbf{x} \in \cap_{i=1}^{r} \mathrm{Ker}(f_i)$ tel que $\mathbf{x} \neq \mathbf{0}$. Ce vecteur non nul satisfait $q(\mathbf{x}) = 0$ ce qui est impossible car q est positive.

$\trianglerighteq$ Réciproquement, supposons que la signature de q soit $(n, 0)$. La décomposition de q en carrés de formes linéaires indépendantes (voir le théorème 8.4) s'écrit $q = \sum_{i=1}^{n} \alpha_i f_i^2$ où tous les coefficients α_i sont strictement positifs. Il s'ensuit que pour tout $\mathbf{x} \in E$, $q(\mathbf{x}) \geqslant 0$ (donc que la forme quadratique q est positive) et que si $q(\mathbf{x}) = 0$ alors $f_i(\mathbf{x}) = 0$ pour tout $i \in \{1, \ldots, n\}$. Puisque la famille $(f_1, \ldots, f_n)$ est une base de E^*, ceci implique que $f(\mathbf{x}) = 0$ pour toute forme linéaire $f \in E^*$. On en conclut, d'après la proposition 8.13 que $\mathbf{x} = \mathbf{0}$. □

Remarque On dit qu'une forme quadratique *permet de définir* un produit scalaire, lorsque sa forme polaire associée *est* un produit scalaire. ◇

Exemples

1. La forme quadratique q donnée sur $\mathbb{R}^4$ par

$$q(\mathbf{x}) = x_1x_2 + x_1x_3 + x_1x_4 + x_2x_3 + x_2x_4 + x_3x_4$$

admet pour écriture en somme de carrés de formes linéaires indépendantes (voir l'exemple de la p. 470) :

$$q(\mathbf{x}) = f_1(\mathbf{x})^2 - f_2(\mathbf{x})^2 - f_3(\mathbf{x})^2 - 3f_4(\mathbf{x})^2$$

où les formes linéaires f_i sont données par : $f_1(\mathbf{x}) = \frac{1}{2}(x_1 + x_2 + 2x_3 + 2x_4)$, $f_2(\mathbf{x}) = \frac{1}{2}(x_1 - x_2)$, $f_3(\mathbf{x}) = x_3 + \frac{1}{2}x_4$ et $f_4(\mathbf{x}) = \frac{1}{2}x_4$. On en déduit que la signature de q vaut $(1, 3)$. En vertu de la proposition 8.18, cette forme quadratique ne permet pas de définir un produit scalaire sur $\mathbb{R}^4$. À partir de la décomposition de q en somme de carrés de formes linéaires, on peut trouver un vecteur non nul $\mathbf{x} \in \mathbb{R}^4$ tel que $q(\mathbf{x}) < 0$ (ce qui prouve que q n'est pas une forme quadratique positive), par exemple pour $\mathbf{x} = (1, -1, 0, 0)$ on a $q(\mathbf{x}) = -1$ (plus généralement n'importe quel vecteur $\mathbf{x}$ non nul dans le noyau de l'application linéaire f_1 est tel que $q(\mathbf{x}) < 0$).

2. La forme quadratique q donnée sur $\mathbb{R}^4$ par

$$q(\mathbf{x}) = f_1(\mathbf{x})^2 + 2f_2(\mathbf{x})^2 + 2f_3(\mathbf{x})^2,$$

où f_1, f_2 et f_3 sont les trois formes linéaires de l'exemple précédent, a pour signature $(3, 0)$ et ne définit donc pas un produit scalaire sur $\mathbb{R}^4$. Bien entendu, cette forme quadratique est positive mais puisque que sa décomposition en somme de carrés de formes linéaires indépendantes ne comporte que trois termes, on peut trouver un élément $\mathbf{x} \in \mathbb{R}^4$ non nul tel que $q(\mathbf{x}) = 0$. Par exemple, le vecteur $\mathbf{x} = (1, 1, 1, -2)$ annule les trois formes linéaires f_1, f_2 et f_3 et donc annule aussi la forme quadratique q.

THÉORÈME 8.5 (Inégalité de Cauchy-Schwarz)
Pour tout produit scalaire φ sur un $\mathbb{R}$-espace vectoriel E, on a

$$\forall (\mathbf{x}, \mathbf{y}) \in E^2 \quad \varphi(\mathbf{x}, \mathbf{y})^2 \leqslant \varphi(\mathbf{x}, \mathbf{x})\ \varphi(\mathbf{y}, \mathbf{y}).$$

Démonstration Considérons un produit scalaire φ sur E et $(\mathbf{x}, \mathbf{y}) \in E^2$. L'application bilinéaire φ est étant positive on a $\varphi(\mathbf{x} + t\mathbf{y}, \mathbf{x} + t\mathbf{y}) \geqslant 0$ pour tout $t \in \mathbb{R}$. De plus, en vertu de la bilinéarité de φ on a

$$\varphi(\mathbf{x} + t\mathbf{y}, \mathbf{x} + t\mathbf{y}) = t^2\varphi(\mathbf{y}, \mathbf{y}) + 2t\varphi(\mathbf{x}, \mathbf{y}) + \varphi(\mathbf{x}, \mathbf{x}).$$

On en déduit que l'application $t \in \mathbb{R} \mapsto \varphi(\mathbf{x} + t\mathbf{y}, \mathbf{x} + t\mathbf{y})$ est une application polynomiale de degré 2 qui est à valeurs positives. Elle admet donc au plus une racine réelle et par conséquent, son discriminant Δ est négatif. Ce discriminant vaut :

$$\Delta = 4\varphi(\mathbf{x}, \mathbf{y})^2 - 4\varphi(\mathbf{y}, \mathbf{y})\varphi(\mathbf{x}, \mathbf{x})$$

et l'inégalité $\Delta \leqslant 0$ correspond exactement à l'inégalité de Cauchy-Schwarz. □

Au cours de la preuve, on peut remarquer les deux choses suivantes.

– À aucun moment de la preuve nous avons utilisé le fait que la forme bilinéaire φ est définie. L'inégalité de Cauchy-Schwarz est donc plus généralement vraie pour une forme bilinéaire symétrique et positive.
– On peut trouver les conditions sur les vecteurs $\mathbf{x}$ et $\mathbf{y}$ pour que l'inégalité de Cauchy-Schwarz devienne l'égalité

$$\varphi(\mathbf{x}, \mathbf{y})^2 = \varphi(\mathbf{x}, \mathbf{x})\varphi(\mathbf{y}, \mathbf{y}). \tag{18}$$

En effet, en reprenant les notations de la démonstration, l'égalité (18) a lieu si le discriminant Δ est nul. Ceci correspond à l'existence d'une unique racine t_0 pour l'application polynomiale $t \in \mathbb{R} \mapsto \varphi(\mathbf{x} + t\mathbf{y}, \mathbf{x} + t\mathbf{y})$:

$$\exists t_0 \in \mathbb{R} \quad \varphi(\mathbf{x} + t_0\mathbf{y}, \mathbf{x} + t_0\mathbf{y}) = 0.$$

Compte tenu du fait que l'application φ est définie, cela équivaut à

$$\exists t_0 \in \mathbb{R} \quad \mathbf{x} + t_0\mathbf{y} = \mathbf{0}.$$

SCHWARZ, Hermann-Amandus (1843, Hermsdorf - 1921, Berlin).



Schwarz fit des études de chimie et de mathématiques à Berlin. Les travaux de Schwarz sont marqués par une forte interaction entre l'analyse et la géométrie. Son mémoire de thèse avait trait aux surfaces d'aire minimale. En 1884, il résolut le problème isopérimétrique en dimension 3 qui consiste à trouver une surface d'aire minimale qui délimite un volume maximal. Schwarz réalisa également des travaux importants en analyse fonctionnelle : étude des fonctions analytiques, des fonctions holomorphes et de la théorie du potentiel.

Ainsi, l'inégalité de Cauchy-Schwarz devient une égalité si et seulement si[67] les vecteurs $\mathbf{x}$ et $\mathbf{y}$ sont liés.

L'inégalité de Cauchy-Schwarz permet, en effectuant des choix particuliers de produits scalaires, d'obtenir de nombreuses inégalités classiques[68] :

– Dans $\mathbb{R}^n$, l'inégalité de Cauchy-Schwarz pour le produit scalaire canonique (voir l'exemple p. 479) s'écrit :

$$\forall (\mathbf{x},\mathbf{y}) \in \mathbb{R}^n \times \mathbb{R}^n \quad \left(\sum_{i=1}^{n} x_i y_i\right)^2 \leqslant \left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{i=1}^{n} y_i^2\right).$$

Cette relation a été donnée dans le *Cours de première année* p. 101, proposition 3.6.

– En utilisant le produit scalaire $\varphi : (f,g) \in \mathcal{C}([0,1],\mathbb{R})^2 \mapsto \int_0^1 f(t)g(t)\,\mathrm{d}t$, l'inégalité de Cauchy-Schwarz devient

$$\forall (f,g) \in \mathcal{C}([0,1],\mathbb{R})^2 \quad \left(\int_0^1 f(t)g(t)\,\mathrm{d}t\right)^2 \leqslant \left(\int_0^1 f(t)^2\,\mathrm{d}t\right)\left(\int_0^1 g(t)^2\,\mathrm{d}t\right).$$

Cette relation a été donnée dans le *Cours de première année* p. 846, proposition 18.6.

PROPOSITION 8.19 *Soit φ est un produit scalaire sur un $\mathbb{R}$-espace vectoriel E. L'application $N : \mathbf{x} \in E \mapsto \sqrt{\varphi(\mathbf{x},\mathbf{x})}$ est une norme sur E, appelée* norme associée au produit scalaire φ.

Démonstration D'après la définition 7.3 p. 358, pour montrer que l'application N est une norme, il faut montrer que :

[67] La réciproque est immédiate : si $\mathbf{x}$ et $\mathbf{y}$ sont liés alors l'égalité $\varphi(\mathbf{x},\mathbf{y})^2 = \varphi(\mathbf{x},\mathbf{x})\varphi(\mathbf{y},\mathbf{y})$ est vraie, d'après la bilinéarité de l'application φ.

[68] Voir par exemple le *Cours de première année* p. 101 et 846.

1. $\forall \mathbf{x} \in E \quad N(\mathbf{x}) \geqslant 0$,
2. $\forall \mathbf{x} \in E \quad N(\mathbf{x}) = 0 \implies \mathbf{x} = 0$,
3. $\forall \mathbf{x} \in E \quad \forall \lambda \in \mathbb{R} \quad N(\lambda \mathbf{x}) = |\lambda| N(\mathbf{x})$,
4. $\forall (\mathbf{x}, \mathbf{y}) \in E^2 \quad N(\mathbf{x} + \mathbf{y}) \leqslant N(\mathbf{x}) + N(\mathbf{y})$.

La première condition est évidente d'après la définition de N. La deuxième résulte du fait que φ est une forme bilinéaire définie. La troisième condition découle de manière immédiate de la bilinéarité de φ. Afin de montrer la dernière condition, considérons $(\mathbf{x}, \mathbf{y}) \in E^2$. Par définition de l'application N, on a

$$N(\mathbf{x} + \mathbf{y})^2 = \varphi(\mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y}).$$

L'application φ étant bilinéaire et symétrique, on en déduit que

$$N(\mathbf{x} + \mathbf{y})^2 = N(\mathbf{x})^2 + 2\varphi(\mathbf{x}, \mathbf{y}) + N(\mathbf{y})^2.$$

En utilisant l'inégalité de Cauchy-Schwarz (voir le théorème 8.5), on a

$$\varphi(\mathbf{x}, \mathbf{y}) \leqslant \sqrt{\varphi(\mathbf{x}, \mathbf{x})}\ \sqrt{\varphi(\mathbf{y}, \mathbf{y})} = N(\mathbf{x})\, N(\mathbf{y}).$$

On en déduit que

$$N(\mathbf{x} + \mathbf{y})^2 \leqslant N(\mathbf{x})^2 + 2N(\mathbf{x})N(\mathbf{y}) + N(\mathbf{y})^2 = (N(\mathbf{x}) + N(\mathbf{y}))^2.$$

La quatrième condition est établie. □

Exercice 8.5 En utilisant l'inégalité de Cauchy-Schwarz dans le cas du produit scalaire canonique de $\mathbb{R}^n$ (voir p. 482), montrer l'inégalité :

$$\forall (x_1, \ldots, x_n) \in (\mathbb{R}_+^*)^n \qquad (x_1 + \cdots + x_n)\left(\frac{1}{x_1} + \cdots + \frac{1}{x_n}\right) \geqslant n^2,$$

puis étudier pour quelles valeurs de $(x_1, \ldots, x_n) \in (\mathbb{R}_+^*)^n$ on a égalité des deux membres de l'inégalité précédente.

EUCLIDE (vers 295 avant J.-C., Grèce).



On ne sait que très peu de choses sur la vie d'Euclide en dehors du fait qu'il ait enseigné les mathématiques à Alexandrie. Il est connu pour être le fondateur de l'École d'Alexandrie qui influença entre autres les travaux d'Archimède. Les théories d'Euclide, dont l'œuvre maîtresse est constituée des *Éléments*[69], sont une référence dans l'histoire des mathématiques. Parmi les cinq axiomes énoncés dans les *Éléments*, on trouve le célèbre postulat d'Euclide : « par tout point du plan passe une et une seule droite parallèle à une autre droite ». Cet axiome constitue le fondement de la « géométrie euclidienne ».

8.2.2 Espace euclidien

DÉFINITION 8.16

✗ *On appelle* espace préhilbertien *un $\mathbb{R}$-espace vectoriel muni d'un produit scalaire.*

✗ *On appelle* espace euclidien *un espace préhilbertien de dimension finie.*

Lorsqu'il n'y aura pas d'ambiguïté, on notera $\langle \mathbf{x}, \mathbf{y} \rangle$ le produit scalaire entre deux éléments $\mathbf{x}$ et $\mathbf{y}$, et $\|\mathbf{x}\|$ la norme associée. Un espace euclidien correspond à la donnée d'un couple $(E, \langle \cdot, \cdot \rangle)$ formé d'un $\mathbb{R}$-espace vectoriel de dimension finie et d'un produit scalaire sur cet espace.

Avec ces notations, l'inégalité de Cauchy-Schwarz (théorème 8.5) se réécrit

$$\forall (\mathbf{x}, \mathbf{y}) \in E^2 \qquad |\langle \mathbf{x}, \mathbf{y} \rangle| \leqslant \|\mathbf{x}\| \, \|\mathbf{y}\|.$$

Puisqu'à un produit scalaire $\langle \cdot, \cdot \rangle$ on peut toujours associer une norme $\| \cdot \|$ (voir la proposition 8.19), un espace euclidien $(E, \langle \cdot, \cdot \rangle)$ est toujours un espace vectoriel normé $(E, \| \cdot \|)$.

Réciproquement, on peut montrer que tout espace vectoriel normé $(E, \| \cdot \|)$ où la norme satisfait la relation suivante, appelée identité du parallélogramme,

$$\forall (\mathbf{x}, \mathbf{y}) \in E^2 \quad \|\mathbf{x} + \mathbf{y}\|^2 + \|\mathbf{x} - \mathbf{y}\|^2 = 2 \left(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 \right)$$

est un espace préhilbertien.

Le produit scalaire étant une forme bilinéaire, dans un espace préhilbertien, il existe une notion d'orthogonalité intrinsèque qui est définie par ce produit scalaire et on a le théorème suivant.

[69] On pourra consulter une version numérisée de cette œuvre sur GALLICA, la bibliothèque numérique de la Bibliothèque Nationale de France : `http://gallica.bnf.fr/ark:/12148/bpt6k110982q`

HILBERT, David (1862, Königsberg - 1943, Göttingen).



Hilbert, dont l'œuvre est immense, est unanimement reconnu comme la figure emblématique des mathématiques du 20^{e} siècle. Professeur à l'université de Göttingen, il fit de cette ville le centre des mathématiques du début du 20^{e} siècle. En août 1900, au Second Congrès International des Mathématiciens, Hilbert a exposé 23 problèmes. Ces problèmes ont depuis été au cœur de nombreuses recherches et il en reste à l'heure actuelle trois qui ne sont pas résolus[70]. Hilbert a introduit les espaces qui portent son nom dans le cadre de son travail sur des équations intégrales.

THÉORÈME 8.6 (Théorème de Pythagore)
Soient $\mathbf{x}$ *et* $\mathbf{y}$ *deux vecteurs d'un espace préhilbertien* E. *Les deux assertions suivantes sont équivalentes :*

- $\mathbf{x}$ *et* $\mathbf{y}$ *sont orthogonaux,*
- $\|\mathbf{x}+\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$.

Démonstration Dans un espace préhilbertien E, pour tout $(\mathbf{x},\mathbf{y}) \in E^2$ on a[71]

$$\|\mathbf{x}+\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + 2\langle \mathbf{x},\mathbf{y}\rangle + \|\mathbf{y}\|^2.$$

Par définition, deux vecteurs $\mathbf{x}$ et $\mathbf{y}$ sont orthogonaux si et seulement si $\langle \mathbf{x},\mathbf{y}\rangle = 0$ ce qui établit le résultat. □

Dans le cas où l'espace préhilbertien E est l'espace vectoriel $\mathbb{R}^2$ muni du produit scalaire canonique, le théorème de Pythagore s'interprète de la manière suivante[72] : « le carré de la longueur de l'hypoténuse d'un triangle rectangle est égal à la somme des carrés des deux autres longueurs des cotés du triangle ».

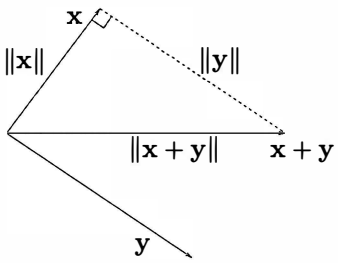
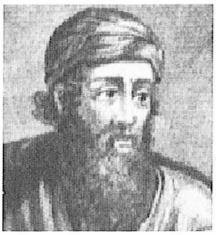


Fig. 1 Illustration du théorème de Pythagore dans le plan $\mathbb{R}^2$.

[70] Voir par exemple à ce sujet : J. GRAY, *Le défi de Hilbert, un siècle de mathématiques*, Dunod, 2003.

[71] Cette relation a été établie dans la preuve de la proposition 8.19.

PYTHAGORE (vers 500 avant J.-C., Grèce).



Mathématicien et philosophe grec, Pythagore est né dans l'île de Samos. Sa vie et son œuvre restent mal connues. Pythagore créa à Crotone dans le sud de l'Italie en 529 avant J.-C. une École où il professait les mathématiques et la philosophie, attirant autour de lui un certain nombre de disciples. Ils forment une communauté aux règles de vie très strictes dont le but est d'abord philosophique puis politique. La philosophie pythagoricienne met en avant la toute puissance du nombre qui doit pouvoir expliquer le monde. Les disciples doivent prêter serment de ne pas divulguer les connaissances qui leur sont révélées. On doit aux pythagoriciens la découverte de l'incommensurabilité de $\sqrt{2}$, autrement dit le fait que $\sqrt{2}$ est un nombre irrationnel. Pythagore fut assassiné avec de nombreux disciples dans l'incendie de l'école.

8.2.3 Procédé d'orthogonalisation de Schmidt

D'après le théorème 8.1, dans tout $\mathbb{R}$-espace vectoriel E de dimension finie et pour toute forme quadratique q sur E, il existe une base q-orthogonale. La démonstration de ce théorème établit l'existence d'une telle base mais ne fournit pas de moyen de construire une telle base. Dans le cas d'un espace préhilbertien, en plus d'une notion d'orthogonalité « naturelle » (celle associée à la forme quadratique issue du produit scalaire) nous disposons de la notion de produit scalaire. La méthode d'orthogonalisation de Schmidt est une méthode permettant, à l'aide du produit scalaire, de construire des bases orthogonales à partir d'une base quelconque.

PROPOSITION 8.20 (Orthogonalisation de Schmidt)
Soient E un espace préhilbertien et $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_p\}$ une famille libre de E. La famille $(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_p)$ définie par récurrence par

$$\varepsilon_1 = \mathbf{e}_1 \quad \text{et pour tout } k \in \{2, \ldots, p\} \quad \varepsilon_k = \mathbf{e}_k - \sum_{i=1}^{k-1} \frac{\langle \varepsilon_i, \mathbf{e}_k \rangle}{\|\varepsilon_i\|^2} \varepsilon_i \qquad (19)$$

est une famille orthogonale qui engendre le même espace vectoriel que la famille $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_p\}$.

Démonstration Montrons par récurrence sur l'entier $p \in \mathbb{N}^*$ la propriété $(\mathcal{P}_p)$ suivante :

(72) On connaissait ce résultat avant l'époque de Pythagore ; on a en particulier retrouvé des tablettes d'argile gravées par les Babyloniens vers 1800 av. J.-C. donnant les longueurs des cotés de triangles rectangles.

Pour toute famille libre à p éléments $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_p\}$, la famille $\{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_p\}$ définie par la relation (19) est orthogonale et ces deux familles engendrent le même espace vectoriel.

$\trianglerighteq$ La propriété $(\mathcal{P}_1)$ est vraie : en effet, si une famille à un élément $\{\varepsilon_1\}$ est libre alors cet élément est non nul. La famille $\{\varepsilon_1\}$ est alors une famille orthogonale puisque toute famille à un seul élément non nul est orthogonale.

$\trianglerighteq$ Supposons que la propriété $(\mathcal{P}_q)$ est vraie pour $q \in \mathbb{N}^*$ et montrons que la propriété $(\mathcal{P}_{q+1})$ est vraie. Considérons une famille à $q+1$ éléments $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_{q+1}\}$ et construisons la famille $\{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{q+1}\}$ en utilisant la relation (19). Puisque le procédé de construction est itératif, les q premiers vecteurs $\{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_q\}$ sont construit à partir des q premiers vecteurs $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_q\}$. Par hypothèse de récurrence, la famille $\{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_q\}$ est donc orthogonale et engendre le même espace vectoriel que la famille $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_q\}$.

Pour montrer que la famille $\{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{q+1}\}$ est orthogonale, montrons que le vecteur ε_{q+1} est orthogonal aux autres vecteurs ε_k pour $k \in \{1, \ldots, q\}$. On a

$$\varepsilon_{q+1} = \mathbf{e}_{q+1} - \sum_{i=1}^{q} \frac{\langle \varepsilon_i, \mathbf{e}_{q+1} \rangle}{\|\varepsilon_i\|^2} \varepsilon_i,$$

et en effectuant le produit scalaire de ε_{q+1} avec un vecteur ε_k où $k \in \{1, \ldots, q\}$, on obtient

$$\langle \varepsilon_{q+1}, \varepsilon_k \rangle = \langle \mathbf{e}_{q+1}, \varepsilon_k \rangle - \sum_{i=1}^{q} \frac{\langle \varepsilon_i, \mathbf{e}_{q+1} \rangle}{\|\varepsilon_i\|^2} \langle \varepsilon_i, \varepsilon_k \rangle.$$

Puisque la famille $\{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_q\}$ est orthogonale, les termes $\langle \varepsilon_i, \varepsilon_k \rangle$ apparaissant dans la dernière somme sont nuls excepté le terme correspondant à la valeur $i = k$. On obtient donc

$$\langle \varepsilon_{q+1}, \varepsilon_k \rangle = \langle \mathbf{e}_{q+1}, \varepsilon_k \rangle - \frac{\langle \varepsilon_k, \mathbf{e}_{q+1} \rangle}{\|\varepsilon_k\|^2} \langle \varepsilon_k, \varepsilon_k \rangle = 0,$$

ce qui permet de conclure que la famille $\{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{q+1}\}$ est orthogonale.

Il reste à prouver que la famille $\{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{q+1}\}$ engendre le même espace que la famille $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_{q+1}\}$. Par hypothèse de récurrence, les deux familles $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_q\}$ et $\{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_q\}$ engendrent le même espace. Puisque $\mathbf{e}_{q+1}$ s'exprime comme combinaison linéaire de ε_{q+1} et des autres ε_k, on en déduit que l'espace engendré par $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_{q+1}\}$ est inclus dans celui engendré par $\{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{q+1}\}$. De même, ε_{q+1} est défini (voir la proposition 8.20) comme combinaison linéaire de $\mathbf{e}_{q+1}$ et des ε_k, $k \in \{1, \ldots, q\}$, ces derniers étant eux-mêmes des combinaisons des $\mathbf{e}_k$, $k \in \{1, \ldots, q\}$; l'espace engendré par $\{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{q+1}\}$ est donc inclus dans celui engendré par $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_{q+1}\}$. Finalement on peut conclure que les espaces vectoriels engendrés par les deux familles $\{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{q+1}\}$ et $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_{q+1}\}$ sont égaux. □

Remarque Puisque les vecteurs ε_i sont non nuls, on a $\|\varepsilon_i\| \neq 0$ et on peut diviser chacun des vecteurs de $\{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_p\}$ par sa norme afin d'obtenir une famille orthonormale. ◇

SCHMIDT, Erhard (1876, Tartu, en Estonie - 1959, Berlin).



En 1905, sous la direction de David Hilbert, Schmidt obtient son doctorat à l'université de Göttingen. L'activité de recherche de Schmidt a alors trait aux équations intégrales et aux espaces dits de Hilbert. Dès 1908, il publie des articles où apparaît le procédé d'orthogonalisation donnant ainsi naissance à ce qui s'appelle aujourd'hui les opérateurs de Hilbert-Schmidt. En 1917, Schmidt obtint un poste de professeur à l'université de Berlin. Ses travaux de recherche ont progressivement évolué vers la topologie et il a en particulier étudié les liens entre l'aire d'une surface et son périmètre.

Cas de la dimension finie

Dans un $\mathbb{R}$-espace vectoriel E de dimension finie certaines familles de vecteurs ont un rôle particulier : ce sont les bases de E. Dans un espace euclidien l'orthogonalisation de Schmidt permet de construire, à partir d'une base quelconque de E, une base orthonormale.

Exemple Considérons le $\mathbb{R}$-espace vectoriel $\mathbb{R}_2[X]$ des polynômes à coefficients réels de degré inférieur à 2. Il s'agit d'un espace vectoriel de dimension 3 dont la base canonique est $(1, X, X^2)$. En tant que sous-espace vectoriel de $\mathcal{C}^0([0,1],\mathbb{R})$, l'espace vectoriel des fonctions continues sur $[0,1]$, l'espace vectoriel $\mathbb{R}_2[X]$ peut être muni du produit scalaire suivant[73] :

$$\varphi : (P,Q) \in \mathbb{R}_2[X] \times \mathbb{R}_2[X] \longmapsto \int_0^1 P(t)Q(t)\,\mathrm{d}t.$$

À partir de la base canonique de l'espace vectoriel $\mathbb{R}_2[X]$, construisons une base orthonormée de l'espace euclidien $(\mathbb{R}_2[X], \varphi)$ en utilisant le procédé d'orthogonalisation de Schmidt. On a

$$\varepsilon_1 = 1, \quad \varepsilon_2 = X - \frac{\varphi(\varepsilon_1, X)}{\varphi(\varepsilon_1,\varepsilon_1)}\varepsilon_1, \quad \varepsilon_3 = X^2 - \frac{\varphi(\varepsilon_1, X^2)}{\varphi(\varepsilon_1,\varepsilon_1)}\varepsilon_1 - \frac{\varphi(\varepsilon_2, X^2)}{\varphi(\varepsilon_2,\varepsilon_2)}\varepsilon_2.$$

Pour faciliter l'obtention de l'expression de chaque terme ε_k, il est utile de calculer la matrice associée au produit scalaire φ dans la base que l'on souhaite orthogonaliser car cette matrice contient toutes les valeurs $\varphi(X^i, X^j)$ utile pour le calcul de chaque ε_k. On a

$$\mathrm{Mat}(\varphi, (1, X, X^2)) = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{pmatrix},$$

[73] Voir l'exemple de la p. 479.

et on obtient $\varepsilon_1 = 1$, $\varepsilon_2 = X - \frac{1}{2}$ et $\varepsilon_3 = X^2 - \frac{1}{3} - (X - \frac{1}{2}) = X^2 - X + \frac{1}{3}$. Si on souhaite obtenir une base orthonormée, il suffit de diviser chaque vecteur de la base orthogonale $(\varepsilon_1, \varepsilon_2, \varepsilon_3)$ par sa norme pour le produit scalaire utilisé :

$$\mathbf{f}_1 = \frac{\varepsilon_1}{\sqrt{\varphi(\varepsilon_1, \varepsilon_1)}}, \quad \mathbf{f}_2 = \frac{\varepsilon_2}{\sqrt{\varphi(\varepsilon_2, \varepsilon_2)}}, \quad \mathbf{f}_3 = \frac{\varepsilon_3}{\sqrt{\varphi(\varepsilon_3, \varepsilon_3)}}.$$

On obtient ainsi la base orthonormale $(\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3)$ où :

$$\mathbf{f}_1 = 1, \quad \mathbf{f}_2 = 2\sqrt{3}\,(X - \tfrac{1}{2}) \quad \text{et} \quad \mathbf{f}_3 = 3\sqrt{\tfrac{10}{23}}\,(X^2 - X + \tfrac{1}{3}).$$

Remarque Considérons une base orthonormale $(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)$ d'un espace vectoriel E de dimension n. Pour tous $\mathbf{x}, \mathbf{y} \in E$ de composantes $(x_1, \ldots, x_n)$ et $(y_1, \ldots, y_n)$ dans la base $(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_p)$, on a

$$\langle \mathbf{x}, \mathbf{y} \rangle = \langle \sum_{i=1}^{n} x_i\, \varepsilon_i, \sum_{j=1}^{n} x_j\, \varepsilon_j \rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i y_j \underbrace{\langle \varepsilon_i, \varepsilon_j \rangle}_{=\delta_{ij}} = \sum_{i=1}^{n} x_i y_i$$

et $\|\mathbf{x}\|^2 = \sum_{i=1}^{n} x_i^2$. En terme matriciel, on a $\langle \mathbf{x}, \mathbf{y} \rangle = {}^t\mathrm{XY}$ et $\|\mathbf{x}\|^2 = {}^t\mathrm{XX}$, X et Y étant les matrices colonnes ayant pour coefficients les composantes de $\mathbf{x}$ et $\mathbf{y}$ dans une base orthonormale de E.

Ainsi, dans un espace euclidien de dimension n muni d'une base orthonormale, tout se passe comme dans l'espace euclidien $\mathbb{R}^n$ muni du produit scalaire canonique. ◇

Exercice 8.6 Considérons l'espace vectoriel $E = \mathcal{C}([0, \pi], \mathbb{R})$ des applications continues de $[0, \pi]$ à valeurs dans $\mathbb{R}$, muni du produit scalaire défini pour tout $(f, g) \in E^2$ par $\langle f, g \rangle = \int_0^\pi f(t)g(t)\,\mathrm{d}t$. Désignons par $\|\cdot\|$ la norme associée au produit scalaire.
1 - Que valent $\|\sin\|$, $\|\cos\|$ et $\|\mathrm{id}\|$ où sin et cos désignent les applications sinus et cosinus et où $\mathrm{id} : t \in [0, \pi] \mapsto t \in \mathbb{R}$ est l'application identité ?
2 - On note F le sous-espace vectoriel de E engendré par les applications sin, cos et id. Montrer que les trois applications sin, cos et id ne sont pas liées et en déduire que F est un espace vectoriel réel de dimension 3, dont une base est donnée par $\mathcal{B} = (\sin, \cos, \mathrm{id})$.
3 - Quelle est la matrice de l'application bilinéaire $\langle \cdot, \cdot \rangle$ dans cette base ?
4 - Soit G le sous-espace vectoriel de E engendré par les applications sin et cos. Montrer que G est un espace vectoriel réel de dimension 2.
5 - Construire une base orthonormale $\mathcal{B}_0 = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ (pour le produit scalaire $\langle \cdot, \cdot \rangle$) de l'espace vectoriel F telle que la famille $\{\mathbf{e}_1, \mathbf{e}_2\}$ soit une base orthonormale de l'espace vectoriel G.

8.3 Isométries et matrices orthogonales

8.3.1 Isométries

Dans un espace euclidien E, certains endomorphismes[74] ont la propriété de « conserver » le produit scalaire de l'espace euclidien E, ce qui signifie que le produit scalaire entre les images de deux éléments $\mathbf{x}$ et $\mathbf{y}$ de E est égal au produit scalaire entre les deux éléments $\mathbf{x}$ et $\mathbf{y}$. Ces endomorphismes sont appelés isométries.

DÉFINITION 8.17 (Isométrie)
Soient E un espace euclidien et u un endomorphisme de E. On dit que u est une isométrie *si*

$$\forall(\mathbf{x},\mathbf{y}) \in E^2 \quad \langle u(\mathbf{x}), u(\mathbf{y})\rangle = \langle \mathbf{x}, \mathbf{y}\rangle. \tag{20}$$

En choisissant $\mathbf{x} = \mathbf{y}$ on établit aisément qu'une isométrie u est une application qui conserve la norme, i.e. que $\|u(\mathbf{x})\| = \|\mathbf{x}\|$ pour tout $\mathbf{x} \in E$. On a plus précisément le résultat suivant.

PROPOSITION 8.21 *Soient E un espace euclidien et u un endomorphisme de E. L'endomorphisme u est une isométrie si et seulement si*

$$\forall \mathbf{x} \in E \quad \|u(\mathbf{x})\| = \|\mathbf{x}\|.$$

Démonstration Nous avons montré en préambule que toute isométrie conserve la norme. Montrons la réciproque ; supposons que l'endomorphisme u vérifie

$$\forall \mathbf{x} \in E \quad \|u(\mathbf{x})\| = \|\mathbf{x}\| \tag{21}$$

et considérons deux vecteurs $\mathbf{x}$ et $\mathbf{y}$ de E. En utilisant la formule de polarisation[75], on obtient

$$\langle u(\mathbf{x}), u(\mathbf{y})\rangle = \tfrac{1}{4}(\|u(\mathbf{x}) + u(\mathbf{y})\|^2 - \|u(\mathbf{x}) - u(\mathbf{y})\|^2).$$

La linéarité de l'application u et l'assertion (21) indiquent que

$$\|u(\mathbf{x}) + u(\mathbf{y})\|^2 = \|u(\mathbf{x}+\mathbf{y})\|^2 = \|\mathbf{x}+\mathbf{y}\|^2,$$
$$\text{et} \quad \|u(\mathbf{x}) - u(\mathbf{y})\|^2 = \|u(\mathbf{x}-\mathbf{y})\|^2 = \|\mathbf{x}-\mathbf{y}\|^2.$$

---

[74] Un endomorphisme d'un $\mathbb{R}$-espace vectoriel E est une application linéaire de E dans E, voir le *Cours de première année* p. 358.

[75] Plus précisément, la proposition 8.6, p. 449 indique que

$$\forall(\mathbf{a},\mathbf{b}) \in E^2 \quad \langle \mathbf{a}, \mathbf{b}\rangle = \tfrac{1}{4}(\|\mathbf{a}+\mathbf{b}\|^2 - \|\mathbf{a}-\mathbf{b}\|^2).$$

Pour obtenir l'égalité souhaitée, on prend $\mathbf{a} = u(\mathbf{x})$ et $\mathbf{b} = u(\mathbf{y})$.

Ainsi, on obtient

$$\langle u(\mathbf{x}), u(\mathbf{y})\rangle = \tfrac{1}{4}(\|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2).$$

La formule de polarisation[76] permet de conclure que $\langle u(\mathbf{x}), u(\mathbf{y})\rangle = \langle \mathbf{x}, \mathbf{y}\rangle$. □

Remarque L'ensemble des isométries d'un espace euclidien E est un groupe[77], plus précisément c'est un sous-groupe du groupe des isomorphismes de E muni de la loi de composition des endomorphisme $(\mathrm{GL}(E), \circ)$. Le groupe des isométries est noté $O(E)$. ◇

En dimension finie, une autre manière de caractériser les isométries est de considérer leur action sur une base orthonormale. On a le résultat suivant.

PROPOSITION 8.22 *Soient E un espace euclidien de dimension finie n et $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ une base orthonormale de E. Un endomorphisme u de E est une isométrie si et seulement si $(u(\mathbf{e}_1), \ldots, u(\mathbf{e}_n))$ est une base orthonormale de E.*

Démonstration ⊵ Supposons que l'endomorphisme u soit une isométrie et considérons une base orthonormale $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ de E. D'après la définition 8.17, pour tout $(i,j) \in \{1,\ldots,n\}^2$ avec $i \neq j$ on a $\langle u(\mathbf{e}_i), u(\mathbf{e}_j)\rangle = \langle \mathbf{e}_i, \mathbf{e}_j\rangle = 0$ et pour tout $i \in \{1,\ldots,n\}$ on a $\|u(\mathbf{e}_i)\| = \|\mathbf{e}_i\| = 1$. On en déduit que la famille $\{u(\mathbf{e}_1), \ldots, u(\mathbf{e}_n)\}$ est orthonormale. Puisque toute famille orthonormale est libre[78], la famille $\{u(\mathbf{e}_1), \ldots, u(\mathbf{e}_n)\}$ est libre. Étant donné que le cardinal de cette famille est égale à la dimension de E, on en déduit[79] que cette famille est une base de E.

⊵ Réciproquement, supposons que les familles $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ et $\{u(\mathbf{e}_1), \ldots, u(\mathbf{e}_n)\}$ soient deux bases orthonormales. Pour montrer que l'endomorphisme u est une isométrie, il suffit d'après la proposition 8.21 de montrer que l'application u conserve la norme. Considérons un vecteur $\mathbf{x}$ de E dont la décomposition dans la base $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ est $\mathbf{x} = \sum_{i=1}^n x_i \mathbf{e}_i$. L'application u étant linéaire, on a

$$u(\mathbf{x}) = \sum_{i=1}^n u(x_i \mathbf{e}_i) = \sum_{i=1}^n x_i\, u(\mathbf{e}_i).$$

Cette dernière égalité donne la décomposition du vecteur $u(\mathbf{x})$ dans la base $(u(\mathbf{e}_1), \ldots, u(\mathbf{e}_n))$. Les bases considérées étant toutes les deux orthonormées, on a[80] $\|\mathbf{x}\| = \sum_{i=1}^n x_i^2$ et $\|u(\mathbf{x})\| = \sum_{i=1}^n x_i^2$ ce qui permet de conclure que $\|u(\mathbf{x})\| = \|\mathbf{x}\|$. □

---

[76] Considérée cette fois-ci avec $\mathbf{a} = \mathbf{x}$ et $\mathbf{b} = \mathbf{y}$

[77] Voir la définition 2.34 p. 61 du *Cours de première année*.

[78] Voir la proposition 8.9 p. 460.

[79] Voir la proposition 8.12 p. 330 du *Cours de première année*.

[80] Voir la remarque de la p. 489.

8.3.2 Matrices orthogonales

Rappels sur le lien entre les applications linéaires et les matrices

On rappelle qu'une application linéaire est entièrement déterminée par les images des vecteurs d'une base. On peut ainsi caractériser une application linéaire par une matrice[81]. Plus précisément, on a la définition suivante.

DÉFINITION 8.18 *Soient E un $\mathbb{R}$-espace vectoriel de dimension n muni d'une base $\mathcal{B}$ et f une application linéaire de E dans E.*
On appelle matrice associée à f relativement à la base $\mathcal{B}$ *la matrice de $\mathcal{M}_n(\mathbb{R})$ dont la j-ième colonne est constituée par les composantes de l'image du j-ième vecteur de la base $\mathcal{B}$ exprimé dans la base $\mathcal{B}$.*

Cette définition permet d'affirmer que tout endomorphisme d'un $\mathbb{R}$-espace vectoriel E de dimension finie peut être représenté, relativement à une base, sous forme matricielle. Inversement, une matrice carrée est toujours la représentation d'une application linéaire. Plus précisément, on a la proposition suivante[82].

PROPOSITION 8.23 *Soit* A *une matrice carré de taille n à coefficients réels. Il existe une unique application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^n$ admettant* A *pour matrice associée relativement à la base canonique. On dit qu'une telle application linéaire est* canoniquement *associée à la matrice* A.

La définition 8.18 ainsi que la proposition 8.23 permettent[83] d'établir un isomorphisme d'espace vectoriel entre l'ensemble $\mathcal{L}_{\mathbb{R}}(E)$ des endomorphismes d'un $\mathbb{R}$-espace vectoriel E de dimension n et l'ensemble $\mathcal{M}_n(\mathbb{R})$ des matrices carrées de taille n. En dimension finie les résultats concernant les applications linéaires peuvent ainsi toujours se traduire en terme matriciel.

Considérons par exemple un endomorphisme $f \in \mathcal{L}_{\mathbb{R}}(E)$ et un couple de vecteurs $(\mathbf{x}, \mathbf{y}) \in E^2$ tel que

$$\mathbf{y} = f(\mathbf{x}).$$

Matriciellement, si on désigne par $\mathrm{A} \in \mathcal{M}_n(\mathbb{R})$ la matrice de f dans une base $\mathcal{B}$, X et Y les matrices colonnes composées des coordonnées de $\mathbf{x}$ et $\mathbf{y}$ dans la base $\mathcal{B}$ alors l'égalité $\mathbf{y} = f(\mathbf{x})$ s'écrit

$$\mathrm{Y} = \mathrm{AX}.$$

Les isométries étant des endomorphismes possédant des propriétés particulières, les matrices associées auront elles aussi des propriétés spécifiques.

---

[81] Voir le *Cours de première année*, p. 407.
[82] Voir le *Cours de première année*, p. 413.
[83] Voir le *Cours de première année*, p. 415.

DÉFINITION 8.19 (Matrice orthogonale)
On appelle matrice orthogonale *la matrice d'une isométrie dans une base orthonormale d'un espace euclidien.*

Comme pour l'ensemble $\mathrm{O}(E)$ des isométries d'un espace euclidien E, l'ensemble des matrices orthogonales de taille n, noté $\mathrm{O}_n(\mathbb{R})$, possède une structure de groupe. Plus précisément, $\mathrm{O}_n(\mathbb{R})$ est un sous-groupe du groupe des matrice inversible $(\mathrm{GL}_n(\mathbb{R}), \times)$. Il est appelé le *groupe orthogonal.*

Considérons une isométrie u d'un espace euclidien $(E, \langle\cdot,\cdot\rangle)$ de dimension n. Par définition, pour tout couple de vecteurs $(\mathbf{x}, \mathbf{y}) \in E^2$, on a

$$\langle u(\mathbf{x}), u(\mathbf{y})\rangle = \langle \mathbf{x}, \mathbf{y}\rangle.$$

Matriciellement cette relation se traduit par :

$$\langle \mathrm{UX}, \mathrm{UY}\rangle_{\mathbb{R}^n} = \langle \mathrm{X}, \mathrm{Y}\rangle_{\mathbb{R}^n}, \tag{22}$$

où les matrices colonnes X et Y correspondent aux coordonnées des vecteurs $\mathbf{x}$ et $\mathbf{y}$ dans une base orthonormale de E, U est la matrice de l'endomorphisme u dans cette même base, et $\langle\cdot,\cdot\rangle_{\mathbb{R}^n}$ désigne le produit scalaire canonique sur $\mathbb{R}^n$.

PROPOSITION 8.24 *Pour tous vecteurs* X *et* Y *de* $\mathbb{R}^n$ *et pour toute matrice* $\mathrm{M} \in \mathcal{M}_n(\mathbb{R})$ *on a*

$$\langle \mathrm{MX}, \mathrm{Y}\rangle_{\mathbb{R}^n} = \langle \mathrm{X}, {}^t\mathrm{MY}\rangle_{\mathbb{R}^n}.$$

Démonstration Considérons deux vecteurs X et Y de $\mathbb{R}^n$ et une matrice $\mathrm{M} \in \mathcal{M}_n(\mathbb{R})$. On sait[84] que $\langle \mathrm{X}, \mathrm{Y}\rangle_{\mathbb{R}^n} = {}^t\mathrm{XY}$ et ${}^t(\mathrm{MX}) = {}^t\mathrm{X}{}^t\mathrm{M}$. On en déduit aisément que $\langle \mathrm{MX}, \mathrm{Y}\rangle_{\mathbb{R}^n} = \langle \mathrm{X}, {}^t\mathrm{MY}\rangle_{\mathbb{R}^n}$. □

En pratique, étant donné une matrice $\mathrm{U} \in \mathcal{M}_n(\mathbb{R})$, il n'est pas simple, à l'aide de la définition 8.19, de déterminer si cette matrice est orthogonale ou non. On va voir que les matrices orthogonales vérifient certaines propriétés les caractérisant.

Étant donné une matrice orthogonale U, la proposition 8.24 permet d'établir que pour tout $(\mathrm{X}, \mathrm{Y}) \in (\mathbb{R}^n)^2$ on a

$$\langle \mathrm{X}, {}^t\mathrm{UUY}\rangle_{\mathbb{R}^n} = \langle \mathrm{UX}, \mathrm{UY}\rangle_{\mathbb{R}^n} = \langle \mathrm{X}, \mathrm{Y}\rangle_{\mathbb{R}^n}.$$

Le produit scalaire étant une forme bilinéaire définie, on en déduit que[85]

$$\forall \mathrm{Y} \in \mathbb{R}^n \quad {}^t\mathrm{UUY} = \mathrm{Y}.$$

[84] Voir la remarque de la p. 489.
[85] Voir la remarque de la p. 478.

En choisissant pour Y le i-ème vecteur $\mathbf{e}_i$ de la base canonique de $\mathbb{R}^n$, on établit que la i-ème colonne de la matrice carrée ${}^t\mathrm{U}\mathrm{U}$ est égale au vecteur $\mathbf{e}_i$. Autrement dit, ${}^t\mathrm{U}\mathrm{U}$ est égale à la matrice identité :

$$ {}^t\mathrm{U}\mathrm{U} = \mathrm{Id}. $$

On vient ainsi de montrer que toute matrice orthogonale U est une matrice inversible et que sa matrice inverse U^{-1} n'est autre que sa matrice transposée ${}^t\mathrm{U}$. La proposition suivante montre que cette propriété caractérise les matrices orthogonales.

PROPOSITION 8.25 *Pour* $\mathrm{U} \in \mathcal{M}_n(\mathbb{R})$ *les assertions suivantes sont équivalentes :*
(1) *la matrice* U *est une matrice orthogonale ;*
(2) *la matrice* U *est inversible et* $\mathrm{U}^{-1} = {}^t\mathrm{U}$ *;*
(3) *les colonnes (resp. les lignes) de la matrice* U *forment une base orthonormée pour le produit scalaire canonique de* $\mathbb{R}^n$.

Démonstration L'implication (1) $\Rightarrow$ (2) a été justifiée en préambule. Nous allons montrer successivement que (2) implique (3) puis que (3) implique (1) ce qui établira les équivalences par transitivité.

On considère sur $\mathbb{R}^n$ le produit scalaire canonique que l'on note $\langle \cdot, \cdot \rangle$ et on désigne par $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ la base canonique de $\mathbb{R}^n$. Cette base est orthonormée pour le produit scalaire canonique, *i.e.* $\langle \mathbf{e}_i, \mathbf{e}_j \rangle = 0$ si $i \neq j$ et $\langle \mathbf{e}_i, \mathbf{e}_i \rangle = 1$.

$\trianglerighteq$ Supposons que l'assertion (2) soit vraie, c'est-à-dire supposons que la matrice U est inversible et $\mathrm{U}^{-1} = {}^t\mathrm{U}$. On a en particulier ${}^t\mathrm{U}\mathrm{U} = \mathrm{Id}$. Les colonnes de la matrice U sont les vecteurs $\mathrm{U}\mathbf{e}_i$ où les $\mathbf{e}_i$ correspondent aux vecteurs de la base canonique de $\mathbb{R}^n$. En utilisant la proposition 8.24, on obtient

$$ \forall (i,j) \in \{1, \ldots, n\}^2 \quad \langle \mathrm{U}\mathbf{e}_i, \mathrm{U}\mathbf{e}_j \rangle = \langle \mathbf{e}_i, {}^t\mathrm{U}\mathrm{U}\mathbf{e}_j \rangle = \langle \mathbf{e}_i, \mathbf{e}_j \rangle = \delta_{ij}. $$

On en déduit que la famille $\{\mathrm{U}\mathbf{e}_1, \ldots, \mathrm{U}\mathbf{e}_n\}$ est aussi une famille orthonormée et par conséquent[86] que c'est une famille libre. Puisque[87] son cardinal est égal à la dimension de l'espace $\mathbb{R}^n$, c'est une base de l'espace vectoriel $\mathbb{R}^n$. Ainsi, les colonnes de la matrice U forment une base orthonormale de $\mathbb{R}^n$. Le même raisonnement avec la matrice ${}^t\mathrm{U}$ à la place de U montre que les lignes de U forment aussi une base orthonormale de $\mathbb{R}^n$. On a donc établit que (2) $\Rightarrow$ (3).

$\trianglerighteq$ Supposons que l'assertion (3) soit vraie. Pour montrer que la matrice U est une matrice orthogonale, d'après la proposition 8.21, on peut montrer que

[86] Voir la proposition 8.9 p. 460.
[87] Voir la proposition 8.11 p. 330 du *Cours de première année*.

pour tout vecteur $X \in \mathbb{R}^n$ on a $\|UX\|^2 = \|X\|^2$. La décomposition d'un vecteur X de $\mathbb{R}^n$ dans la base canonique $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ sous la forme $X = \sum_{i=1}^n x_i \mathbf{e}_i$ implique, par linéarité, que l'on a

$$UX = U\left(\sum_{i=1}^n x_i \mathbf{e}_i\right) = \sum_{i=1}^n x_i \, (U\mathbf{e}_i).$$

Les familles $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ et $\{U\mathbf{e}_1, \ldots, U\mathbf{e}_n\}$ étant toute les deux orthonormales, on en déduit[88] que $\|X\|^2 = \sum_{i=1}^n x_i^2$ et que $\|UX\|^2 = \sum_{i=1}^n x_i^2$. On en conclut que $\|UX\|^2 = \|X\|^2$. □

Inverser une matrice n'est en général pas très « simple »[89], par contre, calculer la transposée d'une matrice est « simple »[90]. Un des avantages des matrices orthogonales est que, d'après la proposition 8.25, leurs inverses sont leurs transposées.

Exemple Considérons la matrice

$$U = \begin{pmatrix} 2/\sqrt{6} & 0 & 1/\sqrt{3} \\ 1/\sqrt{6} & 1/\sqrt{2} & -1/\sqrt{3} \\ -1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \end{pmatrix}.$$

Il s'agit d'une matrice orthogonale car ses colonnes C_1, C_2 et C_3 forment une base orthonormée de $\mathbb{R}^3$ pour le produit scalaire canonique puisque :

$$\langle C_1, C_2 \rangle = 0, \quad \langle C_2, C_3 \rangle = 0, \quad \langle C_3, C_1 \rangle = 0,$$
$$\|C_1\| = 1, \quad \|C_2\| = 1 \quad \text{et} \quad \|C_3\| = 1.$$

On en déduit d'après la proposition 8.25 que la matrice U est une matrice inversible et que son inverse est donné par sa transposée :

$$U^{-1} = \begin{pmatrix} 2/\sqrt{6} & 1/\sqrt{6} & -1/\sqrt{6} \\ 0 & 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{3} & -1/\sqrt{3} & 1/\sqrt{3} \end{pmatrix}.$$

---

[88] Voir la remarque p. 489.

[89] Le terme « simple » fait référence ici à la complexité des algorithmes d'inversion d'une matrice : pour calculer l'inverse d'une matrice carrée (inversible) de taille n « à la main » il faut un nombre d'opérations (additions, multiplications) de l'ordre de n^3.

[90] Pour calculer la transposée d'une matrice carré de taille n, aucune opération (addition, multiplication...) n'est nécessaire, il suffit seulement d'intervertir certains coefficients.

Exercice 8.7 Montrer que la matrice suivante est inversible et calculer son inverse :

$$\mathrm{A} = \frac{1}{3}\begin{pmatrix} 2 & 2 & 1 \\ -1 & 2 & -2 \\ -2 & 1 & -2 \end{pmatrix}.$$

Nous avons indiqué qu'il y avait une correspondance (un isomorphisme) entre l'espace vectoriel $\mathcal{M}_n(\mathbb{R})$ des matrices carrées de taille n à coefficients réels et l'espace vectoriel $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ des applications linéaires de $\mathbb{R}^n$ dans $\mathbb{R}^n$. Il est donc possible de traduire en terme d'applications linéaires les propriétés liées aux matrices orthogonales.

DÉFINITION 8.20 *Soient u une isométrie d'un espace euclidien E et* U *la matrice de u dans une base orthonormale de E.*

✗ *On appelle* endomorphisme adjoint *de u, et on note ${}^t u$, l'endomorphisme associé à la matrice* ${}^t\mathrm{U}$.

✗ *On dit qu'un sous-ensemble F de E est stable par l'endomorphisme u (ou par la matrice* U*) si pour tout $\mathbf{x} \in F$ on a $u(\mathbf{x}) \in F$.*

Remarque On sait que, pour une matrice $\mathrm{A} \in \mathcal{M}_n(\mathbb{R})$ on a ${}^t({}^t\mathrm{A}) = \mathrm{A}$. D'après la définition précédente, on en déduit que pour isométrie u d'un espace euclidien E, on a ${}^t({}^t u) = \mathrm{u}$. ◇

On a le lien suivant entre les sous-espaces stables par un endomorphisme et ceux stables par l'endomorphisme adjoint.

PROPOSITION 8.26 *Soient u une isométrie d'un espace euclidien E et F un sous-ensemble de E. Si F est stable par l'endomorphisme u alors $F^\perp$ est stable par l'endomorphisme adjoint ${}^t u$.*

Démonstration Considérons une isométrie u de E et un sous-ensemble F de E stable par u. Pour montrer que $F^\perp$ est stable par ${}^t u$, il faut prouver que

$$\forall \mathbf{y} \in E \qquad \Big(\mathbf{y} \in F^\perp \implies {}^t u(\mathbf{y}) \in F^\perp\Big).$$

Soit $\mathbf{y} \in F^\perp$. Pour tout $\mathbf{z} \in F$, formons le produit scalaire $\langle {}^t u(\mathbf{y}), \mathbf{z}\rangle$. Matriciellement, en notant U la matrice de l'endomorphisme u dans une base donnée de E, Y et Z les vecteurs colonnes représentant $\mathbf{y}$ et $\mathbf{z}$ dans la même base, on a[(91)]

$$\langle {}^t u(\mathbf{y}), \mathbf{z}\rangle = \langle {}^t\mathrm{UY}, \mathrm{Z}\rangle_{\mathbb{R}^n}.$$

[(91)] L'entier n correspond à la dimension de l'espace vectoriel E.

En vertu des propriétés du produit scalaire vis-à-vis de la transposition, voir la proposition 8.24 p. 493, on en déduit que

$$\langle {}^t u(\mathbf{y}), \mathbf{z}\rangle = \langle \mathrm{Y}, \mathrm{UZ}\rangle_{\mathbb{R}^n} = \langle \mathbf{y}, u(\mathbf{z})\rangle.$$

Puisque F est stable par u et que $\mathbf{z} \in F$, on a $u(\mathbf{z}) \in F$. De plus, comme $\mathbf{y} \in F^\perp$, on en déduit que

$$\langle {}^t u(\mathbf{y}), \mathbf{z}\rangle = \langle \mathbf{y}, u(\mathbf{z})\rangle = 0.$$

Ceci prouve que ${}^t u(\mathbf{y}) \in F^\perp$. □

Remarque La proposition 8.25 peut se traduire en terme d'applications linéaires de la façon suivante : u est une isométrie si et seulement si u est inversible et $u^{-1} = {}^t u$. ◇

PROPOSITION 8.27
Le déterminant d'une matrice orthogonale vaut $+1$ *ou* -1.

Démonstration Considérons une matrice orthogonale U. Puisque ${}^t\mathrm{UU} = \mathrm{Id}$, on a $\det({}^t\mathrm{UU}) = \det({}^t\mathrm{U})\det(\mathrm{U}) = 1$. Puisque le déterminant d'une matrice est égale au déterminant de sa transposée[92], on obtient $(\det(\mathrm{U}))^2 = 1$. La matrice U étant réelle, son déterminant est un nombre réel ; on en déduit que $\det(\mathrm{U}) = 1$ ou que $\det(\mathrm{U}) = -1$. □

Remarque L'ensemble des matrices orthogonales de déterminant $+1$ est noté $\mathrm{SO}_n(\mathbb{R})$; c'est un sous-groupe[93] de $(\mathrm{O}_n(\mathbb{R}), \times)$ qui est appelé *groupe spécial orthogonal.* ◇

Cas de la dimension 2

La troisième assertion de la proposition 8.25 indique que toutes les isométries du plan $\mathbb{R}^2$ s'écrivent matriciellement sous la forme

$$\mathrm{U} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{avec} \quad \begin{cases} a^2 + b^2 = 1 \\ c^2 + d^2 = 1 \\ ab + cd = 0 \end{cases}. \tag{23}$$

Les deux premières égalités de (23) indiquent qu'il existe $(\theta, \phi) \in [0, 2\pi[^2$ tel que $a = \cos(\theta)$, $b = \sin(\theta)$, $c = \cos(\phi)$ et $d = \sin(\phi)$. En utilisant la dernière

[92] Voir le *Cours de première année* p. 479.

[93] Rappelons, voir p. 493, que $\mathrm{O}_n(\mathbb{R})$ désigne l'ensemble des matrices orthogonales qui muni de la multiplication des matrices est un groupe.

égalité de (23) on a $\phi = \theta \pm \pi/2 \pmod{2\pi}$. La matrice U est donc l'une des deux formes suivantes :

$$R_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \quad \text{ou} \quad S_\theta = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix}.$$

Dans le premier cas, on a $\det(R_\theta) = +1$ alors que dans le second cas, on a $\det(S_\theta) = -1$. Notons que les matrices S_θ possède toujours deux valeurs propres réelles distinctes (1 et -1). La matrice S_θ est donc diagonalisable[94] et s'écrit dans une base de vecteurs propres

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

8.3.3 Description géométrique des isométries

On dispose maintenant de deux points de vue concernant les isométries : le point de vue des applications linéaires où une isométrie est une application linéaire ne « changeant » pas le produit scalaire, et le point de vue matricielle où la matrice d'une isométrie est une matrice orthogonale. On va s'intéresser au point de vue géométrique, tout d'abord dans le cas du plan $\mathbb{R}^2$ puis dans le cas plus général d'un espace vectoriel de dimension finie.

Du point de vue des applications linéaires, la proposition 8.27 permet de distinguer les isométries dont le déterminant[95] est $+1$ de celles dont le déterminant est -1.

Définition 8.21 *Soit u une isométrie d'un espace euclidien.*

✗ *On dit que u est une* rotation *(ou une* isométrie directe*) si son déterminant vaut $+1$,*

✗ *On dit que u est une* isométrie indirecte *si son déterminant vaut -1.*

Contrairement à l'ensemble des isométries directes, l'ensemble des isométries indirectes ne possède pas de structure de groupe. En effet, si deux matrices ont pour déterminant -1 alors leur produit aura pour déterminant $+1$: le produit de deux isométries indirectes n'est pas une isométrie indirecte.

[94] En effet, si une matrice de taille n possède exactement n valeurs propres distinctes alors elle est diagonalisable, voir le *Cours de première année*, Corollaire 12.2, p. 537.

[95] Le déterminant d'un endomorphisme d'un espace vectoriel E de dimension finie est le déterminant d'une matrice représentant cette application dans une base de E, on démontre, voir le *Cours de première année* p. 526, que le déterminant ne dépend pas de la base de E choisie.

Cas de la dimension $n = 2$

On déduit de l'étude effectuée à la p. 497, le résultat suivant.

THÉORÈME 8.7 *Les rotations du plan ont pour matrice*

$$\mathrm{R}_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \quad \textit{avec} \quad \theta \in [0, 2\pi[,$$

et les isométries indirectes du plan ont pour matrice

$$\mathrm{S} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Géométriquement, les isométries indirectes correspondent à des symétries par rapport à une droite vectorielle. Cette droite peut être déterminée en trouvant l'espace propre associé à la valeur propre $+1$.

Exemple Considérons l'endomorphisme u de l'espace vectoriel $\mathbb{R}^2$ dont la matrice dans la base canonique $(\mathbf{e}_1, \mathbf{e}_2)$ de $\mathbb{R}^2$ est

$$\mathrm{A} = \mathrm{Mat}(u, (\mathbf{e}_1, \mathbf{e}_2)) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

On vérifie aisément que la matrice A est orthogonale (par exemple en vérifiant que les deux colonnes de cette matrice forment une base orthonormale de $\mathbb{R}^2$). Par conséquent, l'endomorphisme u est une isométrie. Puisque le déterminant de A vaut -1, il s'agit d'une isométrie indirecte.

Recherchons une base de $\mathbb{R}^2$ dans laquelle la matrice de u est égale à la matrice $\mathrm{S} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Pour cela, on diagonalise la matrice A. Un vecteur propre associé à la valeur propre 1 est le vecteur $\varepsilon_1 = (1, \sqrt{2} - 1)$ et un vecteur propre associé à la valeur propre -1 est le vecteur $\varepsilon_2 = (1, -\sqrt{2} - 1)$. Ainsi, dans la base $(\varepsilon_1, \varepsilon_2)$, la matrice de l'endomorphisme u vaut S :

$$\mathrm{Mat}(u, (\varepsilon_1, \varepsilon_2)) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Si on désigne par x_1 et x_2 les deux composantes d'un vecteur $\mathbf{x} \in \mathbb{R}^2$ dans la base $(\varepsilon_1, \varepsilon_2)$ alors on a

$$u(\mathbf{x}) = u(x_1\varepsilon_1 + x_2\varepsilon_2) = x_1 u(\varepsilon_1) + x_2 u(\varepsilon_2) = x_1\varepsilon_1 - x_2\varepsilon_2.$$

L'application u est la symétrie par rapport à la droite vectorielle dirigée par le vecteur ε_1.

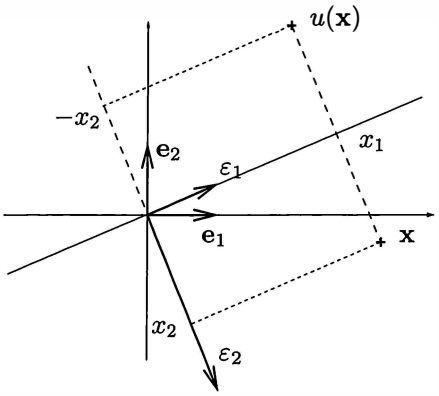


Fig. 2 Symétrie par rapport à une droite vectorielle.

Cas de la dimension $n > 2$

Dans le cas d'un espace euclidien de dimension $n > 2$, on a le résultat suivant.

THÉORÈME 8.8 *Soit u une isométrie d'un espace euclidien E. Il existe une base orthonormée de E dans laquelle la matrice de u est de la forme*

$$\mathrm{U} = \begin{pmatrix} \varepsilon_1 & & & & 0 \\ & \ddots & & & \\ & & \varepsilon_r & & \\ & & & R_{\theta_1} & \\ & 0 & & & \ddots \\ & & & & R_{\theta_s} \end{pmatrix}$$

avec $r \in \mathbb{N}$, $s \in \mathbb{N}$ et où pour tout $i \in \{1, \ldots, r\}$, $\varepsilon_i \in \{-1, +1\}$ et pour tout $j \in \{1, \ldots, s\}$, $\theta_j \in [0, 2\pi[$.

Démonstration La démonstration est basée sur une récurrence sur la dimension n de l'espace euclidien E dont le principe est le suivant.

$\trianglerighteq$ Dans le cas $n = 1$, le résultat est immédiat puisque les seules isométries de $\mathbb{R}$ sont les applications $x \in \mathbb{R} \mapsto x$ et $x \in \mathbb{R} \mapsto -x$ dont les matrices dans une base de $\mathbb{R}$ sont respectivement (1) et (-1).

$\trianglerighteq$ Supposons que le théorème soit vrai pour tout espace euclidien de dimension inférieure strictement à n. Considérons une isométrie u d'un espace euclidien de dimension n. On distingue deux cas.

– Si u possède au moins une valeur propre réelle alors on note $\mathbf{e}_1$ un vecteur propre associé à cette valeur propre. On remarque[(96)] que la droite $\mathbb{R}\mathbf{e}_1$ est stable

[(96)] C'est-à-dire que si $\mathbf{x} \in \mathbb{R}\mathbf{e}_1$ alors $u(\mathbf{x}) \in \mathbb{R}\mathbf{e}_1$, voir la définition 8.20.

par u. D'après la proposition 8.26, on en déduit le sous-espace $H = (\mathbb{R}\mathbf{e}_1)^\perp$ est stable par ${}^t u$. La dimension de H est strictement inférieure à n et on applique l'hypothèse de récurrence à l'endomorphisme ${}^t u : H \to H$.

– Si u ne possède pas de valeur propre réelle alors on exhibe un plan vectoriel P stable par u en considérant l'endomorphisme[97] $f = u + u^{-1}$. Puisque u est une isométrie on a ${}^t u = u^{-1}$ et la matrice associée à l'endomorphisme f est une matrice symétrique. On verra à la proposition 8.28 que tout endomorphisme symétrique admet une valeur propre réelle. Notons $\mathbf{e}_1$ un vecteur propre de f associé à cette valeur propre λ ; on a

$$(u + u^{-1})(\mathbf{e}_1) = u(\mathbf{e}_1) + u^{-1}(\mathbf{e}_1) = \lambda \mathbf{e}_1. \tag{24}$$

Notons $\mathbf{e}_2 = u(\mathbf{e}_1)$; en appliquant l'endomorphisme u à chaque membre de l'égalité (24) on obtient $u(\mathbf{e}_2) + \mathbf{e}_1 = \lambda \mathbf{e}_2$. On a donc $u(\mathbf{e}_1) = \mathbf{e}_2$ et $u(\mathbf{e}_2) = \lambda \mathbf{e}_2 - \mathbf{e}_1$ et on en déduit que le plan vectoriel P engendré par les vecteurs $\mathbf{e}_1$ et $\mathbf{e}_2$ est stable par u. L'endomorphisme u restreint au plan P, $u : P \to P$ est une isométrie et d'après l'étude en dimension 2, on sait[98] qu'il s'agit d'une rotation. D'après la proposition 8.26, l'endomorphisme ${}^t u$ est stable par $P^\perp$ et la dimension de $P^\perp$ est strictement inférieure à n. On peut donc appliquer l'hypothèse de récurrence à l'endomorphisme ${}^t u : P^\perp \to P^\perp$. □

Le théorème 8.8 n'exclut pas d'avoir $r = 0$ ou $s = 0$. Autrement dit, une isométrie d'un espace euclidien peut avoir pour matrice

$$\mathrm{U} = \begin{pmatrix} \varepsilon_1 & & \\ & \ddots & \\ & & \varepsilon_r \end{pmatrix} \quad \text{ou} \quad \mathrm{U} = \begin{pmatrix} R_{\theta_1} & & \\ & \ddots & \\ & & R_{\theta_s} \end{pmatrix}.$$

On notera que le second cas ne peut se produire que si la dimension de E est paire.

Cas de la dimension $n = 3$

Si E est un espace de dimension 3, le théorème 8.8 prend la forme suivante.

THÉORÈME 8.9 *Soit u une isométrie d'un espace euclidien E de dimension 3. Il existe une base orthonormée de E dans laquelle la matrice de u a l'une des formes suivantes :*

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad \textit{ou} \quad \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \textit{ou} \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{pmatrix}$$

avec $\theta \in [0, 2\pi[$.

[97] L'endomorphisme u étant une isométrie, il est en particulier inversible.

[98] Cette isométrie ne peut pas être une isométrie indirecte sinon elle posséderait une valeur propre réelle.

Dans le théorème 8.9, les deux premiers cas correspondent aux isométries indirectes car le déterminant de la matrice vaut -1. Le dernier cas correspond aux rotations (le déterminant de la matrice vaut $+1$).

En pratique, lorsque l'on dispose d'une rotation en dimension 3, c'est-à-dire d'une isométrie de $\mathbb{R}^3$ de déterminant $+1$, on ne dispose pas toujours d'une base dans laquelle cette rotation a l'expression donnée par le théorème 8.9. Néanmoins, on peut déterminer l'angle de cette rotation (la valeur θ dans le théorème 8.9) en calculant la trace de l'isométrie[(99)] ; on a $\mathrm{Tr}(u) = 2\cos(\theta)+1$, ce qui permet d'obtenir la valeur de $\cos(\theta)$. On détermine l'angle θ en connaissant le signe de $\sin(\theta)$. Celui-ci est donné par le signe du déterminant de la famille de trois vecteurs $(\mathbf{e}_1, \mathbf{x}, u(\mathbf{x}))$ où $\mathbf{e}_1$ est le vecteur propre de u associé à la valeur propre 1 (ce vecteur est appelé l'axe de la rotation) et $\mathbf{x}$ un vecteur non colinéaire à l'axe de la rotation.

Exemple Considérons la matrice

$$\mathrm{U} = \begin{pmatrix} 2/\sqrt{6} & 0 & 1/\sqrt{3} \\ 1/\sqrt{6} & 1/\sqrt{2} & -1/\sqrt{3} \\ -1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \end{pmatrix}.$$

On a vu p. 495 que la matrice U est une matrice orthogonale. On a $\det(\mathrm{U}) = 1$ ce qui montre que la matrice U est la matrice d'une rotation.

L'axe de cette rotation est l'ensemble des vecteurs invariants par cette rotation, autrement dit l'axe de la rotation correspond à l'espace propre associé à la valeur propre $+1$ (un vecteur propre $\mathbf{e}_1$ associé à la valeur propre 1 vérifie $\mathrm{U}\mathbf{e}_1 = \mathbf{e}_1$). En résolvant le système linéaire $\mathrm{U}\mathbf{x} = \mathbf{x}$, on établit que ce vecteur propre est

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ (\sqrt{2}+1)(\sqrt{3}-\sqrt{2}) \\ \sqrt{3}-\sqrt{2} \end{pmatrix}.$$

$\trianglerighteq$ Pour déterminer l'angle θ de la rotation, on utilise la trace de la matrice U ; on a $\mathrm{Tr}(\mathrm{U}) = \frac{2+\sqrt{2}+\sqrt{3}}{\sqrt{6}} = 2\cos(\theta)+1$ d'où $\cos(\theta) = \frac{2+\sqrt{2}+\sqrt{3}-\sqrt{6}}{2\sqrt{6}}$. Comme le vecteur $\mathbf{x} = (1,0,0)$ n'est pas colinéaire à $\mathbf{e_1}$, le signe de $\sin(\theta)$ est celui du déterminant de la famille $\{\mathbf{e}_1, \mathbf{x}, u(\mathbf{x})\}$. On a[(100)]

$$\det(\mathbf{e}_1, \mathbf{x}, u(\mathbf{x})) = \begin{vmatrix} 1 & 1 & 2/\sqrt{6} \\ (\sqrt{2}+1)(\sqrt{3}-\sqrt{2}) & 0 & 1/\sqrt{6} \\ \sqrt{3}-\sqrt{2} & 0 & -1/\sqrt{6} \end{vmatrix} = \frac{1+\sqrt{2}}{\sqrt{3}}(\sqrt{3}-\sqrt{2}).$$

(99) La trace d'un endomorphisme est la trace de n'importe quelle matrice représentant cet endomorphisme.

(100) La valeur de ce déterminant est obtenue rapidement en développant selon la deuxième colonne.

Ce déterminant étant positif, on a $\sin(\theta) > 0$ donc $\theta \in]0, \pi[$ et par conséquent la relation $\cos(\theta) = \frac{2+\sqrt{2}+\sqrt{3}-\sqrt{6}}{2\sqrt{6}}$ implique $\theta = \arccos(\frac{2+\sqrt{2}+\sqrt{3}-\sqrt{6}}{(2\sqrt{6})}) \approx 56,6°$.

8.4 Diagonalisation des matrices symétriques réelles

Rappel sur la diagonalisation

De manière générale, diagonaliser une matrice $A \in \mathcal{M}_n(\mathbb{R})$, c'est trouver une matrice semblable à A qui soit diagonale. Plus précisément, une matrice A est *diagonalisable* s'il existe une matrice inversible $P \in GL_n(\mathbb{R})$ et une matrice diagonale $D \in \mathcal{M}_n(\mathbb{R})$ telles que $D = P^{-1}AP$. *Diagonaliser* A, c'est trouver D et P.

En terme d'application linéaire, un endomorphisme f d'un $\mathbb{R}$-espace vectoriel E de dimension finie est dit *diagonalisable* sur $\mathbb{R}$ s'il existe une base de E formée de vecteurs propres de f. *Diagonaliser* f, c'est trouver une telle base.

Tous les endomorphismes ne sont pas diagonalisables. En effet, on montre[101] qu'un endomorphisme d'un $\mathbb{R}$-espace vectoriel E est diagonalisable si et seulement si les deux conditions suivantes sont satisfaites :

1. le nombre de ses valeurs propres (comptées avec leur multiplicité) est égal à la dimension de E ;
2. la dimension de chacun des sous-espaces propres est égale à l'ordre de multiplicité de la valeur propre correspondante.

Il existe des endomorphismes pour lesquels l'une ou l'autre de ces deux conditions (ou les deux à la fois) n'est pas vérifiée. En particulier, on sait que le nombre de valeurs propres correspond au nombre de racines réelles du polynôme caractéristique. Ce polynôme étant de degré égal à n (la dimension de E), on sait que le nombre de racine ne peut excéder n mais rien n'affirme que ce polynôme a exactement n racines réelles.

Nous allons montrer dans cette partie que toutes les matrices symétriques à coefficients réels sont diagonalisables. De plus, nous établirons que les matrices de passage P qui permettent de diagonaliser les matrices symétriques réelles peuvent être choisies orthogonales. Commençons par étudier quelques propriétés des matrices symétriques réelles.

PROPOSITION 8.28 *Toutes les valeurs propres d'une matrice symétrique à coefficients réels sont réelles.*

[101] Voir le *Cours de première année* p. 538

Démonstration Toute matrice à coefficients réels admet au moins [102] une valeur propre dans $\mathbb{C}$. Considérons $\lambda \in \mathbb{C}$ une valeur propre complexe d'une matrice symétrique réelle $M \in \mathcal{M}_n(\mathbb{R})$. Afin de montrer que cette valeur propre est réelle, on va montrer qu'elle est égale à son conjugué, c'est-à-dire que $\lambda = \overline{\lambda}$.

Soit $Z \in \mathbb{C}^n \setminus \{0\}$ un vecteur propre associé à la valeur propre λ ; on a

$$MZ = \lambda Z. \tag{25}$$

D'une part, en multipliant l'égalité (25) par ${}^t\overline{Z}$ à gauche, on obtient

$${}^t\overline{Z}MZ = \lambda\ {}^t\overline{Z}Z. \tag{26}$$

D'autre part, en transposant l'égalité (25), on a ${}^tZ{}^tM = \lambda {}^tZ$. Puisque la matrice M est symétrique, on en déduit ${}^tZM = \lambda {}^tZ$. En prenant le conjugué de cette égalité, on a ${}^t\overline{Z}\,\overline{M} = \overline{\lambda}\ {}^t\overline{Z}$. La matrice M étant réelle, on obtient ${}^t\overline{Z}M = \overline{\lambda}{}^t\overline{Z}$. Enfin, en multipliant cette égalité par Z à droite, on a

$${}^t\overline{Z}MZ = \overline{\lambda}\ {}^t\overline{Z}Z. \tag{27}$$

Les relations (26) et (27) permettent d'affirmer que $\lambda\ {}^t\overline{Z}Z = \overline{\lambda}{}^t\overline{Z}Z$. Puisque $Z = (z_1, \ldots, z_n)$ est non nul, on a ${}^t\overline{Z}Z = |z_1|^2 + \cdots + |z_n|^2 \neq 0$. On en déduit donc que $\lambda = \overline{\lambda}$, autrement dit $\lambda \in \mathbb{R}$. □

Proposition 8.29 *Les sous-espaces propres d'une matrice symétrique à coefficients réels sont deux à deux orthogonaux* [103].

Démonstration Soient λ et μ deux valeurs propres distinctes [104] de la matrice M ; d'après la proposition 8.28, ces valeurs propres sont réelles. Considérons deux sous-espaces propres E_λ et E_μ associés respectivement aux deux valeurs λ et μ. Pour montrer que E_λ est orthogonal à E_μ, il faut montrer que pour tous $X \in E_\lambda$ et $Y \in E_\mu$ on a $\langle X, Y\rangle_{\mathbb{R}^n} = 0$, c'est-à-dire ${}^tYX = 0$.
La matrice M étant symétrique on a ${}^tYMX = {}^t({}^tXMY)$. Or, par définition d'un sous-espace propre, on a $MX = \lambda X$ et $MY = \mu Y$. Cela implique que $\lambda {}^tYX = \mu {}^t({}^tXY)$, c'est-à-dire que $(\lambda - \mu){}^tYX = 0$. Puisque $\lambda \neq \mu$, on en conclut que ${}^tYX = 0$. □

[102] En effet, par définition, une valeur propre est une racine du polynôme caractéristique, et on sait que tout polynôme réel admet au moins une racine dans $\mathbb{C}$, voir le théorème de d'Alembert-Gauss énoncé dans le *Cours de première année* p. 248.

[103] L'orthogonalité fait ici référence au produit scalaire canonique de $\mathbb{R}^n$, n étant la taille de la matrice considérée.

[104] S'il n'existe qu'une seule valeur propre, la proposition est immédiatement satisfaite puisqu'il n'y a qu'un seul sous-espace propre.

THÉORÈME 8.10 *Toute matrice symétrique réelle est diagonalisable dans une base orthonormale. Autrement dit, si* $M \in \mathcal{S}_n(\mathbb{R})$ *alors il existe une matrice orthogonale* $P \in O_n(\mathbb{R})$ *telle que* $P^{-1}MP$ *soit une matrice diagonale réelle.*

Démonstration On procède par récurrence sur la taille n de la matrice M. Si $n = 1$ alors M est diagonale et le résultat est vérifié. Supposons le théorème démontré pour toutes les matrices symétriques réelles de taille $n-1$ et montrons qu'il est vrai pour toutes les matrices symétriques réelles de taille n.
Soit M une matrice symétrique réelle de taille n. On sait d'après la proposition 8.28 que M possède au moins une valeur propre réelle. Notons $\lambda_1 \in \mathbb{R}$ cette valeur propre et $\mathbf{v} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ un vecteur propre associé. Le vecteur $\mathbf{e}_1 = \mathbf{v}/\|\mathbf{v}\|$ est aussi un vecteur propre associé à λ_1 ; il est de norme 1 et la droite $\mathbb{R}\mathbf{e}_1$ est stable[(105)] par M. D'après la proposition 8.26, on en déduit que $H = (\mathbb{R}\mathbf{e}_1)^\perp$ est stable par ${}^tM$. La matrice M étant symétrique, il s'ensuit que l'espace vectoriel H est stable par M.
En notant u l'endomorphisme associé à la matrice M, on a donc montré que u est un endomorphisme de H sur H. Puisque $\dim H = n-1$, on peut appliquer l'hypothèse de récurrence à la matrice de l'endomorphisme $u : H \to H$. Il existe une base orthonormale $(\mathbf{e}_2, \ldots, \mathbf{e}_n)$ de H dans laquelle $u|_H$ est diagonale, c'est-à-dire telle que pour tout $i \in \{2, \ldots, n\}$ on a $u(\mathbf{e}_i) = \lambda_i \mathbf{e}_i$ avec $\lambda_i \in \mathbb{R}$. En complétant cette base avec le vecteur $\mathbf{e}_1$, on obtient une base de $\mathbb{R}^n$ qui convient puisque d'une part on a

$$\forall i \in \{1, \ldots, n\} \quad \exists \lambda_i \in \mathbb{R} \quad u(\mathbf{e}_i) = \lambda_i \mathbf{e}_i,$$

ce qui exprime le fait que dans cette base $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ la matrice de u est diagonale, et d'autre part la base $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ est une base orthonormale puisque la famille $\{\mathbf{e}_2, \ldots, \mathbf{e}_n\}$ est orthogonale et que le vecteur $\mathbf{e}_1$ de norme 1 et orthogonale à H, donc à la famille $\{\mathbf{e}_2, \ldots, \mathbf{e}_n\}$. □

Ne pas confondre ce résultat avec celui énoncé dans la proposition 8.11 qui dit que toute matrice symétrique réelle est congrue à une matrice diagonale. Ici, le résultat est plus fort : le théorème 8.10 indique que la matrice de passage est orthogonale.

Le résultat énoncé dans le théorème 8.10 admet des corollaires intéressants. En effet, le fait d'affirmer que toute matrice symétrique réelle se diagonalise dans une base orthonormale laisse libre choix au sens donné au mot « orthonormale ». Plus précisément, on ne spécifie pas par rapport à quel produit scalaire la base est orthonormale. Prenons ainsi deux formes quadratiques q et q' telles que q est définie positive (et permet donc de définir un produit scalaire). Le théorème 8.10 affirme qu'il existe une base $\mathcal{B}_0$ qui est orthonormale pour q et telle que la matrice de q' dans $\mathcal{B}_0$ est une matrice diagonale. Un tel résultat est appelé un résultat de codiagonalisation puisque la matrice de q dans $\mathcal{B}_0$ est la

(105) C'est-à-dire que si $\mathbf{x} \in \mathbb{R}\mathbf{e}_1$ alors $M\mathbf{x} \in \mathbb{R}\mathbf{e}_1$, voir la définition 8.20.

matrice identité (et par conséquent une matrice diagonale) et celle de q' dans cette même base est elle aussi une matrice diagonale.

Exemple La matrice $\mathrm{A} = \begin{pmatrix} 16 & -12 \\ -12 & 9 \end{pmatrix}$ est une matrice symétrique réelle, elle est par conséquent diagonalisable dans une base orthonormale.
Pour diagonaliser la matrice A, commençons par chercher ses valeurs propres. Le polynôme caractéristique P de A est donné par

$$P(\lambda) = \det(\mathrm{A} - \lambda \mathrm{Id}) = \det \begin{pmatrix} 16-\lambda & -12 \\ -12 & 9-\lambda \end{pmatrix} = \lambda(\lambda - 25).$$

Les valeurs propres de A sont $\lambda_1 = 0$ et $\lambda_2 = 25$. Pour déterminer la matrice de passage orthogonale, calculons les espaces propres associés à λ_1 et λ_2. L'espace propre associé à λ_1 est l'ensemble des vecteurs $\mathbf{x} = {}^t(x, y)$ tels que $\mathrm{A}\mathbf{x} = \lambda_1 \mathbf{x}$. Ce qui donne ici $16x - 12y = 0$. L'espace propre associé à la valeur propre λ_1 est donc engendré par $\mathbf{v}_1 = {}^t(3, 4)$. Afin d'obtenir une matrice de passage orthogonale, considérons un vecteur colinéaire à $\mathbf{v}_1$ et de norme 1 :

$$\mathbf{e}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} = \begin{pmatrix} \frac{3}{5} \\ \frac{4}{5} \end{pmatrix}.$$

Pour calculer le second vecteur propre, on peut soit procéder comme pour le premier vecteur propre, soit utiliser le fait que la matrice de passage (c'est-à-dire la matrice formée par les vecteurs propres associés aux deux valeurs propres) doit être orthogonale. En utilisant cette seconde méthode, on établit qu'un vecteur orthogonal à $\mathbf{v}_1$ et de norme 1 est donné par $\mathbf{e}_2 = {}^t(-\frac{4}{5}, \frac{3}{5})$. La matrice A s'écrit $\mathrm{A} = \mathrm{P}\mathrm{D}{}^t\mathrm{P}$ avec $\mathrm{P} = \begin{pmatrix} \frac{3}{5} & -\frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{pmatrix}$ et $\mathrm{D} = \begin{pmatrix} 0 & 0 \\ 0 & 25 \end{pmatrix}$.

Exercice 8.8 Considérons la matrice symétrique réelle

$$\mathrm{M} = \begin{pmatrix} 7 & 2 & -2 \\ 2 & 4 & -1 \\ -2 & -1 & 4 \end{pmatrix}.$$

1 - Déterminer les valeurs propres de M.
2 - Trouver les sous-espaces propres associés à chacune des valeurs propres et vérifier qu'ils sont orthogonaux deux à deux.
3 - Diagonaliser la matrice M avec une matrice de passage orthogonale.

8.5 Réduction des coniques et quadriques

8.5.1 Étude des coniques

Étant donné un triplet $(a, b, c) \in \mathbb{R}^3$, l'ensemble des points (x, y) du plan $\mathbb{R}^2$ qui sont les solutions de l'équation « linéaire »

$$ax + by + c = 0$$

correspond[106] généralement à une droite. On peut se demander de la même manière à quoi correspond géométriquement l'ensemble des points $(x, y) \in \mathbb{R}^2$ qui sont les solutions de l'équation « quadratique »

$$ax^2 + by^2 + 2cxy + \ell x + my + h = 0 \ ?$$

DÉFINITION 8.22 (Conique)
On appelle conique *l'ensemble des points* $(x, y) \in \mathbb{R}^2$ *tels que*

$$ax^2 + by^2 + 2cxy + \ell x + my + h = 0 \quad \textit{avec } (a, b, c, \ell, m, h) \in \mathbb{R}^6.$$

Étant donné une conique, notons P l'application définie sur $\mathbb{R}^2$ par

$$P(x, y) = ax^2 + by^2 + 2cxy + \ell x + my + h.$$

Elle peut se décomposer en la somme d'une forme quadratique q sur $\mathbb{R}^2$, d'une forme linéaire L sur $\mathbb{R}^2$ et d'une forme constante h ; on a $P = q + L + h$ avec

$$q(x, y) = ax^2 + by^2 + 2cxy \quad \text{et} \quad L(x, y) = \ell x + my.$$

Introduisons la matrice de cette forme quadratique dans la base canonique de $\mathbb{R}^2$

$$\mathrm{A} = \begin{pmatrix} a & c \\ c & b \end{pmatrix} \quad \text{et posons} \quad \mathrm{B} = \begin{pmatrix} \ell \\ m \end{pmatrix}, \quad \mathrm{X} = \begin{pmatrix} x \\ y \end{pmatrix}.$$

L'expression matricielle de l'équation de la conique devient :

$${}^t\mathrm{XAX} + {}^t\mathrm{BX} + h = 0.$$

Autres définitions équivalentes des coniques

✗ Une définition géométrique des coniques d'où est issue la dénomination « conique » est la suivante : les coniques sont des courbes planes résultant de l'intersection d'un plan avec un cône de révolution dans $\mathbb{R}^3$. Selon les positions relatives du plan et du cône, plus exactement selon l'angle d'inclinaison du plan avec l'axe du cône, on obtient les différents types de coniques, voir la figure 3 :

- si cet angle est supérieur à l'angle d'ouverture du cône, l'intersection est une ellipse ;
- si l'angle d'inclinaison est inférieur à l'angle d'ouverture, c'est une hyperbole ;
- si les deux angles sont égaux, c'est une parabole.

[106] Sauf dans quelques cas particuliers comme celui où $a = b = c = 0$ pour lequel tous les points $(x, y) \in \mathbb{R}^2$ satisfont la relation ou celui où $a = b = 0$ et $c \neq 0$ pour lequel aucun point $(x, y) \in \mathbb{R}^2$ ne satisfait la relation.

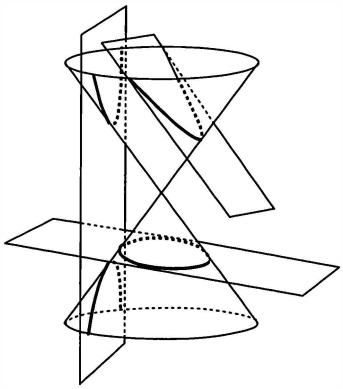


Fig. 3 Intersection d'un plan et d'un cône.

✗ Une définition par foyer et directrice. Dans le plan, on considère une droite D, un point F non situé sur D et un réel e strictement positif. On appelle conique de droite directrice D, de foyer F et d'excentricité e l'ensemble des points M du plan $\mathbb{R}^2$ vérifiant :

$$d(M,F) = e\, d(M,D),$$

$d(M,F)$ mesurant la distance du point M au point F et $d(M,D)$ celle entre le point M et la droite D.

- si $0 < e < 1$ alors l'ensemble des points M est une ellipse ;
- si $e = 1$ alors l'ensemble des points M est une parabole ;
- si $e > 1$ alors l'ensemble des points M est une hyperbole.

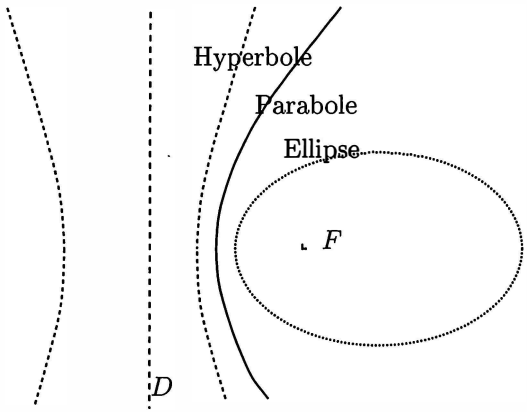


Fig. 4 Conique définie par foyer et directrice.

✗ Une définition bifocale. L'ellipse peut être définie comme le lieu des points dont la somme des distances à deux points fixes appelés foyers de l'ellipse est constante et égale à une valeur fixée, voir la figure 5.

$$d(M,F) + d(M,F') = \text{constante}.$$

Cette définition reste valable dans le cas du cercle, dans lequel les foyers sont confondus. L'hyperbole peut être définie comme le lieu des points dont la différence des distances à deux points fixes appelés foyers de l'hyperbole est constante et égale à une valeur fixée. La parabole n'a pas de définition bifocale.

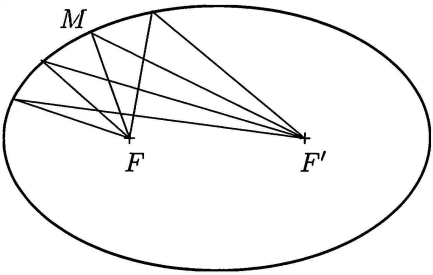


Fig. 5 Ellipse définie par deux foyers.

8.5.2 Classification des coniques

On peut classer les coniques selon la signature $\varepsilon(q) = (\alpha, \beta)$ de la forme quadratique qui lui est attachée[107].

DÉFINITION 8.23 *Soit C une conique d'équation*

$$ax^2 + by^2 + 2cxy + \ell x + my + h = 0.$$

On note q la forme quadratique $(x, y) \in \mathbb{R}^2 \mapsto ax^2 + by^2 + 2cxy$ et $\varepsilon(q)$ la signature de q.

✗ *Si $\varepsilon(q) = (2, 0)$ ou $\varepsilon(q) = (0, 2)$ alors la conique C est appelée* ellipse.

✗ *Si $\varepsilon(q) = (1, 1)$ alors la conique C est appelée* hyperbole.

✗ *Si $\varepsilon(q) = (1, 0)$ ou $\varepsilon(q) = (0, 1)$ alors la conique C est appelée* parabole.

Notons que quasiment tous les cas possibles de coniques sont décrits dans cette définition. En effet, le rang d'un forme quadratique sur $\mathbb{R}^2$ ne pouvant excéder deux, on a nécessairement $\varepsilon(q) = (\alpha, \beta)$ avec $\alpha + \beta \leqslant 2$. L'unique conique non décrite par la définition précédente est donc la conique pour laquelle la signature vaut $\varepsilon(q) = (0, 0)$. Pour une telle conique, on a $q = 0$ et l'ensemble C est une droite.

La signature d'une forme quadratique correspond au nombre de valeurs propres strictement positives et strictement négatives de la matrice de la forme quadratique dans une base. Connaître le déterminant Δ de cette matrice (c'est-à-dire

[107] La signature permet de classifier les formes quadratiques entre elles, voir p. 471.

le produit de ses deux valeurs propres) permet de connaître le genre de la conique. Pour une conique d'équation $ax^2 + by^2 + 2cxy + \ell x + my + h = 0$, ce déterminant vaut $\Delta = ab - c^2$. Si $\Delta > 0$ alors les deux valeurs propres doivent nécessairement être de même signe (et non nulles), la signature de la forme quadratique sera $(2,0)$ ou $(0,2)$ et la conique de genre ellipse. Si $\Delta = 0$ alors l'une des valeurs propres est nulle[108] : la signature est $(1,0)$ ou $(0,1)$ et on est dans le cas du genre parabole. Enfin, si $\Delta < 0$ alors les deux valeurs propres sont de signe opposé, l'une est strictement positive et l'autre strictement négative, la signature est $(1,1)$ et la conique est du genre hyperbole. Pour résumer, on a la classification suivante.

PROPOSITION 8.30 *Soit $\mathcal{C}$ une conique d'équation*

$$ax^2 + by^2 + 2cxy + \ell x + my + h = 0,$$

où $(a,b,c) \neq (0,0,0)$.

✗ *Si $ab - c^2 > 0$ alors $\mathcal{C}$ est une ellipse.*

✗ *Si $ab - c^2 = 0$ alors $\mathcal{C}$ est une parabole.*

✗ *Si $ab - c^2 < 0$ alors $\mathcal{C}$ est une hyperbole.*

8.5.3 Description des différents genres de coniques

Nous venons de classer les coniques en trois catégories en fonction de la signature de la forme quadratique décrivant cette conique. Nous allons dans cette partie étudier géométriquement ces trois types de conique.

Considérons une conique d'équation $ax^2 + by^2 + 2cxy + \ell x + my + h = 0$. Notons $(\mathbf{i},\mathbf{j})$ la base canonique du plan et $(O;\mathbf{i},\mathbf{j})$ un repère du plan. Dans ce repère, la conique correspond à l'ensemble des points $\mathbf{x} = x\mathbf{i} + y\mathbf{j}$ satisfaisant $ax^2 + by^2 + 2cxy + \ell x + my + h = 0$.

On note q la forme quadratique définie sur $\mathbb{R}^2$ par $q : (x,y) \mapsto ax^2 + by^2 + 2cxy$.

Description du genre parabole

Supposons que la signature de q soit $(1,0)$ ou $(0,1)$. La matrice de q dans la base $(\mathbf{i},\mathbf{j})$ qui est une matrice symétrique et à coefficients réels est, d'après le théorème 8.10, diagonalisable dans une base orthonormale $(\mathbf{i}',\mathbf{j}')$. Puisque la signature de q est $(1,0)$ ou $(0,1)$, une (et une seule) des valeurs propres de q est nulle. Dans le repère $(O,\mathbf{i}',\mathbf{j}')$ l'équation de la conique s'écrit[109]

$$x'^2 + \ell' x' + m' y' + h = 0.$$

[108] Excluons le cas où les deux valeurs propres sont nulles, correspondant au cas ou la matrice elle-même serait nulle.

[109] On verra dans les exemples qui suivent comment trouver effectivement ce repère et comment se transforment les termes linéaires $\ell x + my$ et le terme constant h.

1. Supposons que $m' \neq 0$ et effectuons les changements de variables $x'' = x' + \ell'/2$ et $y'' = y' + (4h - \ell'^2)/(4m')$. Géométriquement, ce changement de variables correspond à une translation du repère $(O; \mathbf{i}', \mathbf{j}')$ en le repère $(O'; \mathbf{i}', \mathbf{j}')$ où O' a pour coordonnées $(-\ell'/2, (\ell'^2 - 4h)/(4m'))$ dans le repère $(O; \mathbf{i}', \mathbf{j}')$. Dans ce nouveau repère, l'équation de la conique est :

$$y'' = -\frac{1}{m'} x''^2.$$

La conique correspond donc au graphe de la fonction $x'' \mapsto -x''^2/m'$, voir la figure 6 à gauche. C'est une parabole.

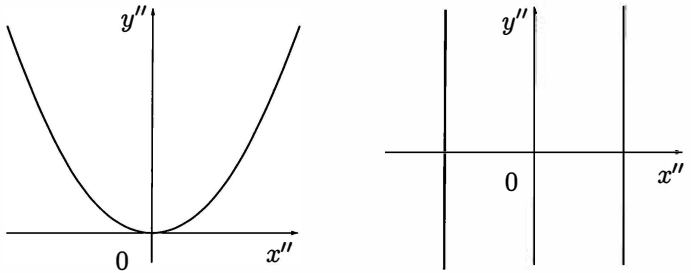


Fig. 6 Une parabole (à gauche) et deux droites parallèles (à droite).

2. Supposons que $m' = 0$ et effectuons le changement de variable $x'' = x' + \ell'/2$. On obtient, en posant $k = \ell'/4 - h$, pour équation de la conique

$$x''^2 = k.$$

Selon le signe du réel k, on obtient deux droites distinctes ($k > 0$, voir la figure 6 à droite), une seule droite ($k = 0$) ou l'ensemble vide ($k < 0$).

Description du genre ellipse

Supposons que la signature de q soit $(2, 0)$ ou $(0, 2)$. La matrice de q dans la base $(\mathbf{i}, \mathbf{j})$ est une matrice symétrique à coefficients réels. D'après le théorème 8.10, elle est diagonalisable dans une base orthonormale $(\mathbf{i}', \mathbf{j}')$. Puisque les deux valeurs propres sont de même signe (et non nulles), dans le repère $(O; \mathbf{i}', \mathbf{j}')$ l'équation de la conique s'écrit :

$$\frac{x'^2}{a^2} + \frac{y'^2}{b^2} + \ell' x' + m' y' + h = 0.$$

Effectuons le changement de variables $x'' = x' + a^2\ell'/2$ et $y'' = y' + b^2 m'/2$. Géométriquement, ce changement de variables correspond à une translation du repère $(O; \mathbf{i}', \mathbf{j}')$ en le repère $(O'; \mathbf{i}', \mathbf{j}')$ où l'origine O' du nouveau repère a pour coordonnées $(-a^2\ell'/2, -b^2 m'/2)$ dans $(O; \mathbf{i}', \mathbf{j}')$. L'équation s'écrit sous la forme (on a posé $k = a^2\ell'^2/4 + b^2 m'^2/4 - h$)

$$\frac{x''^2}{a^2} + \frac{y''^2}{b^2} = k.$$

Si $k > 0$ alors la conique correspond à une ellipse[110], voir la figure 7, si $k = 0$, la conique est réduite à un seul point et si $k < 0$ on obtient l'ensemble vide.

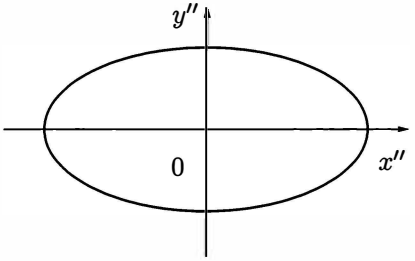


Fig. 7 Une ellipse.

Description du genre hyperbole

Supposons que la forme quadratique q ait pour signature $(1, 1)$. Comme pour les deux cas précédents, le théorème 8.10 permet d'affirmer que dans un repère $(O; \mathbf{i}', \mathbf{j}')$ l'équation de cette conique s'écrit

$$\frac{x'^2}{a^2} - \frac{y'^2}{b^2} + \ell' x' + m' y' + h = 0.$$

En utilisant le changement de variables $x'' = x' + a^2\ell'/2$ et $y'' = y' + b^2 m'/2$, on écrit l'équation de l'hyperbole sous la forme

$$\frac{x''^2}{a^2} - \frac{y''^2}{b^2} = k.$$

Si $k \neq 0$ alors on obtient une hyperbole dont les asymptotes sont les droites d'équation $x'' = y''$ et $x'' = -y''$; si $k = 0$, on obtient une hyperbole dégénérée en ses deux asymptotes.

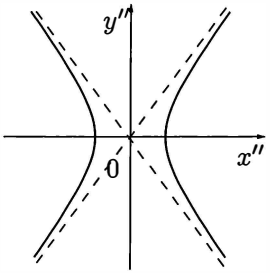


Fig. 8 Une hyperbole.

[110] On prendra garde au fait que l'équation $x''^2 + y''^2 = k$ avec $k > 0$ n'est pas nécessairement l'équation d'un cercle. En effet, le repère $(O'; \mathbf{i}', \mathbf{j}')$ n'étant pas nécessairement orthonormé, les points de composantes $(k, 0)$ et $(0, k)$ dans ce repère appartiennent tous les deux à la conique mais ne sont pas forcément à la même distance (considérée ici dans son sens usuel) de l'origine O'.

Exemple Étudions la conique $\mathcal{C}$ dont l'équation dans le plan $\mathbb{R}^2$, muni de son repère orthonormé canonique $(O;\mathbf{i},\mathbf{j})$, est

$$16x^2 + 9y^2 - 24xy + 35x - 20y = 0.$$

Avec les notations de la proposition 8.30 on a $ab - c^2 = 16 \times 9 - 12^2 = 0$, ce qui permet d'affirmer que la conique $\mathcal{C}$ est une parabole. Matriciellement, l'équation de la conique est ${}^t\mathrm{X}\mathrm{A}\mathrm{X} + {}^t\mathrm{B}\mathrm{X} + C = 0$, où

$$\mathrm{A} = \begin{pmatrix} 16 & -12 \\ -12 & 9 \end{pmatrix}, \qquad \mathrm{B} = \begin{pmatrix} 35 \\ -20 \end{pmatrix}, \qquad C = 0 \quad \text{et} \quad \mathrm{X} = \begin{pmatrix} x \\ y \end{pmatrix}.$$

Étape 1 - Au cours de l'exemple de la p. 506, nous avons diagonalisé la matrice A en écrivant A sous a forme $\mathrm{A} = \mathrm{P}\mathrm{D}{}^t\mathrm{P}$ avec

$$\mathrm{P} = \begin{pmatrix} 3/5 & -4/5 \\ 4/5 & 3/5 \end{pmatrix} \quad \text{et} \quad \mathrm{D} = \begin{pmatrix} 0 & 0 \\ 0 & 25 \end{pmatrix}.$$

Étape 2 - La matrice de changement de base P est une matrice de rotation puisqu'elle est de la forme $\mathrm{P} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$ avec $\theta = \arccos(3/5)$. Autrement dit, dans un repère obtenu par une rotation d'angle θ par rapport au repère initial, la forme quadratique $(x, y) \mapsto 16x^2 + 9y^2 - 24xy$ s'écrira $(x', y') \mapsto 25y'^2$.

Étape 3 - L'équation de la conique $\mathcal{C}$ dans le repère $(O;\mathbf{i},\mathbf{j})$ s'écrit matriciellement

$${}^t\mathrm{X}\mathrm{P}\mathrm{D}{}^t\mathrm{P}\mathrm{X} + {}^t\mathrm{B}\mathrm{X} + C = 0.$$

En posant $\mathrm{X}' = {}^t\mathrm{P}\mathrm{X}$ on a

$${}^t\mathrm{X}'\mathrm{D}\mathrm{X}' + {}^t\mathrm{B}\mathrm{P}\mathrm{X}' + C = 0$$

et en notant x' et y' les composantes de X', on obtient dans le repère[(111)] $(O;\mathbf{i}',\mathbf{j}')$ l'équation de la conique $\mathcal{C}$:

$$25y'^2 + 5x' - 40y' = 0.$$

Étape 4 - On effectue une translation de l'origine du repère $(O;\mathbf{i}',\mathbf{j}')$. Pour cela on regroupe les deux termes contenant y' en faisant apparaître un carré :

$$25y'^2 + 5x' - 40y' = 25\left(y' - \frac{4}{5}\right)^2 - 16 + 5x'.$$

En posant $x'' = x' - \frac{16}{5}$ et $y'' = y' - \frac{4}{5}$, on obtient l'équation réduite dans le nouveau repère $(O';\mathbf{i}',\mathbf{j}')$ déduit par translation du repère $(O;\mathbf{i}',\mathbf{j}')$:

$$5y''^2 + x'' = 0.$$

Étape 5 - Pour tracer la parabole dans le repère initial $(O;\mathbf{i},\mathbf{j})$, on place successivement les repères $(O;\mathbf{i}',\mathbf{j}')$ et $(O';\mathbf{i}',\mathbf{j}')$. On dessine dans $(O';\mathbf{i}',\mathbf{j}')$ la parabole d'équation réduite $5y''^2 + x'' = 0$, voir la figure 9.

(111) Ce repère est celui dans laquelle la matrice A est diagonale, voir l'exemple de la p. 506.

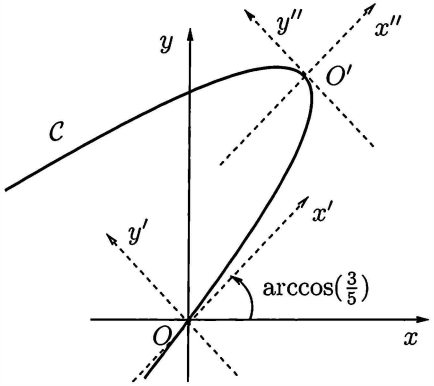


Fig. 9 La conique $\mathcal{C}$ dans les différents repères.

Exercice 8.9 Dans un repère orthonormé du plan $\mathbb{R}^2$, On considère la conique Γ d'équation $2xy + \sqrt{2}x - 1 = 0$.
1 - Préciser la nature de cette conique.
2 - Trouver un repère orthonormé dans lequel l'équation de Γ est réduite.
3 - Représenter la conique dans le repère initial.

8.5.4 Étude des quadriques

Rappelons que dans $\mathbb{R}^3$, l'ensemble des points (x, y, z) vérifiant une équation « linéaire » de la forme :

$$ax + by + cz + d = 0 \quad \text{où } (a, b, c, d) \in \mathbb{R}^4 \text{ avec } (a, b, c) \neq (0, 0, 0)$$

est un plan. On peut là encore se demander ce que représente géométriquement l'ensemble des points $(x, y, z) \in \mathbb{R}^3$ qui sont les solutions de l'équation « quadratique »

$$ax^2 + by^2 + cz^2 + 2dxy + 2eyz + 2fzx + \ell x + my + nz + h = 0.$$

DÉFINITION 8.24 *On appelle* quadrique, *l'ensemble des points* $(x, y, z) \in \mathbb{R}^3$ *tels que*

$$ax^2 + by^2 + cz^2 + 2dxy + 2eyz + 2fzx + \ell x + my + nz + h = 0$$

où $(a, b, c, d, e, f, \ell, m, n, h) \in \mathbb{R}^{10}$.

Comme pour les coniques, on peut classifier les quadriques selon la signature de la forme quadratique $q : (x, y, z) \mapsto ax^2 + by^2 + cz^2 + 2dxy + 2eyz + 2fzx$. Introduisons la matrice de cette forme quadratique

$$\mathrm{A} = \begin{pmatrix} a & d & f \\ d & b & e \\ f & e & c \end{pmatrix} \quad \text{et posons} \quad \mathrm{B} = \begin{pmatrix} \ell \\ m \\ n \end{pmatrix} \quad \text{et} \quad X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

L'écrire matricielle d'une quadrique est :

$$^t\mathrm{X}\mathrm{A}\mathrm{X} + {}^t\mathrm{B}\mathrm{X} + h = 0.$$

8.5.5 Classification des quadriques

La matrice A étant symétrique réelle, d'après le théorème 8.10, elle est diagonalisable et admet des vecteurs propres formant une base orthonormale de $\mathbb{R}^3$. Ces vecteurs propres sont appelés « directions propres » de la quadrique.

Le signature de la forme quadratique q étant déterminée par le signe[112] des valeurs propres de la matrice A, on peut classifier les quadriques selon le signe des valeurs propres λ_1, λ_2 et λ_3 de A.

✗Supposons dans un premier temps que les trois valeurs propres sont non nulles : dans ce cas, comme l'illustre l'exemple suivant, on peut[113] se ramener après transformation orthogonale puis translation à une équation du type

$$\lambda_1\, x''^2 + \lambda_2\, y''^2 + \lambda_3\, z''^2 \;=\; k. \tag{28}$$

Exemple Intéressons-nous à la quadrique dont l'équation dans un repère orthonormé $(O; \mathbf{i}, \mathbf{j}, \mathbf{k})$ est

$$4z^2 + 2xy - 2x = 0.$$

Le principe de cette étude est similaire à celui mené pour les coniques. En posant

$$\mathrm{X} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \quad \mathrm{A} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 4 \end{pmatrix} \quad \text{et} \quad \mathrm{B} = \begin{pmatrix} -2 \\ 0 \\ 0 \end{pmatrix},$$

on obtient pour équation matricielle de la quadrique : $^t\mathrm{X}\mathrm{A}\mathrm{X} + {}^t\mathrm{B}\mathrm{X} = 0$. Les valeurs propres de la matrice A sont $\lambda_1 = 1$, $\lambda_2 = -1$ et $\lambda_3 = 4$. La matrice A est donc diagonalisable et la matrice de passage entre la base canonique et la base des vecteurs propres est la matrice orthogonale

$$\mathrm{P} = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Cette matrice correspond à une matrice de rotation d'angle $\pi/4$ et d'axe $(O; \mathbf{k})$.

En effectuant le changement de variable $\mathrm{X} = \mathrm{P}\mathrm{X}'$, l'équation devient

$$x'^2 - y'^2 + 4z'^2 - \sqrt{2}x' + \sqrt{2}y' = 0.$$

---

[112] Voir la classification des formes quadratiques p. 471.

[113] En pratique, on utilise le théorème 8.10 en diagonalisant la matrice A avec une matrice de passage orthogonale, puis on utilise des changements de variable afin de ne faire apparaître qu'une seule fois chacune des variables.

En regroupant les termes contenant les mêmes variables, on obtient :

$$\left(x' - \frac{1}{\sqrt{2}}\right)^2 - \left(y' - \frac{1}{\sqrt{2}}\right)^2 + 4z'^2 = 0.$$

Ainsi, en posant $x'' = x' - \frac{1}{\sqrt{2}}$, $y'' = y' - \frac{1}{\sqrt{2}}$ et $z'' = z'$, c'est-à-dire en effectuant une translation, l'équation réduite de la quadrique est : $x''^2 - y''^2 + 4z''^2 = 0$.

Dans le cas où les trois valeurs propres de la matrice A sont non nulles, on a l'un des deux genres suivants.

Genre ellipsoïde

On dit qu'une quadrique est du genre ellipsoïde lorsque les trois valeurs propres λ_1, λ_2 et λ_3 sont strictement positives. Dans ce cas, on peut écrire l'équation (28) de la quadrique sous la forme

$$\frac{x''^2}{a^2} + \frac{y''^2}{b^2} + \frac{z''^2}{c^2} = k.$$

Lorsque $k > 0$ on obtient un ellipsoïde (voir la figure 10), lorsque $k = 0$ on obtient un point et lorsque $k < 0$ on obtient l'ensemble vide.

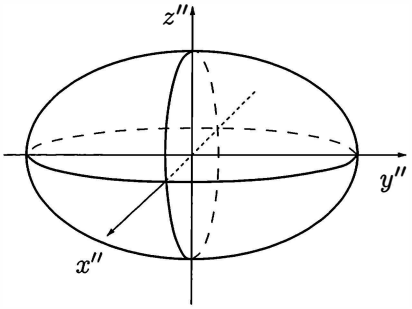


Fig. 10 Un ellipsoïde.

Genre hyperboloïde

Si $\lambda_1 > 0$, $\lambda_2 > 0$ et $\lambda_3 < 0$ alors on peut alors écrire l'équation (28) sous la forme

$$\frac{x''^2}{a^2} + \frac{y''^2}{b^2} - \frac{z''^2}{c^2} = k.$$

Lorsque $k > 0$ la quadrique est un hyperboloïde à une nappe (voir la figure 11 à gauche), lorsque $k = 0$ on obtient un cône et lorsque $k < 0$ la quadrique est un hyperboloïde à deux nappes (voir la figure 11 à droite).

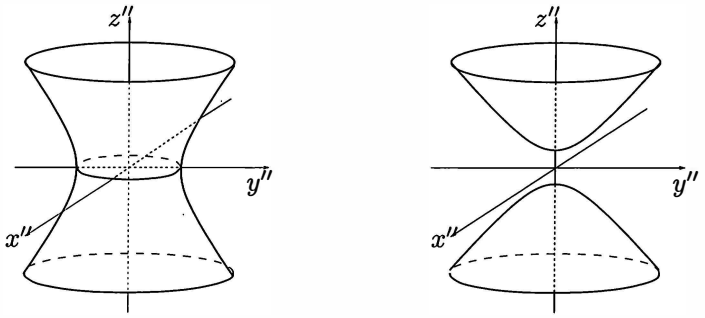


Fig. 11 Un hyperboloïde à une nappe (à gauche) et un hyperboloïde à deux nappes (à droite).

✗Supposons maintenant qu'une des valeurs propres λ_1, λ_2 ou λ_3 et une seule soit nulle ; pour fixer les choses, on supposera qu'il s'agit de λ_3. Dans ce cas, l'équation de la quadrique peut s'écrire sous la forme

$$\lambda_1 \, x''^2 + \lambda_2 \, y''^2 + \alpha \, z'' \; = \; k. \tag{29}$$

On distingue les deux genres suivants.

Genre paraboloïde

Si $\alpha \neq 0$ alors en posant $Z = z'' - k/\alpha$, c'est-à-dire en effectuant une translation du repère, on écrit l'équation (29) de la quadrique sous la forme

$$\frac{x''^2}{a^2} + \varepsilon \frac{y''^2}{b^2} + \alpha Z \; = \; 0 \quad \text{où } \varepsilon = +1 \text{ ou } \varepsilon = -1.$$

Si $\varepsilon = 1$ alors la quadrique est un paraboloïde elliptique (voir la figure 12 à gauche) et si $\varepsilon = -1$ alors la quadrique est un paraboloïde hyperbolique (voir la figure 12 à droite).

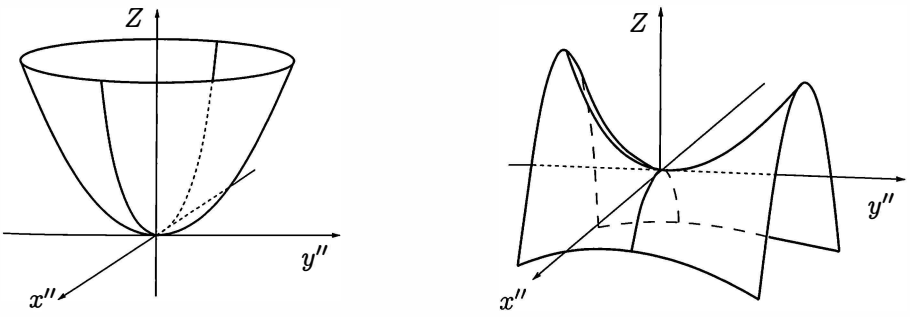


Fig. 12 Paraboloïdes elliptique (à gauche) et hyperbolique (à droite).

Genre cylindrique

Si $\alpha = 0$ alors l'équation (29) ne dépend plus de z'' et on obtient

$$\frac{x''^2}{a^2} + \varepsilon \frac{y''^2}{b^2} = k \quad \text{où } \varepsilon = +1 \text{ ou } \varepsilon = -1.$$

Lorsque $k = 0$, ou bien $\varepsilon = +1$ et l'ensemble cherché est la droite correspondant à l'axe des z'', ou bien $\varepsilon = -1$ et on obtient deux plans sécants d'équation $bx'' = \pm ay''$. Lorsque $k > 0$, la quadrique est un cylindre elliptique ($\varepsilon = 1$, voir la figure 13 à gauche) ou un cylindre hyperbolique ($\varepsilon = -1$, voir la figure 13 à droite). Enfin, lorsque $k < 0$, soit $\varepsilon = +1$ et on obtient l'ensemble vide, soit $\varepsilon = -1$ et on obtient à nouveau un cylindre hyperbolique (en « échangeant » les rôles de x'' et de y'').

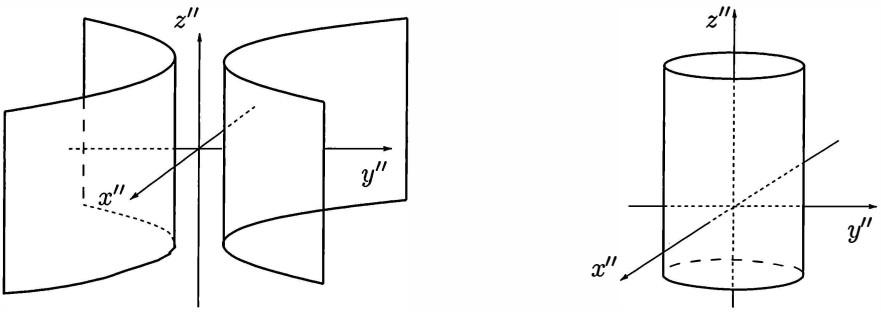


Fig. 13 Cylindres elliptique (à gauche) et hyperbolique (à droite).

✗ Supposons enfin que deux valeurs propres parmi λ_1, λ_2 et λ_3 soient nulles (par exemple $\lambda_2 = \lambda_3 = 0$). On se trouve dans l'une des deux situations suivantes. Soit on peut enlever les termes constants et l'équation de la quadrique s'écrit :

$$x''^2 + ky'' = 0.$$

On a un cylindre parabolique (voir la figure 14 à gauche), soit on peut ôter les termes linéaire et l'équation de la quadrique peut s'écrire sous la forme

$$x''^2 = k$$

auquel cas selon le signe du réel k on a deux plans parallèles ($k > 0$, voir la figure 14 à droite) qui dégénère en un seul plan lorsque $k = 0$, ou bien l'ensemble vide si $k < 0$.

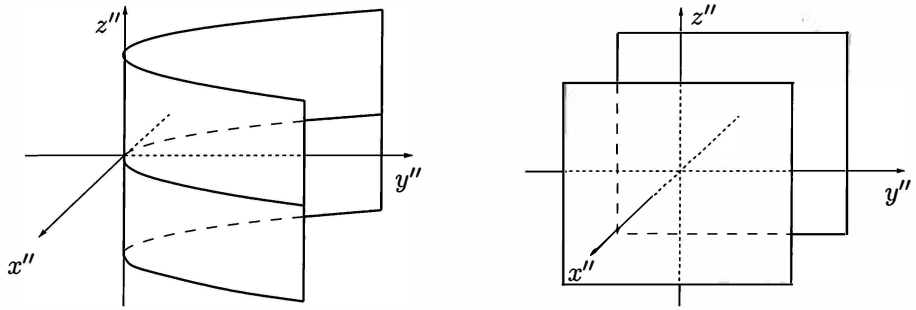


Fig. 14 Un cylindre parabolique (à gauche) et deux plans parallèles (à droite).

Remarque Les noms donnés à ces surfaces sont issus de la terminologie des coniques. Par exemple, un cylindre elliptique est une surface dont l'intersection avec certains plans (par exemple le plan d'équation $z'' = 0$ dans la présentation ci-dessus) est une ellipse. De la même façon, le paraboloïde hyperbolique admet pour section avec les axes coordonnées des paraboles ou des hyperboles... ◇

Exemple On a vu lors de l'exemple de la p. 515 que la quadrique d'équation $4z^2 + 2xy - 2x = 0$ dans le repère $(O; \mathbf{i}, \mathbf{j}, \mathbf{k})$ avait dans un autre repère $(O; \mathbf{i}'', \mathbf{j}'', \mathbf{k}'')$ orthonormé du plan pour équation $x''^2 - y''^2 + 4z''^2 = 0$. D'après la classification précédente, cette quadrique correspond à un cône. Pour obtenir la forme de ce cône dans le repère $(O; \mathbf{i}'', \mathbf{j}'', \mathbf{k}'')$, on peut identifier quelles sont les courbes obtenues par intersection du cône avec des plans.

- L'intersection du cône avec le plan $x'' = 0$, a pour équation $y'' = \pm 2z''$ ce qui correspond à deux droites.
- L'intersection du cône avec le plan $y'' = 1$, a pour équation $x''^2 + 4z''^2 = 1$ ce qui est l'équation d'une ellipse.

Ces informations permettent d'obtenir la représentation graphique de la quadrique de la figure 15.

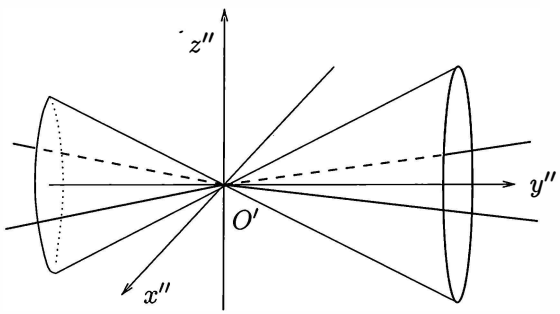


Fig. 15 La quadrique d'équation $x''^2 - y''^2 + 4z''^2 = 0$.

Si on veut connaître la position du cône dans le repère initial, il faut faire intervenir les transformations effectuées au cours de la réduction ayant permis de passer du repère initial au repère final. Il s'agit d'une part un rotation d'angle $\pi/4$ et d'axe $(O;\mathbf{k})$ puis une translation de vecteur $(1/\sqrt{2}, 1/\sqrt{2}, 0)$. Ils sont représentées sur la figure 16.

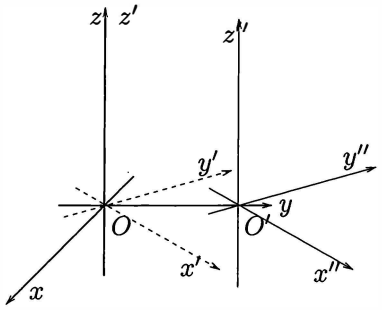


Fig. 16 Position des différents repères.

Exercice 8.10 Déterminer la nature la quadrique d'équation

$$x^2 + 4y^2 + 5z^2 - 4xy - 2x + 4y = 0.$$

8.6 Exercices de synthèse

Exercice 8.11

Dans cet exercice, q désigne une forme quadratique sur $\mathbb{R}^n$. Les termes « orthonormé » et « norme » utilisés dans l'énoncé font référence au produit scalaire canonique sur $\mathbb{R}^n$.

1 - Montrer que s'il existe une base orthonormée $\{\mathbf{f}_i\}_{i\in\{1,\dots,n\}}$ de $\mathbb{R}^n$ telle que

$$\forall i \in \{1, 2, \dots, n\} \quad q(\mathbf{f}_i) = 0$$

alors la trace[(114)] de q est nulle.

2 - Supposons maintenant que la trace de q est nulle.

a) Rappeler pourquoi il existe une base orthonormée $\{\mathbf{e}_i\}_{i\in\{1,\dots,n\}}$ de $\mathbb{R}^n$ dans laquelle la forme quadratique s'écrit

$$q(x_1, x_2, \dots, x_n) = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \cdots + \lambda_n x_n^2$$

où les λ_i sont des réels vérifiant $\lambda_1 + \lambda_2 + \cdots + \lambda_n = 0$.

[(114)] On rappelle que la trace d'une forme quadratique est la trace de la matrice représentant q dans une base orthonormée quelconque.

b) Trouver un vecteur $\mathbf{f}_1$ de $\mathbb{R}^n$ de norme égale à 1 et tel que $q(\mathbf{f}_1) = 0$.

c) En déduire par récurrence sur la dimension n qu'il existe une base orthonormée $\{\mathbf{f}_i\}_{i\in\{1,\dots,n\}}$ de $\mathbb{R}^n$ telle que

$$\forall i \in \{1, 2, \dots, n\} \quad q(\mathbf{f}_i) = 0.$$

Exercice 8.12 ***Projection sur un convexe fermé***

Soit $(E, \langle\cdot,\cdot\rangle)$ un espace préhilbertien *complet*. On note $\|\cdot\|$ la norme associée au produit scalaire $\langle\cdot,\cdot\rangle$. Soit K un sous-ensemble convexe[115] *fermé* non vide de E. L'objectif de cet exercice est de montrer le résultat suivant :

$$\forall \mathbf{f} \in E \quad \exists! \mathbf{u} \in K \quad \|\mathbf{f} - \mathbf{u}\| = \inf_{\mathbf{w}\in K} \|\mathbf{f} - \mathbf{w}\|. \tag{30}$$

1 - Dans le cas où $E = \mathbb{R}^2$ est muni du produit scalaire canonique, illustrer à l'aide d'un dessin la propriété (30). On pourra prendre, par exemple, pour K une boule fermée du plan et pour $\mathbf{f}$ un point extérieur à cette boule.

2 - Montrer que

$$\forall (\mathbf{x}, \mathbf{y}) \in E^2 \quad \|\mathbf{x}+\mathbf{y}\|^2 + \|\mathbf{x}-\mathbf{y}\|^2 = 2\left(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2\right). \tag{31}$$

3 - La propriété (31) a pour nom *identité du parallélogramme*. En considérant le cas où $E = \mathbb{R}^2$ est muni du produit scalaire canonique, justifier cette terminologie.

4 - On désigne par $\mathbf{f}$ un élément quelconque de E et on considère une suite $(\mathbf{w}_n)_n$ telle que[116]

$$\forall n \in \mathbb{N} \quad \mathbf{w}_n \in K \qquad \text{et} \qquad \lim_{n\to+\infty} \|\mathbf{f} - \mathbf{w}_n\| = \inf_{\mathbf{w}\in K} \|\mathbf{f} - \mathbf{w}\|.$$

a) En utilisant l'identité du parallélogramme, montrer que pour tout $(p, q) \in \mathbb{N}^2$ on a

$$2\left(\left\|\mathbf{f} - \tfrac{1}{2}(\mathbf{w}_p + \mathbf{w}_q)\right\|^2 + \left\|\tfrac{1}{2}(\mathbf{w}_p - \mathbf{w}_q)\right\|^2\right) = \|\mathbf{f} - \mathbf{w}_p\|^2 + \|\mathbf{f} - \mathbf{w}_q\|^2.$$

b) En déduire que la suite $(\mathbf{w}_n)_n$ est une suite de Cauchy.

c) En déduire qu'il existe $\mathbf{u} \in K$ tel que

$$\|\mathbf{f} - \mathbf{u}\| = \inf_{\mathbf{w}\in K} \|\mathbf{f} - \mathbf{w}\|. \tag{32}$$

---

[115] Rappelons qu'un sous-ensemble K d'un espace vectoriel est dit convexe si pour tout $(\mathbf{x}, \mathbf{y}) \in K^2$ on a l'inclusion $\{(1-t)\mathbf{x} + t\mathbf{y}, t \in [0,1]\} \subset K$. En particulier, si K est convexe et si $\mathbf{x}$ et $\mathbf{y}$ sont dans K alors $\frac{\mathbf{x}+\mathbf{y}}{2}$ est aussi dans K.

[116] Par définition de la borne inférieure, une telle suite existe. On ne demande pas de le démontrer.

5 - Soient $\mathbf{f}$ un élément de E et $\mathbf{u}$ un élément de K vérifiant la relation (32). Pour $\mathbf{w} \in K$ et $t \in [0,1]$, on note $\mathbf{v} = (1-t)\mathbf{u} + t\mathbf{w}$.

a) Montrer que $\|\mathbf{f}-\mathbf{v}\|^2 = \|\mathbf{f}-\mathbf{u}\|^2 - 2t\langle \mathbf{f}-\mathbf{u}, \mathbf{w}-\mathbf{u}\rangle + t^2\|\mathbf{w}-\mathbf{u}\|^2$.

b) En déduire que $\|\mathbf{f}-\mathbf{u}\|^2 \leqslant \|\mathbf{f}-\mathbf{u}\|^2 - 2t\langle \mathbf{f}-\mathbf{u}, \mathbf{w}-\mathbf{u}\rangle + t^2\|\mathbf{w}-\mathbf{u}\|^2$.

c) En déduire que $\mathbf{u}$ vérifie

$$\forall \mathbf{w} \in K \qquad \langle \mathbf{f}-\mathbf{u}, \mathbf{w}-\mathbf{u}\rangle \leqslant 0. \tag{33}$$

6 - Soient $\mathbf{f}$ un élément de E et $\mathbf{u}$ un élément de K vérifiant l'assertion (33).

a) Montrer que pour tout $\mathbf{w} \in K$, on a $\|\mathbf{f}-\mathbf{w}\|^2 \geqslant \|\mathbf{f}-\mathbf{u}\|^2$.

b) En déduire que $\mathbf{u}$ vérifie (32).

7 - a) Montrer qu'il existe un seul $\mathbf{u} \in K$ vérifiant (33).

b) En déduire le résultat annoncé par l'assertion (30).

8.7 Solutions des exercices

Solution de l'exercice 8.1

L'espace $\mathcal{M}_2(\mathbb{R})$ est un $\mathbb{R}$-espace vectoriel de dimension 4. La base canonique est constituée des quatre matrices suivantes :

$$\mathrm{E}^{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \mathrm{E}^{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \mathrm{E}^{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \quad \text{et} \quad \mathrm{E}^{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Notons E_1, E_2, E_3 et E_4 ces quatre matrices (dans l'ordre où elles sont décrites ci-dessus). En calculant les valeurs de chaque $\varphi(\mathrm{E}_i, \mathrm{E}_j)$, on obtient les coefficients de la matrice recherchée :

$$\mathrm{Mat}_{\mathcal{B}_{\mathrm{can}}}(\varphi) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Solution de l'exercice 8.2

On utilise la méthode de Gauss en écrivant successivement

$$\begin{aligned} q_a(x_1, x_2, x_3) &= x_1^2 + x_1(2x_2) + \big((1+a)x_2^2 + (1+a+a^2)x_3^2 - 2ax_2x_3\big) \\ &= (x_1 + x_2)^2 + \underbrace{\big(ax_2^2 + (1+a+a^2)x_3^2 - 2ax_2x_3\big)}_{= \widetilde{q_a}(x_2, x_3)} \end{aligned}$$

puis pour la forme quadratique $\widetilde{q_a}$:

$$\widetilde{q_a}(x_2, x_3) = ax_2^2 + x_2(-2ax_3) + (1+a+a^2)x_3^2 = a(x_2 - x_3)^2 + (1+a^2)x_3^2.$$

On obtient finalement la décomposition suivante :

$$q_a(x_1, x_2, x_3) = (x_1 + x_2)^2 + a(x_2 - x_3)^2 + (1+a^2)x_3^2.$$

Solution de l'exercice 8.3

Puisque la forme quadratique q_a se décompose sous la forme

$$q_a(\mathbf{x}) = (x_1 + x_2)^2 + a(x_2 - x_3)^2 + (1+a^2)x_3^2,$$

le théorème 8.4 nous donne directement la signature et le rang de cette forme quadratique :

- Si $a < 0$ alors $\varepsilon(q_a) = (2, 1)$ et $\mathrm{rg}(q_a) = 3$,
- Si $a = 0$ alors $\varepsilon(q_a) = (2, 0)$ et $\mathrm{rg}(q_a) = 2$,
- Si $a > 0$ alors $\varepsilon(q_a) = (3, 0)$ et $\mathrm{rg}(q_a) = 3$.

Solution de l'exercice 8.4

1 - L'application φ est symétrique car pour tout couple $(A, B) \in \mathcal{M}_n(\mathbb{R})^2$ on a $\mathrm{Tr}(AB) = \mathrm{Tr}(AB)$. Pour montrer qu'il s'agit d'une application bilinéaire il suffit donc de montrer qu'elle est linéaire par rapport à la première variable. Pour tout $(A, B, C) \in \mathcal{M}_n(\mathbb{R})^3$ et pour tout $(\lambda, \mu) \in \mathbb{R}^2$, il résulte des propriétés de l'application trace que[117] :

$$\begin{aligned}\varphi(\lambda A + \mu B, C) &= \mathrm{Tr}(\lambda AC + \mu AC) = \lambda \,\mathrm{Tr}(AC) + \mu \,\mathrm{Tr}(AC) \\ &= \lambda \varphi(A, C) + \mu \varphi(B, C).\end{aligned}$$

L'application φ est donc bien une forme bilinéaire symétrique sur $\mathcal{M}_n(\mathbb{R})$ et par conséquent sur $\mathcal{S}_n(\mathbb{R})$. Par ailleurs, si $A = (a_{i,j})$ est une matrice symétrique, on a

$$\varphi(A, A) = \mathrm{Tr}(A^2) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{i,j}a_{j,i} = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{i,j}^2.$$

Ainsi, l'application φ est positive et est définie car une somme de nombres réels élevés au carré est nulle si et seulement si tous les termes de la somme sont nuls. On en conclut que φ est un produit scalaire sur $\mathcal{S}_n(\mathbb{R})$.

2 - L'application φ n'est pas un produit scalaire sur $\mathcal{M}_n(\mathbb{R})$ car la forme quadratique de φ n'est pas une forme quadratique définie : la matrice A_0 contenant que zéro excepté en position $(1, n)$ vérifie $A_0^2 = 0$ donc $\varphi(A_0, A_0) = 0$ alors qu'elle n'est pas nulle.

Solution de l'exercice 8.5

Si on utilise le produit scalaire canonique sur $\mathbb{R}^n$, l'inégalité de Cauchy-Schwarz s'écrit

$$\forall\big((a_1, \ldots, a_n), (b_1, \ldots, b_n)\big) \in (\mathbb{R}^n)^2 \quad \left(\sum_{i=1}^{n} a_i b_i\right)^2 \leqslant \left(\sum_{i=1}^{n} a_i^2\right)\left(\sum_{i=1}^{n} b_i^2\right).$$

Pour $(x_1, \ldots, x_n) \in (\mathbb{R}^{*+})^n$, on choisit $a_i = \sqrt{x_i}$ et $b_i = 1/\sqrt{a_i}$ de sorte que l'inégalité ci-dessus devient

$$\left(\sum_{i=1}^{n} 1\right)^2 \leqslant \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} \frac{1}{x_i}\right)$$

ce qui correspond au résultat demandé.

D'après la remarque de la p. 481, il y a égalité si et seulement si $(a_1, \ldots, a_n)$ et $(b_1, \ldots, b_n)$ sont colinéaires, c'est-à-dire lorsqu'il existe $\lambda \in \mathbb{R}$ tel que

$$\forall i \in \{1, \ldots, n\} \quad b_i = \lambda a_i.$$

---

[117] Voir l'exercice 5 p. 448 du *Cours de première année*.

Autrement dit, il y a égalité si et seulement si il existe $\lambda \in \mathbb{R}$ tel que

$$\forall i \in \{1, \dots, n\} \quad x_i = \lambda.$$

donc lorsque tous les x_i sont égaux.

Solution de l'exercice 8.6

1 - Par définition de la norme $\|\cdot\|$, on a

$$\|\sin\| = \sqrt{\int_0^\pi \sin(t)^2 \, \mathrm{d}t} = \sqrt{\frac{\pi}{2}} \quad \text{et} \quad \|\cos\| = \sqrt{\int_0^\pi \cos(t)^2 \, \mathrm{d}t} = \sqrt{\frac{\pi}{2}}.$$

De même, on a $\|\mathrm{id}\| = \sqrt{\int_0^\pi t^2 \, \mathrm{d}t} = \sqrt{\frac{\pi^3}{3}}$.

2 - Supposons qu'il existe trois réels a, b et c tels que $a\sin + b\cos + c\,\mathrm{id} = 0$. On a donc, pour tout $t \in [0; \pi]$,

$$a\sin(t) + b\cos(t) + c\,t = 0.$$

En choisissant $t = 0$, ceci implique que $b = 0$. En choisissant ensuite $t = \pi$ on obtient $c\pi = 0$, c'est-à-dire $c = 0$ et en choisissant finalement $t = \pi/2$ on a $a = 0$. Ceci montre que la famille $\{\sin, \cos, \mathrm{id}\}$ est une famille libre.

3 - Pour déterminer la matrice de l'application bilinéaire $\langle\cdot,\cdot\rangle$ dans cette base $(\sin, \cos, \mathrm{id})$ de F, on évalue les réels $\langle f, g\rangle$ où f et g sont des éléments de la famille $\{\sin, \cos, \mathrm{id}\}$. Remarquons que l'on a déjà calculé les termes $\langle f, f\rangle$ pour les trois valeurs $f \in \{\sin, \cos, \mathrm{id}\}$ à la question 1. On a

$$\langle \sin, \cos\rangle = \int_0^\pi \sin(t)\cos(t) \, \mathrm{d}t = \int_0^\pi \frac{1}{2}\sin(2t) \, \mathrm{d}t = 0,$$

et en effectuant une intégration par partie dans les deux intégrales suivantes, on a

$$\langle \sin, \mathrm{id}\rangle = \int_0^\pi t\sin(t) \, \mathrm{d}t = \pi \quad \text{et} \quad \langle \cos, \mathrm{id}\rangle = \int_0^\pi t\cos(t) \, \mathrm{d}t = -2.$$

La matrice A de l'application bilinéaire $\langle\cdot,\cdot\rangle$ dans la base $(\sin, \cos, \mathrm{id})$ vaut donc

$$\mathrm{A} = \begin{pmatrix} \pi/2 & 0 & \pi \\ 0 & \pi/2 & -2 \\ \pi & -2 & \pi^3/3 \end{pmatrix}.$$

4 - Comme à la question 2, on montre aisément que les deux applications sin et cos ne sont pas liées, elles engendrent donc un sous-espace vectoriel de dimension 2.

5 - On utilise le procédé d'orthogonalisation de Schmidt à partir de la base (sin, cos, id) de F dont les deux premiers vecteurs forment une base de G : on pose

$$\varepsilon_1 = \sin, \quad \varepsilon_2 = \cos - \frac{\langle \varepsilon_1, \cos \rangle}{\langle \varepsilon_1, \varepsilon_1 \rangle} \varepsilon_1 \quad \text{et} \quad \varepsilon_3 = \mathrm{id} - \frac{\langle \varepsilon_1, \mathrm{id} \rangle}{\langle \varepsilon_1, \varepsilon_1 \rangle} \varepsilon_1 - \frac{\langle \varepsilon_2, \mathrm{id} \rangle}{\langle \varepsilon_2, \varepsilon_2 \rangle} \varepsilon_2.$$

En utilisant les calculs précédents (en particuliers les calculs effectués pour obtenir la matrice A), on obtient

$$\varepsilon_1 = \sin, \quad \varepsilon_2 = \cos \quad \text{et} \quad \varepsilon_3 = \mathrm{id} - 2\sin + \frac{4}{\pi} \cos.$$

Selon l'orthogonalisation de Schmidt, la famille $\{\varepsilon_1, \varepsilon_2\}$ est orthogonale et engendre G, et la famille $\{\varepsilon_1, \varepsilon_2, \varepsilon_3\}$ est orthogonale et engendre F. Pour obtenir des familles orthonormales, il suffit de diviser chacun de ces vecteurs de la base $\{\varepsilon_1, \varepsilon_2\}$ par leur norme : on pose

$$\mathbf{f}_1 = \frac{\varepsilon_1}{\|\varepsilon_1\|}, \quad \mathbf{f}_2 = \frac{\varepsilon_2}{\|\varepsilon_2\|}, \quad \mathbf{f}_3 = \frac{\varepsilon_3}{\|\varepsilon_3\|}.$$

On obtient ainsi la base orthonormale suivante : $\mathbf{f}_1 = \sqrt{2/\pi}\,\sin$, $\mathbf{f}_2 = \sqrt{2/\pi}\,\cos$ et

$$\mathbf{f}_3 = \frac{1}{\sqrt{\pi^3/3 - 2\pi - 8/\pi}} \left(\mathrm{id} - 2\sin + \frac{4}{\pi} \cos \right).$$

Solution de l'exercice 8.7

Montrons que la matrice A est une matrice orthogonale ; en vertu de la proposition 8.25, elle sera donc inversible et on aura $\mathrm{A}^{-1} = {}^t\mathrm{A}$. Pour prouver que A est une matrice orthogonale, on vérifie par exemple que les trois colonnes forment une base orthonormale de $\mathbb{R}^3$ pour le produit scalaire canonique sur $\mathbb{R}^3$. Ces trois colonnes sont

$$\mathrm{C}_1 = \begin{pmatrix} 2/3 \\ -1/3 \\ -2/3 \end{pmatrix}, \quad \mathrm{C}_2 = \begin{pmatrix} 2/3 \\ 2/3 \\ 1/3 \end{pmatrix} \quad \text{et} \quad \mathrm{C}_2 = \begin{pmatrix} 1/3 \\ -2/3 \\ -2/3 \end{pmatrix}$$

et il est aisé de vérifier que

$$\langle \mathrm{C}_1, \mathrm{C}_2 \rangle = 0, \quad \langle \mathrm{C}_2, \mathrm{C}_3 \rangle = 0, \quad \langle \mathrm{C}_3, \mathrm{C}_1 \rangle = 0,$$
$$\|\mathrm{C}_1\| = 1, \quad \|\mathrm{C}_2\| = 1 \quad \text{et} \quad \|\mathrm{C}_3\| = 1.$$

Solution de l'exercice 8.8

1 - Les valeurs propres de la matrice M correspondent aux racines du polynôme caractéristique $\det(\mathrm{M} - X\mathrm{I}_3)$ qui est donné par

$$\det(\mathrm{M} - X\mathrm{I}_3) = \begin{vmatrix} 7 - X & 2 & -2 \\ 2 & 4 - X & -1 \\ -2 & -1 & 4 - X \end{vmatrix}.$$

On calcul ce déterminant en utilisant les méthodes vues dans le *Cours de première année*, chap. 11. En remplaçant la dernière colonne de ce déterminant par la somme des deux dernières, on a

$$\det(\mathrm{M}-X\mathrm{I}_3)=\begin{vmatrix}7-X & 2 & 0\\ 2 & 4-X & 3-X\\ -2 & -1 & 3-X\end{vmatrix}=(3-X)\begin{vmatrix}7-X & 2 & 0\\ 2 & 4-X & 1\\ -2 & -1 & 1\end{vmatrix}.$$

Ensuite, en remplaçant la dernière ligne par la différence des deux dernières, on a

$$\det(\mathrm{M}-X\mathrm{I}_3)=(3-X)\begin{vmatrix}7-X & 2 & 0\\ 2 & 4-X & 1\\ 4 & 5-X & 0\end{vmatrix}.$$

Enfin, on développe ce déterminant suivant la dernière colonne. On obtient

$$\begin{aligned}\det(\mathrm{M}-X\mathrm{I}_3)&=-(3-X)\begin{vmatrix}7-X & 2\\ 4 & 5-X\end{vmatrix}=-(3-X)(27-12X+X^2)\\&=-(3-X)^2(9-X).\end{aligned}$$

Les racines de ce polynôme sont 9 et 3 (cette dernière étant double). Par conséquent, les valeurs propres de M sont 9 et 3.

2 - Soit $\mathbf{x}={}^t(x,y,z)$ un vecteur propre de la matrice M associé à la valeur propre 3. On a

$$\mathrm{M}\mathbf{x}=3\mathbf{x}\iff\begin{cases}7x+2y-2z=3x\\ 2x+4y-z=3y\\ -2x-y+4z=3z\end{cases}\iff\begin{cases}4x+2y-2z=0\\ 2x+y-z=0.\\ -2x-y+z=0\end{cases}$$

Ces trois dernières égalités étant équivalentes, on en déduit que $\mathrm{M}\mathbf{x}=3\mathbf{x}$ si et seulement si $2x+y-z=0$.

De même, considérons $\mathbf{x}={}^t(x,y,z)$ un vecteur propre de la matrice M associé à la valeur propre 9. On a

$$\mathrm{M}\mathbf{x}=9\mathbf{x}\iff\begin{cases}-2x+2y-2z=0\\ 2x-5y-z=0.\\ -2x-y-5z=0\end{cases}$$

En ajoutant les deux dernières égalités, on obtient $y=-z$. En utilisant cette relation, on déduit que

$$\mathrm{M}\mathbf{x}=9\mathbf{x}\iff(\ x=2y\quad\text{et}\quad y+z=0\ ).$$

Finalement, le sous-espace propre associé à la valeur propre 9 est la droite D d'équations $x=2y$ et $y+z=0$. Le sous-espace propre associé à la valeur propre 3 est le plan P d'équation $2x+y-z=0$. Afin de prouver que ces deux sous-espaces sont orthogonaux, considérons $\mathbf{a}=(x_D,y_D,z_D)\in D$ et $\mathbf{b}=(x_P,y_P,z_P)\in P$. Le produit scalaire canonique de ces deux vecteurs de $\mathbb{R}^3$ vaut

$$\mathbf{a}\cdot\mathbf{b}=x_D\ x_P+y_D\ y_P+z_D\ z_P.\tag{34}$$

Puisque $(x_D, y_D, z_D) \in D$ on a $x_D = 2y_D$ et $z_D = -y_D$ la relation (34) s'écrit

$$\mathbf{a} \cdot \mathbf{b} = 2y_D\, x_P + y_D\, y_P - y_D\, z_P = y_D(2x_P + y_P - z_P).$$

Puisque $(x_P, y_P, z_P) \in P$ on a aussi $2x_P + y_P - z_P = 0$. Le produit scalaire $\mathbf{a} \cdot \mathbf{b}$ est donc nul, ce qui montre que la droite D et le plan P sont orthogonaux.

3 - Puisque la droite D est un espace vectoriel de dimension 1, tout élément non nul de D est une base orthogonale de D. Ainsi, le vecteur $\mathbf{e}_1 = {}^t(2, 1, -1)$, qui est dans D, est une base orthogonale de D. Sa norme valant $\|\mathbf{e}_1\| = \sqrt{6}$, la famille $\{{}^t(\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{-1}{\sqrt{6}})\}$ est une base orthogonale de D.

Afin de déterminer une base orthonormale du plan P nous allons utiliser la méthode d'orthogonalisation de Schmidt. Une base du plan est donnée par $\mathbf{f}_2 = {}^t(0, 1, 1)$ et $\mathbf{f}_3 = {}^t(1, 0, 2)$. La méthode de Schmidt[118] permet, à partir de cette base, d'obtenir une base $\{\mathbf{e}_2, \mathbf{e}_3\}$ qui est orthogonale :

$$\mathbf{e}_2 = \mathbf{f}_2 \quad \text{et} \quad \mathbf{e}_3 = \mathbf{f}_3 - \frac{\mathbf{f}_2 \cdot \mathbf{f}_3}{\|\mathbf{f}_2\|^2}\mathbf{f}_2 = \mathbf{f}_3 - \mathbf{f}_2.$$

On a $\mathbf{e}_2 = {}^t(0, 1, 1)$ et $\mathbf{e}_3 = {}^t(-1, 1, -1)$. Pour obtenir une base orthonormale du plan P, on divise chacun de ces deux vecteurs par leur norme respective. La famille $\{{}^t(0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}), {}^t(\frac{-1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{-1}{\sqrt{3}})\}$ est une base orthogonale de D.

Cette étude permet de conclure que la matrice M est semblable à une matrice diagonale M′ et que la matrice de passage P est orthogonale :

$$\mathrm{M}' = \begin{pmatrix} 9 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix} \quad \text{et} \quad \mathrm{P} = \begin{pmatrix} \frac{2}{\sqrt{6}} & 0 & \frac{-1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{-1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{3}} \end{pmatrix}.$$

Solution de l'exercice 8.9

1 - En utilisant la proposition 8.30 p. 510, on en déduit immédiatement que la quadrique est une hyperbole.

2 - La matrice de la forme quadratique $(x, y) \in \mathbb{R}^2 \mapsto 2xy$ est

$$\mathrm{A} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Cette matrice possède pour valeurs propres 1 et -1, dont des vecteurs propres sont ${}^t(1, 1)$ et ${}^t(1, -1)$. En divisant ces deux vecteurs par leur propre norme, on en déduit que la matrice A s'écrit $\mathrm{A} = \mathrm{PDP}^{-1}$ où

$$\mathrm{D} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{et} \quad \mathrm{P} = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix}.$$

---

[118] Voir la section 8.2.3, p. 486.

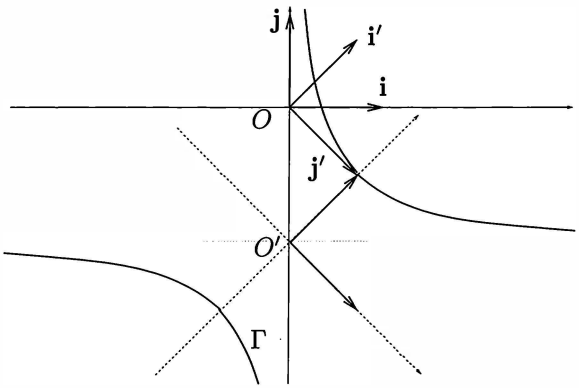
Remarquons que la matrice orthogonale P correspond à la matrice de la symétrie par rapport à la droite vectorielle dirigée par le vecteur $\mathbf{eps} = {}^t(1, \sqrt{2}-1)$ (voir l'exemple de la p. 499).

Notons $\mathbf{i}' = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$ et $\mathbf{j}' = \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}$ les deux vecteurs qui forment une base orthonormale du plan et dans laquelle la matrice de la forme quadratique $(x, y) \in \mathbb{R}^2 \mapsto 2xy$ est diagonale. Afin d'écrire l'équation de la conique dans le repère $(O; \mathbf{i}', \mathbf{j}')$, remarquons que l'équation de la conique s'écrit matriciellement ${}^t\mathrm{XPD}{}^t\mathrm{PX} + {}^t\mathrm{BX} = 1$ où ${}^t\mathrm{B} = (2\sqrt{2}, 0)$. En posant $\mathrm{X}' = {}^t\mathrm{PX}$ on a

$$ {}^t\mathrm{X}'\mathrm{DX}' + {}^t\mathrm{BPX}' = 1, $$

et en notant x' et y' les composantes de X' on obtient dans le repère $(O; \mathbf{i}', \mathbf{j}')$ l'équation de la conique $\mathcal{C}$ suivante : $x'^2 - y'^2 + 2x' + 2y' = 1$. En regroupant les termes comportant x' et ceux comportant y', on obtient pour équation de la conique, l'équation $(x'+1)^2 - (y'-1)^2 = 1$. Autrement dit, en effectuant le changement de variables $x'' = x' + 1$ et $y'' = y' - 1$, ce qui correspond à effectuer une translation du repère $(O; \mathbf{i}', \mathbf{j}')$, l'équation de la conique s'écrit $x''^2 - y''^2 = 1$.

3 - On trace successivement, le repère d'origine $(O; \mathbf{i}, \mathbf{j})$, le repère $(O; \mathbf{i}', \mathbf{j}')$ puis le repère $(O'; \mathbf{i}', \mathbf{j}')$ obtenu par la translation. Dans ce dernier repère, on trace l'hyperbole d'équation $x''^2 - y''^2 = 1$ dont les asymptotes sont les deux droites $x'' = y''$ et $x'' = -y''$. On obtient la représentation graphique suivante pour l'hyperbole Γ :



Solution de l'exercice 8.10

Notons q la forme quadratique $(x, y, z) \in \mathbb{R}^3 \mapsto x^2 + 4y^2 + 5z^2 - 4xy$. La matrice A de cette forme quadratique dans la base canonique de $\mathbb{R}^3$ vaut

$$ \mathrm{A} = \begin{pmatrix} 1 & -2 & 0 \\ -2 & 4 & 0 \\ 0 & 0 & 5 \end{pmatrix}. $$

Pour connaître le type de la quadrique, on peut déterminer le signe des valeurs propres de cette matrice. Ces valeurs propres sont les racines du polynôme caractéristique de la matrice A. On a

$$\det(\mathrm{A}-\lambda\mathrm{Id}) = \begin{vmatrix} 1-\lambda & -2 & 0 \\ -2 & 4-\lambda & 0 \\ 0 & 0 & 5-\lambda \end{vmatrix} = -\lambda(5-\lambda)^2$$

et par conséquent, les valeurs propres de A sont 0 et 5. La quadrique est donc soit du type paraboloïde, soit du type cylindrique. Pour préciser le type de la quadrique, une première méthode consiste à calculer les vecteurs propres associés aux valeurs propres 0 et 5, en déduire une base dans laquelle la forme quadratique q est diagonale et ensuite écrire les termes linéaire $-2x+4y$ dans cette base.

On peut aussi directement obtenir le type de la quadrique en utilisant la décomposition de Gauss de la forme quadratique q : on écrit, pour tout $(x, y, z) \in \mathbb{R}^3$:

$$x^2+4y^2+5z^2-4xy = (x-2y)^2+5z^2,$$

et en posant $X = x-2y$ on a

$$x^2+4y^2+5z^2-4xy-2x+4y=0 \iff (X-1)^2+5Z^2=1.$$

La quadrique est par conséquent un cylindre elliptique.

Solution de l'exercice 8.11

1 - Soit φ la forme polaire associée à la forme quadratique q. La matrice de la forme bilinéaire φ dans une base $\{\mathbf{f}_i\}_{i\in\{1,\dots,n\}}$ vaut

$$\mathrm{Mat}(\varphi,\{\mathbf{f}_i\}) = \begin{pmatrix} \varphi(\mathbf{f}_1,\mathbf{f}_1) & \varphi(\mathbf{f}_1,\mathbf{f}_2) & \varphi(\mathbf{f}_1,\mathbf{f}_n) \\ \varphi(\mathbf{f}_2,\mathbf{f}_1) & \varphi(\mathbf{f}_2,\mathbf{f}_2) & \varphi(\mathbf{f}_2,\mathbf{f}_n) \\ \vdots & \vdots & \vdots \\ \varphi(\mathbf{f}_n,\mathbf{f}_1) & \varphi(\mathbf{f}_n,\mathbf{f}_2) & \varphi(\mathbf{f}_n,\mathbf{f}_n) \end{pmatrix}.$$

En particulier, ses coefficients diagonaux valent $\varphi(\mathbf{f}_i,\mathbf{f}_i) = q(\mathbf{f}_i)$ de sorte que la trace de $\mathrm{Mat}(\varphi,\{\mathbf{f}_i\})$, c'est-à-dire la trace de la forme quadratique q vaut $q(\mathbf{f}_1)+q(\mathbf{f}_2)+\cdots+q(\mathbf{f}_n)$. Par conséquent, si tous les réels $q(\mathbf{f}_i)$ sont nuls alors la trace de q est nulle.

2 - a) On utilise le théorème 8.10 de la page 505. Celui-ci indique qu'il existe une base orthonormale pour le produit scalaire canonique de $\mathbb{R}^n$ et dans laquelle la matrice de la forme quadratique q est diagonale. Si on désigne par $\{\mathbf{e}_i\}_{i\in\{1,\dots,n\}}$ la base obtenue et par λ_i $(i \in \{1,\dots,n\})$ les coefficients diagonaux de la matrice de q dans cette base alors on a, pour tout $(x_1,\dots,x_n) \in \mathbb{R}^n$

$$q(x_1,x_2,\dots,x_n) = \lambda_1x_1^2+\lambda_2x_2^2+\cdots+\lambda_nx_n^2.$$

Puisque, par hypothèse, la trace de la forme quadratique q est nulle, la trace de la matrice représentant q dans la base $\{\mathbf{e}_i\}_{i\in\{1,\dots,n\}}$ est nulle. Or cette matrice étant la matrice diagonale composée des réels λ_i, on a $\lambda_1+\lambda_2+\cdots+\lambda_n=0$.

b) En remarquant que $q(1,1,\ldots,1) = \lambda_1 + \lambda_2 + \cdots + \lambda_n = 0$, il suffit de normer le vecteur $(1,1,\ldots,1)$, c'est-à-dire de le diviser par sa norme, pour obtenir un vecteur $\mathbf{f}_1$ de $\mathbb{R}^n$ de norme égale à 1 et tel que $q(\mathbf{f}_1) = 0$. On peut donc choisir $\mathbf{f}_1 = \frac{1}{\sqrt{n}}(1,1,\ldots,1)$.

c) Intéressons-nous au cas de la dimension 1. Dans $\mathbb{R}$, une forme quadratique est représentée par sa matrice qui est une matrice à un seul élément. Si la trace d'une telle matrice est nulle, c'est que la matrice et par conséquent la forme quadratique est nulle. La seule forme quadratique de trace nulle sur $\mathbb{R}$ est la forme quadratique nulle. Il existe donc un vecteur $f_1 \in \mathbb{R}$ non nul tel que $q(f_1) = 0$.
Supposons que la propriété à démontrer soit satisfaite pour un entier $n-1 \in \mathbb{N}^*$, c'est-à-dire que pour toute forme quadratique $\widetilde{q}$ sur $\mathbb{R}^{n-1}$ dont la trace est nulle il existe une base orthonormale $\{\mathbf{f}_i\}_{i\in\{1,\ldots,n-1\}}$ de $\mathbb{R}^{n-1}$ telle que pour tout $i \in \{1,\ldots,n-1\}$ on ait $\widetilde{q}(\mathbf{f}_i) = 0$. Montrons que cette propriété est vraie pour l'entier n. Considérons une forme quadratique q sur $\mathbb{R}^n$ dont la trace est nulle. D'après la question précédente, il existe un vecteur $\mathbf{f}_1$ de $\mathbb{R}^n$ de norme égale à 1 et tel que $q(\mathbf{f}_1) = 0$. Notons E l'orthogonal du vecteur $\mathbf{f}_1$ pour le produit scalaire canonique. Ce sous-espace vectoriel est de dimension $n-1$ et ne contient pas $\mathbf{f}_1$. En appliquant l'hypothèse de récurrence à la forme quadratique q restreinte à E qui est aussi de trace nulle, on en déduit qu'il existe une base orthonormée $\{\mathbf{f}_i\}_{i\in\{2,\ldots,n\}}$ de E telle que pour tout $i \in \{1,2,\ldots,n\}$ on ait $q(\mathbf{f}_i) = 0$. La base $\{\mathbf{f}_i\}_{i\in\{1,\ldots,n\}}$ est une base orthonormée de $\mathbb{R}^n$ qui satisfait ainsi

$$\forall i \in \{1,2,\ldots,n\} \quad q(\mathbf{f}_i) = 0.$$

Solution de l'exercice 8.12

1 - La[119] figure 17 représente une boule dans $\mathbb{R}^2$, un point $\mathbf{u}$ à l'extérieur de la boule, sa projection $\mathbf{f}$ et un point $\mathbf{w}$ à l'intérieur de la boule « indiquant » que la distance entre $\mathbf{u}$ et $\mathbf{f}$ est plus petite que celle entre $\mathbf{u}$ et $\mathbf{w}$.

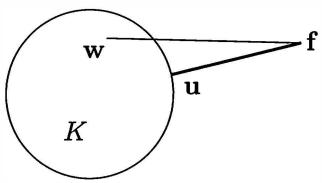


Fig. 17 Illustration graphique de la propriété.

[119] L'un des auteurs ayant eu à modifier cet exercice le vendredi 13 juin 2008, journée du désastre de l'équipe de France de football face aux Pays-Bas lors du premier tour de l'Euro 2008, nous ne saurions garantir que le ton neutre et courtois habituellement utilisé pour les corrections des exercices ait été respecté pour celui ci.

2 - Soit $(\mathbf{x},\mathbf{y}) \in E^2$. Par définition de la norme associée à un produit scalaire, on a $\|\mathbf{x}+\mathbf{y}\|^2 = \langle \mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y}\rangle$. En utilisant la bilinéarité et symétrie du produit scalaire, on en déduit que

$$\|\mathbf{x}+\mathbf{y}\|^2 = \langle \mathbf{x},\mathbf{x}\rangle + \langle \mathbf{x},\mathbf{y}\rangle + \langle \mathbf{y},\mathbf{x}\rangle + \langle \mathbf{y},\mathbf{y}\rangle = \|\mathbf{x}\|^2 + 2\langle \mathbf{x},\mathbf{y}\rangle + \|\mathbf{y}\|^2.$$

De la même manière, on a $\|\mathbf{x}-\mathbf{y}\|^2 = \|\mathbf{x}\|^2 - 2\langle \mathbf{x},\mathbf{y}\rangle + \|\mathbf{y}\|^2$ de sorte que la somme $\|\mathbf{x}+\mathbf{y}\|^2 + \|\mathbf{x}-\mathbf{y}\|^2$ est égale à $2\left(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2\right)$.

3 - Dans le cas où $E = \mathbb{R}^2$ est muni du produit scalaire canonique, la figure 18 indique que l'identité du parallélogramme signifie que la somme des carrés des côtés d'un parallélogramme est égale à la somme des carrés de ces diagonales.

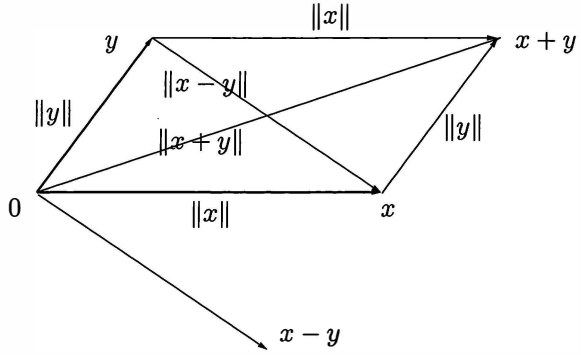


Fig. 18 Illustration graphique de l'identité du parallélogramme.

4 - a) Pour tout $(p,q) \in \mathbb{N}^2$, en utilisant la relation (31) avec $\mathbf{x} = \mathbf{f} - \frac{1}{2}(\mathbf{w}_p + \mathbf{w}_q)$ et $\mathbf{y} = \frac{1}{2}(\mathbf{w}_p - \mathbf{w}_q)$, on obtient :

$$2\left(\left\|\mathbf{f} - \tfrac{1}{2}(\mathbf{w}_p+\mathbf{w}_q)\right\|^2 + \left\|\tfrac{1}{2}(\mathbf{w}_p-\mathbf{w}_q)\right\|^2\right) = \left(\|\mathbf{f}-\mathbf{w}_p\|^2 + \|\mathbf{f}-\mathbf{w}_q\|^2\right).$$

b) Pour $n \in \mathbb{N}$, notons $d_n = \|\mathbf{f}-\mathbf{w}_n\|$ et $d = \inf_{\mathbf{w}\in K}\|\mathbf{f}-\mathbf{w}\|$. Puisque le sous-ensemble K est convexe, pour tous entiers p et q, on a $\frac{1}{2}(\mathbf{w}_p+\mathbf{w}_q) \in K$ et par définition de d, on a $\left\|\mathbf{f} - \frac{1}{2}(\mathbf{w}_p+\mathbf{w}_q)\right\| \geqslant d$. On déduit de la question précédente que pour tout $(p,q) \in \mathbb{N}^2$

$$d_p^2 + d_q^2 = 2\left(\left\|\mathbf{f} - \tfrac{1}{2}(\mathbf{w}_p+\mathbf{w}_q)\right\|^2 + \left\|\tfrac{1}{2}(\mathbf{w}_p-\mathbf{w}_q)\right\|^2\right) \geqslant 2d^2 + \tfrac{1}{2}\|\mathbf{w}_p-\mathbf{w}_q\|^2.$$

Autrement dit, on a $\|\mathbf{w}_p-\mathbf{w}_q\|^2 \leqslant 2(d_p^2 + d_q^2 - 2d^2)$.

Or, par définition de d, on a $\lim_{n\to+\infty} d_n^2 = d^2$ ce qui implique que[120] :

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad \left(n \geqslant N \Longrightarrow |d_n^2 - d^2| \leqslant \tfrac{1}{4}\varepsilon\right).$$

Ainsi, pour tout réel $\varepsilon \in \mathbb{R}_+^*$ il existe un entier $N \in \mathbb{N}$ tel que pour tout couple d'entiers $(p,q) \in \mathbb{N}^2$

$$(\ p \geqslant N \quad \text{et} \quad q \geqslant N\ ) \quad \Longrightarrow \quad \|\mathbf{w}_p - \mathbf{w}_q\|^2 \leqslant 2(\tfrac{1}{4}\varepsilon + \tfrac{1}{4}\varepsilon) = \varepsilon.$$

[120] Voir la définition 5.1 p. 168 du *Cours de première année*.

On en déduit, voir la définition 7.20 p. 398, que la suite $(\mathbf{w}_n)_n$ est une suite de Cauchy. P@!?#>, ils viennent d'en prendre un troisième...

c) Puisque l'espace vectoriel E est complet, d'après la définition 7.21 p. 401, la suite de Cauchy $(\mathbf{w}_n)_n$ est une suite convergente. Notons $\mathbf{u} \in E$ sa limite ; puisque l'ensemble K est fermé, et que les éléments $\mathbf{w}_n$ appartiennent à K, la limite $\mathbf{u}$ est aussi un élément de K et vérifie

$$\|\mathbf{f} - \mathbf{u}\| = \inf_{\mathbf{w} \in K} \|\mathbf{f} - \mathbf{w}\|.$$

5 - a) Écrivons $\mathbf{f} - \mathbf{v} = (\mathbf{f} - \mathbf{u}) - t(\mathbf{w} - \mathbf{u})$. On a alors[121]

$$\|\mathbf{f} - \mathbf{v}\|^2 = \|\mathbf{f} - \mathbf{u}\|^2 - 2t\langle \mathbf{f} - \mathbf{u}, \mathbf{w} - \mathbf{u}\rangle + t^2\|\mathbf{w} - \mathbf{u}\|^2.$$

b) Par définition de $\mathbf{u}$, pour tout $\mathbf{v} \in K$ on a $\|\mathbf{f} - \mathbf{u}\| \leqslant \|\mathbf{f} - \mathbf{v}\|$. Par conséquent, d'après la question précédente on obtient

$$\|\mathbf{f} - \mathbf{u}\|^2 \leqslant \|\mathbf{f} - \mathbf{u}\|^2 - 2t\langle \mathbf{f} - \mathbf{u}, \mathbf{w} - \mathbf{u}\rangle + t^2\|\mathbf{w} - \mathbf{u}\|^2.$$

c) D'après la question précédente, pour tout $\mathbf{w} \in K$ et pour tout $t \in]0, 1]$ on a $2\langle \mathbf{f} - \mathbf{u}, \mathbf{w} - \mathbf{u}\rangle \leqslant t\|\mathbf{w} - \mathbf{u}\|^2$ de sorte que lorsque t tend vers 0, on obtient

$$\langle \mathbf{f} - \mathbf{u}, \mathbf{w} - \mathbf{u}\rangle \leqslant 0. \tag{35}$$

6 - a) Écrivons $\mathbf{f} - \mathbf{w} = (\mathbf{f} - \mathbf{u}) - (\mathbf{w} - \mathbf{u})$. On a alors

$$\|\mathbf{f} - \mathbf{w}\|^2 = \|\mathbf{f} - \mathbf{u}\|^2 - 2\langle \mathbf{f} - \mathbf{u}, \mathbf{w} - \mathbf{u}\rangle + \|\mathbf{w} - \mathbf{u}\|^2.$$

En utilisant le fait que $\mathbf{u}$ vérifie la relation (33), on obtient immédiatement l'inégalité suivante : $\|\mathbf{f} - \mathbf{w}\|^2 \geqslant \|\mathbf{f} - \mathbf{u}\|^2$.

b) D'après la question précédente l'élément $\mathbf{u}$ vérifie la relation (32). Et de 4 ! Ils sont vraiment très forts ces Hollandais.

7 - a) Supposons qu'il existe deux éléments $\mathbf{u}_1 \in K$ et $\mathbf{u}_2 \in K$ satisfaisant la relation (35), *i.e.*

$$\forall \mathbf{w} \in K \quad \langle \mathbf{f} - \mathbf{u}_1, \mathbf{w} - \mathbf{u}_1\rangle \leqslant 0, \qquad \forall \mathbf{w} \in K \quad \langle \mathbf{f} - \mathbf{u}_2, \mathbf{w} - \mathbf{u}_2\rangle \leqslant 0.$$

En choisissant $\mathbf{w} = \mathbf{u}_2$ dans la première de ces deux relations et $\mathbf{w} = \mathbf{u}_1$ dans la seconde, on obtient d'une part $\langle \mathbf{f} - \mathbf{u}_1, \mathbf{u}_2 - \mathbf{u}_1\rangle \leqslant 0$ et d'autre part $\langle \mathbf{f} - \mathbf{u}_2, \mathbf{u}_1 - \mathbf{u}_2\rangle \leqslant 0$. En effectuant la somme de ces deux relations, on en déduit que $\|\mathbf{u}_1 - \mathbf{u}_2\|^2 \leqslant 0$ et donc que $\mathbf{u}_1 = \mathbf{u}_2$ ce qui prouve qu'il existe un seul $\mathbf{u} \in K$ vérifiant (35).

b) D'après la question 5-c, on sait que pour tout $\mathbf{f} \in E$ il existe $\mathbf{u} \in K$ tel que $\|\mathbf{f} - \mathbf{u}\| = \inf_{\mathbf{w} \in K} \|\mathbf{f} - \mathbf{w}\|$. Selon la question 6 cet élément $\mathbf{u}$ vérifie la relation (35) et il y a d'après la question 8 au plus un élément vérifiant cette relation. On en déduit l'unicité de l'élément $\mathbf{u} \in K$ vérifiant $\|\mathbf{f} - \mathbf{u}\| = \inf_{\mathbf{w} \in K} \|\mathbf{f} - \mathbf{w}\|$.

---

[121] Rappelons que dans un espace préhilbertien E on a pour tout $(\mathbf{x}, \mathbf{y}) \in E^2$ l'identité $\|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + 2\langle \mathbf{x}, \mathbf{y}\rangle + \|\mathbf{y}\|^2$.

CHAPITRE 9

Limite et continuité

L'objectif de ce chapitre est de généraliser les notions de limite et de continuité, vues dans le cas des fonctions réelles de la variable réelle dans le *Cours de première année* (au chapitre 13) puis dans le cas des fonctions complexes de la variable complexe dans le chapitre 4 de cet ouvrage, aux fonctions d'un espace métrique (E, d_E) dans un espace métrique (F, d_F). Nous insisterons plus particulièrement sur le cas des fonctions de $\mathbb{R}^n$ dans $\mathbb{R}^p$ puis sur le cas des applications linéaires compte tenu de leur importance dans la pratique. Nous aborderons enfin dans la dernière partie du chapitre le *théorème du point fixe* aux nombreuses applications en mathématiques.

9.1 Limite et continuité dans un espace métrique

Les résultats de cette première partie du chapitre énoncés dans le cadre d'espaces métriques sont transposables au cas d'espaces vectoriels normés en prenant la distance subordonnée à la norme.

9.1.1 Limite d'une fonction

Dans le cours de première année nous avions défini la limite pour une application à valeurs réelles f définie sur un sous-ensemble D de $\mathbb{R}$ de la manière suivante : f admet pour limite $\ell \in \mathbb{R}$ en $x_0 \in \overline{D}$ si

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists \eta \in \mathbb{R}_+^* \quad \forall x \in D \quad \big(|x - x_0| \leqslant \eta \implies |f(x) - \ell| \leqslant \varepsilon\big).$$

Dans cette définition $|x - x_0|$ mesure la distance entre x et x_0 et $|f(x) - \ell|$ la distance entre $f(x)$ et ℓ. Elle s'interprète de la manière suivante : f admet pour limite ℓ en x_0 si pour tout réel strictement positif ε la distance de $f(x)$ à ℓ peut être rendue inférieure à ε pour tous les réels x assez proches de x_0 (dont la distance à x_0 est inférieure à un certain réel strictement positif η déterminé en fonction de la valeur prise pour ε).

Dans le cas d'une application d'un espace métrique (E, d_E) dans un espace métrique (F, d_F), cette définition se généralise de manière naturelle : on mesure la distance entre x et x_0 en utilisant la distance d_E et on mesure la distance entre $f(x)$ et ℓ en utilisant la distance d_F.

DÉFINITION 9.1 *Soient (E, d_E) et (F, d_F) deux espaces métriques, D un sous-ensemble de E et f une application de D dans F. On dit que f admet une limite en $x_0 \in \overline{D}$ s'il existe un élément $\ell \in F$ tel que :*

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists \eta \in \mathbb{R}_+^* \quad \forall x \in D \quad \big(d_E(x, x_0) \leqslant \eta \implies d_F(f(x), \ell) \leqslant \varepsilon\big).$$

L'élément ℓ de F est appelé limite *de f en x_0 .*

Remarques

1. Il résulte de la définition 9.1 que l'application f n'a pas de limite en $x_0 \in \overline{D}$ si[1] quel que soit $\ell \in F$,

$$\exists \varepsilon \in \mathbb{R}_+^* \quad \forall \eta \in \mathbb{R}_+^* \quad \exists x \in D \quad \big(d_E(x, x_0) \leqslant \eta \text{ et } d_F(f(x), \ell) > \varepsilon\big).$$

2. L'existence et la valeur d'une limite ne changent pas si on remplace les distances d_E et d_F par des distances équivalentes. Supposons que f admette $\ell \in F$ pour limite en $x_0 \in \overline{D}$ et considérons deux distances d'_E et d'_F respectivement équivalentes à d_E et d_F avec[2]

$$\begin{aligned} &\forall (x_1, x_2) \in E^2 \quad C_1\, d_E(x_1, x_2) \leqslant d'_E(x_1, x_2) \leqslant C_2\, d_E(x_1, x_2) \\ \text{et} \quad &\forall (y_1, y_2) \in F^2 \quad C_3\, d_F(y_1, y_2) \leqslant d'_F(y_1, y_2) \leqslant C_4\, d_F(y_1, y_2), \end{aligned}$$

où C_1, C_2, C_3 et C_4 désignent des réels strictement positifs. Pour tout $\varepsilon' \in \mathbb{R}_+^*$, on obtient en considérant l'assertion de la définition 9.1 avec $\varepsilon = \varepsilon'/C_4$, l'existence d'un réel strictement positif η tel que pour tout $x \in D$ vérifiant $d_E(x, x_0) \leqslant \eta$ on ait $d'_F(f(x), \ell) \leqslant C_4\, d_F(f(x), \ell) \leqslant \varepsilon'$. Par ailleurs, puisque $d_E(x, x_0) \leqslant \eta \Leftrightarrow d'_E(x, x_0) \leqslant \eta'$ où $\eta' = C_1\, \eta$, on a

$$\forall \varepsilon' \in \mathbb{R}_+^* \quad \exists \eta' \in \mathbb{R}_+^* \quad \forall x \in D \quad \big(d'_E(x, x_0) \leqslant \eta' \implies d'_F(f(x), \ell) \leqslant \varepsilon'\big)$$

ce qui signifie que f admet également pour limite ℓ en x_0 lorsque E est muni de la distance d'_E et F est muni de la distance d'_F.

3. Dans le cas où $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ sont deux $\mathbb{K}$-espaces vectoriels normés, la définition 9.1 indique qu'une application f d'un sous-ensemble D de E à valeurs dans F admet pour limite ℓ en $x_0 \in \overline{D}$ si

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists \eta \in \mathbb{R}_+^* \quad \forall x \in D \quad \big(\|x - x_0\|_E \leqslant \eta \implies \|f(x) - \ell\|_F \leqslant \varepsilon\big).$$

Par ailleurs, dans le cas où les $\mathbb{K}$-espaces vectoriels sont de dimension finie, toutes les normes étant équivalentes, en pratique on choisit celle pour laquelle l'assertion précédente est la plus aisée à établir. ◇

---

[1] Cette assertion est la négation de celle donnée à la définition 9.1.

[2] Voir la définition 7.2 p. 352.

Dans la pratique, on est souvent amené à s'intéresser à la limite de fonctions de $\mathbb{R}^n$ dans $\mathbb{R}^p$ et nous consacrerons une partie de ce chapitre aux méthodes spécifiques de calcul de limites pour les fonctions de $\mathbb{R}^n$ dans $\mathbb{R}^p$, voir p. 565. La définition 9.1 permet déjà dans certains cas d'étudier la limite de telles fonctions.

Exemples

1. Montrons que l'application $f : (x_1, x_2) \in \mathbb{R}^2_* \mapsto \dfrac{x_1 x_2}{|x_1| + |x_2|}$ admet pour limite 0 en $(0,0)$ lorsque $\mathbb{R}^2$ est muni de la distance subordonnée à la norme infinie $(x_1, x_2) \in \mathbb{R}^2 \mapsto \max(|x_1|, |x_2|)$, notée $\|\cdot\|_\infty$, et $\mathbb{R}$ est muni de la distance usuelle. D'après la définition 9.1, il s'agit d'établir que

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists \eta \in \mathbb{R}_+^* \quad \forall (x_1, x_2) \in \mathbb{R}^2_* \quad \big(\|(x_1, x_2)\|_\infty \leqslant \eta \implies |f(x_1, x_2)| \leqslant \varepsilon\big).$$

Compte tenu des inégalités

$$|x_1||x_2| \leqslant \max(|x_1|, |x_2|)^2 \quad \text{et} \quad |x_1| + |x_2| \geqslant \max(|x_1|, |x_2|),$$

on a pour tout $(x_1, x_2) \in \mathbb{R}^2_*$

$$|f(x_1, x_2)| = \frac{|x_1||x_2|}{|x_1| + |x_2|} \leqslant \max(|x_1|, |x_2|) = \|(x_1, x_2)\|_\infty.$$

Soit ε un réel strictement positif fixé et $\eta = \varepsilon$; on déduit de la majoration précédente que $|f(x_1, x_2)| \leqslant \varepsilon$ pour tout élément $(x_1, x_2) \in \mathbb{R}^2_*$ tel que $\|(x_1, x_2)\|_\infty \leqslant \eta$. L'assertion est démontrée ; l'application f admet pour limite 0 en $(0,0)$.

2. Considérons l'application $f : (x_1, x_2) \in \mathbb{R}^2_* \mapsto \dfrac{x_1 x_2}{x_1^2 + x_2^2}$ où $\mathbb{R}^2$ est muni de la distance subordonnée à la norme[3] $(x_1, x_2) \in \mathbb{R}^2 \mapsto |x_1| + |x_2|$, notée $\|\cdot\|_1$, et $\mathbb{R}$ est muni de la distance usuelle. On peut remarquer qu'en restriction à l'axe des abscisses $(x_2 = 0)$ ou en restriction à l'axe des ordonnées $(x_1 = 0)$, f est l'application nulle ; toutefois, l'application f n'admet pas pour limite 0 en $(0,0)$. Pour l'établir, il s'agit de montrer que

$$\exists \varepsilon \in \mathbb{R}_+^* \quad \forall \eta \in \mathbb{R}_+^* \quad \exists (x_1, x_2) \in \mathbb{R}^2_* \quad \big(\|(x_1, x_2)\|_1 \leqslant \eta \ \text{ et } \ |f(x_1, x_2)| > \varepsilon\big).$$

Soient $\eta \in \mathbb{R}_+^*$ et $(x_1, x_2) = (\frac{1}{2}\eta, \frac{1}{2}\eta)$; on a $\|(x_1, x_2)\|_1 = |x_1| + |x_2| = \eta$ et

$$|f(x_1, x_2) - 0| = \left|\frac{x_1 x_2}{x_1^2 + x_2^2}\right| = \frac{\frac{1}{4}\eta^2}{\frac{1}{4}\eta^2 + \frac{1}{4}\eta^2} = \frac{1}{2}.$$

On en déduit qu'il existe $\varepsilon \in \mathbb{R}_+^*$ (prendre $\varepsilon = \frac{1}{4}$) tel que pour tout $\eta \in \mathbb{R}_+^*$ l'élément $(x_1, x_2) = (\frac{1}{2}\eta, \frac{1}{2}\eta)$ de $\mathbb{R}^2$ vérifie $|f(x_1, x_2)| > \varepsilon$. L'application f n'admet donc pas pour limite 0 en $(0,0)$. On peut en fait montrer que l'application f n'admet pas de limite en $(0,0)$. En effet, pour tout $\ell \in \mathbb{R}$, il existe $\varepsilon \in \mathbb{R}_+^*$ (prendre $\varepsilon = \frac{1}{4}|1 - 2\ell|$) tel que pour tout $\eta \in \mathbb{R}_+^*$ l'élément $(x_1, x_2) = (\frac{1}{2}\eta, \frac{1}{2}\eta)$ de $\mathbb{R}^2$ est tel que $|f(x_1, x_2) - \ell| > \varepsilon$.

---

[3] Voir la proposition 7.4 p. 362.

PROPOSITION 9.1 *Soient (E, d_E) et (F, d_F) deux espaces métriques, D un sous-ensemble de E et f une application de D dans F. Si f admet une limite en $x_0 \in \overline{D}$ alors cette limite est unique. Elle est notée* $\lim\limits_{\substack{x \to x_0 \\ x \in D}} f(x)$.

Démonstration Raisonnons par l'absurde et supposons que l'application f admette deux limites ℓ_1 et ℓ_2 distinctes en $x_0 \in \overline{D}$. Considérons le réel strictement positif $\varepsilon = \frac{1}{3} d_F(\ell_1, \ell_2)$; d'après la définition 9.1, on a d'une part

$$\exists \eta_1 \in \mathbb{R}_+^* \quad \forall x \in D \quad \big(d_E(x, x_0) \leqslant \eta_1 \implies d_F(f(x), \ell_1) \leqslant \varepsilon\big)$$

et d'autre part

$$\exists \eta_2 \in \mathbb{R}_+^* \quad \forall x \in D \quad \big(d_E(x, x_0) \leqslant \eta_2 \implies d_F(f(x), \ell_2) \leqslant \varepsilon\big).$$

Notons $\eta = \min(\eta_1, \eta_2)$; pour tout $x \in B(x_0, \eta) \cap D$, d'après les assertions précédentes, on a :

$$d_F(\ell_1, \ell_2) \leqslant d_F(f(x), \ell_1) + d_F(f(x), \ell_2) \leqslant 2\epsilon \leqslant \tfrac{2}{3} d_F(\ell_1, \ell_2).$$

Cela constitue une contradiction : comme $\ell_1 \neq \ell_2$ on a $d_F(\ell_1, \ell_2) \neq 0$ et cette assertion implique que $1 \leqslant \frac{2}{3}$. Il y a donc unicité de la limite. □

Cas des fonctions à valeurs réelles

Dans le cas de fonctions à valeurs réelles, on peut introduire la notion de limite infinie en généralisant les définitions données pour les fonctions de $\mathbb{R}$ dans $\mathbb{R}$ dans le *Cours de première année* p. 573.

DÉFINITION 9.2 *Soient (E, d_E) un espace métrique, D un sous-ensemble de E et f une application de D dans $\mathbb{R}$. On munit $\mathbb{R}$ de la distance usuelle.*

✗ *On dit que f admet pour limite $+\infty$ en $x_0 \in \overline{D}$ si*

$$\forall \kappa \in \mathbb{R} \quad \exists \eta \in \mathbb{R}_+^* \quad \forall x \in D \quad \big(d_E(x, x_0) \leqslant \eta \implies f(x) \geqslant \kappa\big).$$

✗ *On dit que f admet pour limite $-\infty$ en $x_0 \in \overline{D}$ si*

$$\forall \kappa \in \mathbb{R} \quad \exists \eta \in \mathbb{R}_+^* \quad \forall x \in D \quad \big(d_E(x, x_0) \leqslant \eta \implies f(x) \leqslant \kappa\big).$$

Ces définitions sont conformes à la définition 9.1 lorsque l'on munit $\overline{\mathbb{R}}$ de la distance[(4)] :

$$d : (x, y) \in \overline{\mathbb{R}} \times \overline{\mathbb{R}} \mapsto |\phi(x) - \phi(y)| \quad \text{où} \quad \phi : x \in \overline{\mathbb{R}} \longmapsto \begin{cases} \dfrac{x}{1+|x|} & \text{si } x \in \mathbb{R} \\ 1 & \text{si } x = +\infty \\ -1 & \text{si } x = -\infty \end{cases}$$

[(4)] Voir le chap. 7 p. 347.

En effet, d'après la définition 9.1, l'application f de D dans $\overline{\mathbb{R}}$ muni de la distance d admet pour limite $+\infty$ en $x_0 \in \overline{D}$ si

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists \eta \in \mathbb{R}_+^* \quad \forall x \in D \quad \big(d_E(x, x_0) \leqslant \eta \implies d(f(x), +\infty) \leqslant \varepsilon\big).$$

Or,

$$d(f(x), +\infty) = \left| \frac{f(x)}{1 + f(x)} - 1 \right| = \frac{1}{|1 + f(x)|}$$

donc pour une fonction f à valeurs positives, $d(f(x), +\infty) \leqslant \varepsilon$ équivaut à $f(x) + 1 \geqslant 1/\varepsilon$. En posant $\kappa = 1/\varepsilon - 1$, on établit le lien entre les deux assertions.

Dans le cas des fonctions à valeurs réelles, on dispose également du résultat suivant qui généralise le théorème d'encadrement pour les fonctions réelles de la variable réelle[5].

THÉORÈME 9.1 (Théorème d'encadrement)
Soient (E, d_E) un espace métrique, D un sous-ensemble de E et f, g, h trois applications de D dans $\mathbb{R}$;

- *si les applications f et h ont pour limite le réel ℓ en $x_0 \in \overline{D}$,*
- *et si pour tout $x \in D$ on a $f(x) \leqslant g(x) \leqslant h(x)$,*

alors g admet pour limite ℓ en x_0.

Démonstration Puisque les applications f et h ont pour limite le réel ℓ en x_0, d'après la définition 9.1 on a :

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists \eta_1 \in \mathbb{R}_+^* \quad \forall x \in D \quad \big(d_E(x, x_0) \leqslant \eta_1 \implies |f(x) - \ell| \leqslant \varepsilon\big),$$
$$\text{et} \quad \forall \varepsilon \in \mathbb{R}_+^* \quad \exists \eta_2 \in \mathbb{R}_+^* \quad \forall x \in D \quad \big(d_E(x, x_0) \leqslant \eta_2 \implies |h(x) - \ell| \leqslant \varepsilon\big).$$

Soit ε un réel strictement positif fixé ; pour tout $x \in V$ où V est l'ensemble $\overline{B}(x_0, \eta_1) \cap \overline{B}(x_0, \eta_2) \cap D$ on a $-\varepsilon \leqslant f(x) - \ell \leqslant \varepsilon$ et $-\varepsilon \leqslant h(x) - \ell \leqslant \varepsilon$. Comme par hypothèse, pour tout $x \in D$ on a $f(x) \leqslant g(x) \leqslant h(x)$, on en déduit que pour $x \in V$ on a $|g(x) - \ell| \leqslant \varepsilon$ car

$$-\varepsilon \leqslant f(x) - \ell \leqslant g(x) - \ell \leqslant h(x) - \ell \leqslant \varepsilon.$$

Puisque V est un voisinage de x_0, il existe $\eta \in \mathbb{R}_+^*$ tel que $\overline{B}(x_0, \eta) \subset V$. On en conclut que :

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists \eta \in \mathbb{R}_+^* \quad \forall x \in D \quad \big(d_E(x, x_0) \leqslant \eta \implies |g(x) - \ell| \leqslant \varepsilon\big),$$

c'est-à-dire que g admet pour limite ℓ en x_0. □

[5] Voir le théorème 13.1 p. 577 du *Cours de première année*.

Remarque On s'apercoit dans la démonstration du théorème 9.1, que l'encadrement $f(x) \leqslant g(x) \leqslant h(x)$ n'est requis « qu'au voisinage » de x_0, c'est-à-dire sur un ensemble de la forme $B(x_0, r) \cap D$ où $r \in \mathbb{R}_+^*$. ◇

La proposition suivante, utilisée conjointement avec le théorème d'encadrement est d'un usage pratique très intéressant. Elle ramène l'étude de la limite d'une fonction entre deux espaces métriques, à l'étude de la limite d'une fonction à valeurs réelles.

PROPOSITION 9.2 *Soient (E, d_E) et (F, d_F) deux espaces métriques, D un sous-ensemble de E et f une application de D dans F.*
L'application f admet pour limite $\ell \in F$ en $x_0 \in \overline{D}$ si et seulement si l'application $x \in D \mapsto d_F(x, \ell) \in \mathbb{R}$, où l'ensemble $\mathbb{R}$ est muni de la distance usuelle, admet pour limite le réel 0 en x_0.

Démonstration La justification de ce résultat est immédiate en utilisant la définition 9.1. En effet, l'application f admet pour limite $\ell \in F$ en $x_0 \in \overline{D}$ si

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists \eta \in \mathbb{R}_+^* \quad \forall x \in D \quad \big(d_E(x, x_0) \leqslant \eta \implies d_F(f(x), \ell) \leqslant \varepsilon\big).$$

Munissons $\mathbb{R}$ de la distance usuelle ; l'application $\phi : x \in D \mapsto d_F(x, \ell) \in \mathbb{R}$ admet pour limite le réel 0 en x_0 si

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists \eta \in \mathbb{R}_+^* \quad \forall x \in D \quad \big(d_E(x, x_0) \leqslant \eta \implies |\phi(x) - 0| \leqslant \varepsilon\big).$$

Comme $|\phi(x) - 0| = |d_F(f(x), \ell) - 0| = d_F(f(x), \ell)$, ces deux assertions sont identiques. □

Lien avec la notion de limite d'une suite

Intéressons-nous au cas particulier où $E = \overline{\mathbb{R}}$ est muni de la distance (voir p. 347)

$$d : (x, y) \in \overline{\mathbb{R}} \times \overline{\mathbb{R}} \mapsto |\phi(x) - \phi(y)| \quad \text{où} \quad \phi : x \in \overline{\mathbb{R}} \longmapsto \begin{cases} \dfrac{x}{1+|x|} & \text{si } x \in \mathbb{R} \\ 1 & \text{si } x = +\infty \\ -1 & \text{si } x = -\infty \end{cases}$$

Considérons une application f de $D = \mathbb{N}$ dans (F, d_F) ; cette application admet une limite en $x_0 = +\infty$ s'il existe un élément ℓ de F tel que

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists \eta \in \mathbb{R}_+^* \quad \forall x \in \mathbb{N} \quad \big(d(x, +\infty) \leqslant \eta \implies d_F(f(x), \ell) \leqslant \varepsilon\big). \qquad (1)$$

Pour $x \in \mathbb{N}$, on a :

$$d(x, +\infty) \leqslant \eta \iff \left|1 - \frac{x}{1+x}\right| \leqslant \eta \iff \frac{1}{1+x} \leqslant \eta \iff x \geqslant \frac{1}{\eta} - 1.$$

En posant $N = \mathrm{E}(\eta^{-1})$, l'assertion (1) est donc équivalente à l'assertion :

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad \big(n \geqslant N \implies d_F(f(n), \ell) \leqslant \varepsilon\big),$$

qui indique que la suite de terme général $u_n = f(n)$ converge vers ℓ dans (F, d_F), voir la définition 7.16 p. 388.

On a ainsi explicité le lien existant entre la limite d'une application et la limite d'une suite : la définition donnée pour la limite d'une suite $(u_n)_n$ dans un espace métrique (F, d_F) correspond à la limite de l'application $f : n \in \mathbb{N} \mapsto u_n \in F$ en $+\infty$ au sens de la définition 9.1 pour la distance d, ou pour toute distance qui lui est équivalente.

On a également le résultat suivant qui relie limite d'une application et limite d'une suite.

PROPOSITION 9.3 *Soient (E, d_E) et (F, d_F) deux espaces métriques, D un sous-ensemble de E et f une application de D dans F. Une condition nécessaire et suffisante pour que f admette pour limite l'élément ℓ de F en $x_0 \in \overline{D}$ est que pour toute suite $(u_n)_n$ d'éléments de D convergeant vers x_0 dans (E, d), la suite de terme général $f(u_n)$ converge dans (F, d_F) vers ℓ.*

Démonstration $\trianglerighteq$ Supposons que f admette pour limite l'élément ℓ de F en x_0 et considérons une suite $(u_n)_n$ convergeant vers x_0. Soit $\varepsilon \in \mathbb{R}_+^*$ fixé ; puisque f admet pour limite ℓ en x_0,

$$\exists \eta \in \mathbb{R}_+^* \quad \forall x \in D \quad \big(d_E(x, x_0) \leqslant \eta \implies d_F(f(x), \ell) \leqslant \varepsilon\big) \tag{2}$$

et puisque[6] la suite $(u_n)_n$ converge vers x_0 dan E,

$$\exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad (n \geqslant N \implies d_E(u_n, x_0) \leqslant \eta)\,.$$

Pour $n \geqslant N$ on a $d_E(u_n, x_0) \leqslant \eta$ donc $d_F(f(u_n), \ell) \leqslant \varepsilon$ d'après l'assertion (2). On a donc établi que :

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad (n \geqslant N \implies d_F(f(u_n), \ell) \leqslant \varepsilon)\,,$$

autrement dit[6] que la suite de terme général $f(u_n)$ converge vers ℓ.

$\trianglelefteq$ Pour montrer la réciproque, raisonnons par l'absurde. Supposons que pour toute suite $(u_n)_n$ convergeant vers x_0 dans (E, d_E), la suite de terme général $f(u_n)$ converge vers ℓ dans (F, d_F) et que f n'admet pas pour limite ℓ en x_0. Montrons que l'on aboutit alors à une contradiction. La négation de l'assertion « f admet pour limite ℓ en x_0 » s'écrit :

$$\exists \varepsilon \in \mathbb{R}_+^* \quad \forall \eta \in \mathbb{R}_+^* \quad \exists u_\eta \in D \quad \big(d_E(u_\eta, x_0) \leqslant \eta \text{ et } d_F(f(u_\eta), \ell) > \varepsilon\big).$$

Pour la valeur de ε donnée par l'assertion précédente, en prenant pour η des valeurs de la forme $1/n$ où $n \in \mathbb{N}^*$, on obtient :

$$\forall n \in \mathbb{N} \quad \exists u_n \in D \quad \big(d_E(u_n, x_0) \leqslant 1/n \text{ et } d_F(f(u_n), \ell) > \varepsilon\big).$$

[6] Voir la définition 7.16 p. 388.

On dispose donc d'une suite $(u_n)_n$ qui converge vers x_0 (car pour tout $n \in \mathbb{N}^*$ on a $d_E(u_n, x_0) \leqslant 1/n$) mais comme

$$\exists \varepsilon \in \mathbb{R}_+^* \quad \forall n \in \mathbb{N} \quad d_F(f(u_n), \ell) > \varepsilon,$$

la suite de terme général $f(u_n)$ ne converge pas vers ℓ. C'est en contradiction avec nos hypothèses. □

Opérations algébriques sur la limite

Le résultat suivant précise les conditions d'existence et la valeur de la limite de la composée de deux applications.

PROPOSITION 9.4 *Soient (E, d_E), (F, d_F) et (G, d_G) trois espaces métriques, f une application de E dans F et g une application de F dans G. Si f admet pour limite $y_0 \in F$ en $x_0 \in E$ et que g admet pour limite $\ell \in G$ en $y_0 \in F$ alors l'application composée $g \circ f$ admet pour limite ℓ en x_0.*

Démonstration Par hypothèse, comme f admet pour limite y_0 en x_0, on a

$$\forall \varepsilon_1 \in \mathbb{R}_+^* \quad \exists \eta_1 \in \mathbb{R}_+^* \quad \forall x \in E \quad \big(d_E(x, x_0) \leqslant \eta_1 \Longrightarrow \ d_F(f(x), y_0) \leqslant \varepsilon_1\big) \quad (3)$$

et comme g admet pour limite ℓ en y_0, on a

$$\forall \varepsilon_2 \in \mathbb{R}_+^* \quad \exists \eta_2 \in \mathbb{R}_+^* \quad \forall y \in F \quad \big(d_F(y, y_0) \leqslant \eta_2 \Longrightarrow \ d_G(g(y), \ell) \leqslant \varepsilon_2\big). \quad (4)$$

Pour montrer que $g \circ f$ admet pour limite ℓ en x_0, il faut montrer que

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists \eta \in \mathbb{R}_+^* \quad \forall x \in E \quad \big(d_E(x, x_0) \leqslant \eta \Longrightarrow \ d_G(g(f(x)), \ell) \leqslant \varepsilon\big). \quad (5)$$

Soit ε un réel strictement positif fixé, x un élément de E et $y = f(x)$ un élément de F. D'après l'assertion (4) considérée avec $\varepsilon_2 = \varepsilon$, il existe $\eta_2 \in \mathbb{R}_+^*$ tel que si $d_F(y, y_0) = d_F(f(x), y_0) \leqslant \eta_2$ alors $d_G(g(y), \ell) = d_G(g(f(x)), \ell) \leqslant \varepsilon$. D'après l'assertion (3) considérée avec $\varepsilon_1 = \eta_2$, il existe $\eta_1 \in \mathbb{R}_+^*$ tel que si $d(x, x_0) \leqslant \eta_1$ alors $d_F(f(x), y_0) \leqslant \eta_2$. En combinant ces deux propriétés, on en déduit que si $x \in E$ est tel que $d_E(x, x_0) \leqslant \eta_1$ alors $d_G(g(f(x)), \ell) \leqslant \varepsilon$. L'assertion (5) est démontrée. □

Un espace métrique n'étant pas nécessairement muni de lois de composition interne, il n'est pas possible d'énoncer de résultat sur la limite d'une somme ou d'un produit d'applications. Dans le cas d'un $\mathbb{K}$-espace vectoriel normé, on dispose toutefois d'une loi somme et d'une loi de composition externe pour lesquelles il est possible de préciser le comportement vis à vis de la limite.

Propriétés de la limite dans un espace vectoriel normé

Proposition 9.5 *Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux $\mathbb{K}$-espaces vectoriels normés, D un sous-ensemble de E et f_1 et f_2 deux applications de D dans F. Si f_1 et f_2 admettent pour limites respectives ℓ_1 et ℓ_2 en $x_0 \in \overline{D}$, alors quels que soient λ_1, λ_2 dans $\mathbb{K}$, l'application $\lambda_1 \cdot f_1 + \lambda_2 \cdot f_2$ admet une limite en $x_0 \in \overline{D}$ qui est $\lambda_1\ell_1 + \lambda_2\ell_2$.*

Démonstration Supposons que les deux applications f_1 et f_2 de D dans F admettent pour limites respectives ℓ_1 et ℓ_2 en $\mathbf{x}_0 \in \overline{D}$ et notons $f = \lambda_1 \cdot f_1 + \lambda_2 \cdot f_2$ et $\ell = \lambda_1\ell_1 + \lambda_2\ell_2$. Pour tout $\mathbf{x} \in D$, on a $f(\mathbf{x}) = \lambda_1 f_1(\mathbf{x}) + \lambda_2 f_2(\mathbf{x})$ et en exploitant les propriétés de la norme, on obtient :

$$\begin{aligned} \|f(\mathbf{x}) - \ell\|_F &= \|\lambda_1(f_1(\mathbf{x}) - \ell_1) + \lambda_2(f_2(\mathbf{x}) - \ell_2)\|_F \\ &\leqslant |\lambda_1|\,\|f_1(\mathbf{x}) - \ell_1\|_F + |\lambda_2|\,\|f_2(\mathbf{x}) - \ell_2\|_F. \end{aligned}$$

Commençons par supposer que $\lambda_1 \neq 0$ et $\lambda_2 \neq 0$ et considérons un réel strictement positif ε ; d'après la définition 9.1 on a

$$\begin{aligned} &\exists \eta_1 \in \mathbb{R}_+^* \quad \forall \mathbf{x} \in D \quad \big(\|\mathbf{x} - \mathbf{x}_0\|_E \leqslant \eta_1 \implies \|f_1(\mathbf{x}) - \ell_1\|_F \leqslant \tfrac{1}{2|\lambda_1|}\,\varepsilon\big) \\ \text{et } &\exists \eta_2 \in \mathbb{R}_+^* \quad \forall \mathbf{x} \in D \quad \big(\|\mathbf{x} - \mathbf{x}_0\|_E \leqslant \eta_2 \implies \|f_2(\mathbf{x}) - \ell_2\|_F \leqslant \tfrac{1}{2|\lambda_2|}\,\varepsilon\big). \end{aligned}$$

On en déduit que pour $\mathbf{x} \in D$ avec $d_E(\mathbf{x}, \mathbf{x}_0) \leqslant \eta$ où $\eta = \min(\eta_1, \eta_2)$ on a

$$|\lambda_1|\,\|f_1(\mathbf{x}) - \ell_1\|_F + |\lambda_2|\,\|f_2(\mathbf{x}) - \ell_2\|_F \leqslant |\lambda_1|\,\tfrac{1}{2|\lambda_1|}\,\varepsilon + |\lambda_2|\,\tfrac{1}{2|\lambda_2|}\,\varepsilon = \varepsilon.$$

On a donc démontré que

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists \eta_1 \in \mathbb{R}_+^* \quad \forall \mathbf{x} \in D \quad \big(\|\mathbf{x} - \mathbf{x}_0\|_E \leqslant \eta \implies \|f(\mathbf{x}) - \ell\|_F \leqslant \varepsilon\big),$$

autrement dit que $f = \lambda_1 \cdot f_1 + \lambda_2 \cdot f_2$ admet pour limite $\ell = \lambda_1\ell_1 + \lambda_2\ell_2$ en $\mathbf{x}_0 \in \overline{D}$. Dans le cas où $\lambda_1 = 0$ et $\lambda_2 \neq 0$ et dans le cas où $\lambda_2 = 0$ et $\lambda_1 \neq 0$, le raisonnement précédent s'adapte sans difficulté en ne considérant qu'une des deux majorations à chaque fois. Si $\lambda_1 = 0$ et $\lambda_2 = 0$ alors f est l'application nulle ; elle admet pour limite 0 en $\mathbf{x}_0$ qui est bien de la forme $\lambda_1\ell_1 + \lambda_2\ell_2$. □

Limite d'un produit dans une algèbre normée

Afin de pouvoir considérer, comme dans le cas des fonctions réelles de la variable réelle, la limite du produit de deux fonctions à valeurs dans un $\mathbb{K}$-espace vectoriel normé, il faut pouvoir munir cet espace d'arrivée d'une loi de composition interne produit $\times$. Il n'est pas toujours possible de le faire. On est amené à définir la notion d'*algèbre normée*.

DÉFINITION 9.3 (Algèbre normée)

✗ *Un ensemble structuré*[7] $(A, +, \cdot, \times)$ *est qualifié d'*algèbre *sur un corps* $\mathbb{K}$ *(ou encore de* $\mathbb{K}$*-algèbre) si*

1. $(A, +, \cdot)$ *est un* $\mathbb{K}$*-espace vectoriel,*
2. $(A, +, \times)$ *est un anneau*[8]*,*
3. *les lois* $\cdot$ *et* $\times$ *sont compatibles, i.e.*

$$\forall(\lambda, \mathbf{a}, \mathbf{b}) \in K \times A^2 \quad \lambda \cdot (\mathbf{a} \times \mathbf{b}) = (\lambda \cdot \mathbf{a}) \times \mathbf{b} = \mathbf{a} \times (\lambda \cdot \mathbf{b}).$$

✗ *On appelle* $\mathbb{K}$-algèbre normée *une algèbre sur le corps* $\mathbb{K}$ *pour laquelle le* $\mathbb{K}$*-espace vectoriel* $(A, +, \cdot)$ *est muni d'une norme* N *vérifiant :*

$$\forall(\mathbf{x}, \mathbf{y}) \in A^2 \quad N(\mathbf{x} \times \mathbf{y}) \leqslant N(\mathbf{x})\ N(\mathbf{y}).$$

Exemples

1. L'ensemble structuré $(\mathbb{R}, +, \times, \times)$ des nombres réels muni de l'addition $+$ et de la multiplication $\times$ des réels est une $\mathbb{R}$-algèbre (les deux multiplications $\cdot$ et $\times$ coïncident). Muni de la valeur absolue, il s'agit d'une algèbre normée.

2. Il n'y a pas de structure naturelle d'algèbre sur l'espace vectoriel $\mathbb{R}^3$. Autrement dit, on ne sait pas construire une multiplication entre les vecteurs de $\mathbb{R}^3$ qui satisfasse la définition 9.3.

3. L'ensemble $\mathcal{M}_n(\mathbb{R})$ des matrices carrées d'ordre $n \in \mathbb{N}^*$ à coefficients dans $\mathbb{R}$ est une $\mathbb{R}$-algèbre lorsqu'on munit cet ensemble des lois[9]
   - addition de matrices $(+)$,
   - multiplication d'une matrice par un réel $(\cdot)$,
   - produit de matrices $(\times)$.

   Si on munit $\mathcal{M}_n(\mathbb{R})$ de la norme de Schur, on dispose d'une algèbre normée puisque[10]
$$\forall(A, B) \in \mathcal{M}_n(\mathbb{K})^2 \quad \|A \times B\|_s \leqslant \|A\|_s\ \|B\|_s.$$

4. L'ensemble $\mathcal{B}(X, \mathbb{K})$, des applications bornées sur l'ensemble X et à valeurs dans $\mathbb{K}$, muni de l'addition et de la multiplication entre fonctions (les deux multiplication $\cdot$ et $\times$ coïncidant) est une $\mathbb{K}$-algèbre. Si on munit $\mathcal{B}(X, \mathbb{K})$ de la norme infinie, on dispose d'une algèbre normée puisque[11] pour tous $f, g \in \mathcal{B}(X, \mathbb{K})$, on a $\|f \times g\|_\infty \leqslant \|f\|_\infty\ \|g\|_\infty$.

---

[7] Voir la définition 2.30 p. 56 du *Cours de première année*.

[8] Voir la définition 2.35 p. 63 du *Cours de première année*.

[9] Voir les définitions 10.7 et 10.8 p. 397 du *Cours de première année*.

[10] Voir la définition 7.5 p. 365 et l'exercice 7.8 p. 366.

[11] Voir la proposition 13.7 p. 567 du *Cours de première année*.

PROPOSITION 9.6 *Soient $(E, \|\cdot\|_E)$ un $\mathbb{K}$-espace vectoriel normé, $(F, \|\cdot\|_F)$ une $\mathbb{K}$-algèbre normée, D un sous-ensemble de E et f_1 et f_2 deux applications de D dans F. Si f_1 et f_2 admettent pour limite respective ℓ_1 et ℓ_2 en $\mathbf{x}_0 \in \overline{D}$, alors l'application $\mathbf{x} \in E \mapsto f_1(\mathbf{x}) \times f_2(\mathbf{x}) \in F$ admet une limite en $\mathbf{x}_0 \in \overline{D}$ qui est $\ell_1 \times \ell_2$.*

Démonstration Pour tout $\mathbf{x} \in D$, en utilisant les propriétés de la norme d'algèbre, on obtient :

$$\begin{aligned}\|f_1(\mathbf{x}) \times f_2(\mathbf{x}) - \ell_1 \times \ell_2\|_F &= \|f_1(\mathbf{x}) \times (f_2(\mathbf{x}) - \ell_2) + \ell_2 \times (f_1(\mathbf{x}) - \ell)\|_F \\ &\leqslant \|f_1(\mathbf{x})\|_F \, \|f_2(\mathbf{x}) - \ell_2\|_F + \|\ell_2\|_F \, \|f_1(\mathbf{x}) - \ell\|_F.\end{aligned}$$

Par ailleurs, $\|f_1(\mathbf{x})\|_F = \|(f_1(\mathbf{x}) - \ell_1) + \ell_1\|_F \leqslant \|f_1(\mathbf{x}) - \ell_1\|_F + \|\ell_1\|_F$. En combinant ces deux majorations, on obtient :

$$\begin{aligned}0 \leqslant \|f_1(\mathbf{x}) \times f_2(\mathbf{x}) - \ell_1 \times \ell_2\|_F \leqslant \|f_1(\mathbf{x}) - \ell_1\|_F \, \|f_2(\mathbf{x}) - \ell_2\|_F \\ + \|\ell_2\|_F \, \|f_1(\mathbf{x}) - \ell_1\|_F + \|\ell_1\|_F \, \|f_2(\mathbf{x}) - \ell_2\|_F. \quad (6)\end{aligned}$$

Supposons que $\ell_1 \neq 0$ et que $\ell_2 \neq 0$. Pour montrer que $f_1 \times f_2$ admet pour limite $\ell_1 \times \ell_2$ en $\mathbf{x}_0$, il faut montrer que :

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists \eta \in \mathbb{R}_+^* \quad \forall \mathbf{x} \in D \qquad \big(\|\mathbf{x}-\mathbf{x}_0\|_E \leqslant \eta \Longrightarrow \|f_1(\mathbf{x})f_2(\mathbf{x}) - \ell_1\ell_2\|_F \leqslant \varepsilon\big).$$

Considérons un réel strictement positif ε ; Comme f_1 admet pour limite ℓ_1 en $\mathbf{x}_0$, on a

$$\exists \eta_1 \in \mathbb{R}_+^* \quad \forall \mathbf{x} \in D \qquad \Big(\|\mathbf{x} - \mathbf{x}_0\|_E \leqslant \eta_1 \Longrightarrow \|f_1(\mathbf{x}) - \ell_1\|_F \leqslant \frac{\varepsilon}{3\|\ell_2\|_F}\Big)$$

$$\text{et} \quad \exists \eta_2 \in \mathbb{R}_+^* \quad \forall \mathbf{x} \in D \qquad \big(\|\mathbf{x} - \mathbf{x}_0\|_E \leqslant \eta_2 \Longrightarrow \|f_1(\mathbf{x}) - \ell_1\|_F \leqslant \sqrt{\varepsilon/3}\big).$$

Et comme f_2 admet pour limite ℓ_2 en $\mathbf{x}_0$, on a

$$\exists \eta_3 \in \mathbb{R}_+^* \quad \forall \mathbf{x} \in D \qquad \Big(\|\mathbf{x} - \mathbf{x}_0\|_E \leqslant \eta_3 \Longrightarrow \|f_2(\mathbf{x}) - \ell_2\|_F \leqslant \frac{\varepsilon}{3\|\ell_1\|_F}\Big)$$

$$\text{et} \quad \exists \eta_4 \in \mathbb{R}_+^* \quad \forall \mathbf{x} \in D \qquad \big(\|\mathbf{x} - \mathbf{x}_0\|_E \leqslant \eta_4 \Longrightarrow \|f_2(\mathbf{x}) - \ell_2\|_F \leqslant \sqrt{\varepsilon/3}\big).$$

Posons $\eta = \min\{\eta_1, \eta_2, \eta_3, \eta_4\}$ et prenons $\mathbf{x} \in B(\mathbf{x}_0, \eta) \cap D$. D'après la majoration (6), on déduit des assertions précédentes que

$$\|f_1(\mathbf{x})f_2(\mathbf{x}) - \ell_1\ell_2\|_F \leqslant \sqrt{\tfrac{\varepsilon}{3}}\,\sqrt{\tfrac{\varepsilon}{3}} + \|\ell_2\|_F \frac{\varepsilon}{3\|\ell_2\|_F} + \|\ell_1\|_F \frac{\varepsilon}{3\|\ell_1\|_F} = \varepsilon.$$

On a donc établi que pour tout réel strictement positif ε, il existait un réel strictement positif η tel que pour tout réel $\mathbf{x} \in D$,

$$\|\mathbf{x} - \mathbf{x}_0\|_E \leqslant \eta \Longrightarrow \|f_1(\mathbf{x}) \times f_2(\mathbf{x}) - \ell_1 \times \ell_2\|_F \leqslant \varepsilon.$$

D'après la définition 9.1, cela signifie que la fonction $f_1 \times f_2$ admet pour limite $\ell_1 \times \ell_2$ en $\mathbf{x}_0$. Pour démontrer le résultat, dans le cas où ℓ_1 ou ℓ_2 est nul, il est nécessaire d'adapter légèrement la démonstration précédente en ne considérant que les majorations utiles. □

Applications bornées

DÉFINITION 9.4 *Soient X un ensemble et (F, d_F) un espace métrique. Une application f de X dans F est dite* bornée *si l'ensemble $f(X)$ est borné dans F, i.e. si*[12]

$$\exists y_0 \in F \quad \exists M \in \mathbb{R}^+ \quad \forall x \in X \quad d_F(f(x), y_0) \leqslant M.$$

Remarque Si F est un $\mathbb{K}$-espace vectoriel muni de la norme N_F alors une application f de X dans F est bornée si

$$\exists M \in \mathbb{R}^+ \quad \forall x \in X \quad N_F(f(x)) \leqslant M.$$

Cette assertion est obtenu en prenant dans la définition 9.4 pour y_0 le vecteur nul $\mathbf{0}_F$. ◇

PROPOSITION 9.7 *Soient (E, d_E) et (F, d_F) deux espaces métriques, D un sous-ensemble de E et f une application de D dans F. Si f admet une limite en $x_0 \in D$ alors f est bornée au voisinage de x_0.*

Démonstration Notons $\ell \in F$ la limite de f en x_0 ; d'après la définition 9.1 considérée avec $\varepsilon = 1$, on a :

$$\exists \eta \in \mathbb{R}_+^* \quad \forall x \in D \quad \big(d_E(x, x_0) \leqslant \eta \implies d_F(f(x), \ell) \leqslant 1\big).$$

En introduisant l'ensemble $V = \{x \in D \ ; \ d_E(x, x_0) \leqslant \eta\} = \overline{B}(x_0, \eta) \cap D$ qui est un voisinage de x_0, l'assertion précédente indique que pour tout $x \in V$ on a $d_F(f(x), \ell) \leqslant 1$. L'application f est donc bornée sur V. □

9.1.2 Continuité d'une fonction

Soient f une application de D dans F et x_0 un élément de D. Lorsque la limite en x_0 de f est égale à $f(x_0)$, on dit que f est continue en x_0. Plus précisément, on a la définition suivante.

DÉFINITION 9.5 *Soient (E, d_E) et (F, d_F) deux espaces métriques, D un sous-ensemble de E et f une application de D dans F.*

✗ *On dit que f est* continue en $x_0 \in D$ *si*

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists \eta \in \mathbb{R}_+^* \quad \forall x \in D \quad \big(d_E(x, x_0) \leqslant \eta \implies d_F(f(x), f(x_0)) \leqslant \varepsilon\big).$$

✗ *On dit que f est* continue sur D *si pour tout $x_0 \in D$, f est continue en x_0.*

Étant donné deux espaces métriques (E, d_E) et (F, d_F), on note $\mathcal{C}^0(D, F)$ l'ensemble des applications continues de D dans F où D est un sous-ensemble de E.

[12] Voir la définition 7.11 p. 382.

Remarques

1. D'après la définition 9.1, une application f de D dans F est continue en $x_0 \in D$ si $\lim\limits_{\substack{x \to x_0 \\ x \in D}} d_F(f(x), f(x_0)) = 0$.

2. La propriété de continuité d'une application de E dans F n'est pas modifiée si l'on change les distances sur E et sur F par des distances équivalentes [13].

3. Si f n'est pas continue en x_0, on dit que f est discontinue en x_0. Cela se traduit par :

$$\exists \varepsilon \in \mathbb{R}_+^* \quad \forall \eta \in \mathbb{R}_+^* \quad \exists x \in D \quad \big(d_E(x, x_0) \leqslant \eta \quad \text{et} \quad d_F(f(x), f(x_0)) > \varepsilon\big).$$

4. Dans le cas où $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ sont deux $\mathbb{K}$-espaces vectoriels normés, la définition 9.5 indique qu'une application f d'un sous-ensemble D de E à valeurs dans F est continue en $x_0 \in \overline{D}$ si

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists \eta \in \mathbb{R}_+^* \quad \forall x \in D \quad \big(\|x - x_0\|_E \leqslant \eta \implies \|f(x) - f(x_0)\|_F \leqslant \varepsilon\big).$$

Par ailleurs, dans le cas où les $\mathbb{K}$-espaces vectoriels sont de dimension finie, toutes les normes étant équivalentes, en pratique on choisit celle pour laquelle l'assertion précédente est la plus aisée à établir. ◇

Exemples

1. Toute application constante de E dans F est continue sur E puisque pour tous $\varepsilon \in \mathbb{R}_+^*$ et $(x, x_0) \in E^2$ on a $d_F(f(x), f(x_0)) = 0 \leqslant \varepsilon$.

2. Montrons que l'application $f : (x_1, x_2) \in \mathbb{K}^2 \mapsto x_1 + x_2 \in \mathbb{K}$ où $\mathbb{K}$ est muni de la distance usuelle et $\mathbb{K}^2$ de la distance subordonnée à la norme infinie $(x_1, x_2) \in \mathbb{K}^2 \mapsto \max(|x_1|, |x_2|)$, notée $\|\cdot\|_\infty$, est une application continue. Soit $\mathbf{y} = (y_1, y_2) \in \mathbb{K}^2$; on a pour tout $\mathbf{x} = (x_1, x_2) \in \mathbb{K}^2$

$$\begin{aligned} |f(x_1, x_2) - f(y_1, y_2)| &= |x_1 - y_1 + x_2 - y_2| \leqslant |x_1 - y_1| + |x_2 - y_2| \\ &\leqslant 2\,\|\mathbf{x} - \mathbf{y}\|_\infty. \end{aligned}$$

On en déduit que pour tout $\varepsilon \in \mathbb{R}_+^*$, il existe un réel $\eta = \frac{1}{2}\varepsilon$ tel que pour tout $\mathbf{x} \in \mathbb{K}^2$ avec $\|\mathbf{x} - \mathbf{y}\|_\infty \leqslant \eta$ on ait

$$|f(x_1, x_2) - f(y_1, y_2)| \leqslant 2\,\|\mathbf{x} - \mathbf{y}\|_\infty \leqslant 2\,\tfrac{1}{2}\varepsilon = \varepsilon.$$

L'application f est donc continue en $(y_1, y_2) \in \mathbb{K}^2$ pour tout $(y_1, y_2) \in \mathbb{K}^2$.

3. Munissons $E = \mathcal{C}^0([0,1], \mathbb{R})$ de la distance subordonnée à la norme de la convergence uniforme $f \in E \mapsto \sup_{x \in [0,1]} |f(x)|$, notée $\|\cdot\|_\infty$, et $\mathbb{R}$ de la distance usuelle. Montrons que l'application $f : \varphi \in E \mapsto \int_0^1 \cos(\varphi(x))\,dx \in \mathbb{R}$ est une application continue sur E. Il s'agit de montrer que pour tout $\varphi_0 \in E$, l'application f est continue en φ_0, autrement dit d'établir pour $\varphi_0 \in E$ fixé l'assertion suivante :

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists \eta \in \mathbb{R}_+^* \quad \forall \varphi \in E \quad \big(\|\varphi - \varphi_0\|_\infty \leqslant \eta \implies |f(\varphi) - f(\varphi_0)| \leqslant \varepsilon\big).$$

[13] Voir p. 536 pour la justification.

Pour tout $\varphi \in E$, on a

$$|f(\varphi) - f(\varphi_0)| = \left| \int_0^1 \big(\cos(\varphi(x)) - \cos(\varphi_0(x))\big)\, \mathrm{d}x \right| \leqslant \int_0^1 |\cos(\varphi(x)) - \cos(\varphi_0(x))|\, \mathrm{d}x.$$

Pour tout $x \in [0,1]$, le théorème des accroissements finis[14] appliqué à la fonction cosinus entre $\varphi(x)$ et $\varphi_0(x)$ indique qu'il existe un réel c_x tel que

$$\cos(\varphi(x)) - \cos(\varphi_0(x)) = (\varphi(x) - \varphi_0(x))\, \cos(c_x).$$

Comme la fonction cosinus est bornée par 1, on en déduit que

$$|f(\varphi) - f(\varphi_0)| \leqslant \int_0^1 |\varphi(x) - \varphi_0(x))|\, |\cos(c_x)|\, \mathrm{d}x \leqslant \int_0^1 |\varphi(x) - \varphi_0(x)|\, \mathrm{d}x \leqslant \int_0^1 \sup_{y\in[0,1]} |\varphi(y) - \varphi_0(y)|\, \mathrm{d}x = \|\varphi - \varphi_0\|_\infty.$$

Ainsi, pour tout $\varepsilon \in \mathbb{R}_+^*$, en prenant $\eta = \varepsilon$ on a pour tout $\varphi \in E$ tel que $\|\varphi - \varphi_0\|_\infty \leqslant \eta$ la majoration $|f(\varphi) - f(\varphi_0)| \leqslant \varepsilon$. L'assertion est démontrée et on peut en conclure que f est continue sur E.

4. Soit E le sous-espace vectoriel de $\mathcal{C}^1([0,1], \mathbb{R})$ formé des fonctions s'annulant en 0. On considère sur cet espace les deux normes[15]

$$N_1 : f \in E \mapsto \sup_{x\in[0,1]} |f(x)| \quad \text{et} \quad N_2 : f \in E \mapsto \sup_{x\in[0,1]} |f'(x)|.$$

Montrons que l'application identité de E muni de la norme N_2, dans E muni de la norme N_1 est continue. Il s'agit de montrer que pour tout $f_0 \in E$, l'identité est continue en f_0, c'est-à-dire d'établir pour $f_0 \in E$ fixé l'assertion suivante :

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists \eta \in \mathbb{R}_+^* \quad \forall f \in E \quad \big(N_2(f - f_0) \leqslant \eta \implies N_1(f - f_0) \leqslant \varepsilon\big).$$

Pour tout $x \in [0,1]$, on a

$$f(x) = f(x) - f(0) = \int_0^x f'(t)\, \mathrm{d}t \quad \text{et} \quad f_0(x) = f_0(x) - f_0(0) = \int_0^x f_0'(t)\, \mathrm{d}t.$$

On en déduit que

$$|f(x) - f_0(x)| = \left| \int_0^x \big(f'(t) - f_0'(t)\big)\, \mathrm{d}t \right| \leqslant \int_0^x \big|f'(t) - f_0'(t)\big|\, \mathrm{d}t \leqslant \int_0^x \sup_{s\in[0,1]} |f'(s) - f_0'(s)|\, \mathrm{d}t = x \sup_{s\in[0,1]} |f'(s) - f_0'(s)| \leqslant N_2(f - f_0).$$

---

[14] Voir le théorème 16.2 p.731 du *Cours de première année*.

[15] On vérifiera que N_2 est bien une norme sur E. On pourra consulter l'exercice 7.15 p. 409 à ce propos.

On a donc $N_1(f - f_0) = \sup_{x\in[0,1]} |f(x) - f_0(x)| \leqslant N_2(f - f_0)$. Ainsi, pour tout $\varepsilon \in \mathbb{R}_+^*$, en prenant $\eta = \varepsilon$ on a pour tout $f \in E$ tel que $N_2(f - f_0) \leqslant \eta$ la majoration $N_1(f - f_0) \leqslant \varepsilon$. L'assertion est démontrée et on peut en conclure que l'identité de (E, N_2) dans (E, N_1) est continue.

Nous allons à présent montrer que l'application identité de (E, N_1) dans (E, N_2) n'est pas continue[16] en 0_E où 0_E désigne ici l'application nulle. Pour montrer la non-continuité de l'identité en 0_E, il faut montrer que

$$\exists \varepsilon \in \mathbb{R}_+^* \quad \forall \eta \in \mathbb{R}_+^* \quad \exists f \in E \quad \big(N_1(f) \leqslant \eta \quad \text{et} \quad N_2(f) > \varepsilon\big).$$

Pour $\varepsilon = \frac{1}{2}$ et $\eta \in \mathbb{R}_+^*$, considérons l'application $f : x \in [0, 1] \mapsto \frac{1}{n} x^n$ où $n = 1 + \mathrm{E}(\eta^{-1})$. Cette application est bien de classe $\mathcal{C}^1$ et prend la valeur 0 en 0. Elle vérifie $N_1(f) = \frac{1}{n} \leqslant \eta$ et $N_2(f) = 1 > \varepsilon$. L'assertion est établie. L'application identité de (E, N_1) dans (E, N_2) n'est donc pas continue en 0.

Comme par définition une application f est continue en x_0 si elle admet pour limite $f(x_0)$ en x_0, on déduit des propriétés de la limite énoncées dans la partie précédente un certain nombre de résultats concernant les applications continues. En particulier, d'après la proposition 9.7, une application continue sur un ensemble D est bornée sur cet ensemble. On a également les propositions suivantes qui sont des corollaires des propositions 9.3 et 9.4 respectivement.

Proposition 9.8 *Soient (E, d_E) et (F, d_F) deux espaces métriques, D un sous-ensemble de E et f une application de D dans F. Une condition nécessaire et suffisante pour que f soit continue en $x_0 \in D$ est que pour toute suite $(u_n)_n$ d'éléments de D convergeant vers x_0, la suite de terme général $f(u_n)$ converge dans (F, d_F) vers $f(x_0)$.*

Proposition 9.9 *Soient (E, d_E), (F, d_F) et (G, d_G) trois espaces métriques, f une application de E dans F et g une application de F dans G. Si f est continue en $x_0 \in E$ (resp. sur E) et que g est continue en $y_0 = f(x_0)$ (resp. sur F) alors l'application composée $g \circ f$ est continue en x_0 (resp. sur E).*

Par ailleurs, dans le cas d'espaces vectoriels normés, on dispose de la propriété suivante qui est un corollaire de la proposition 9.5.

Proposition 9.10 *Soient (E, N_E) et (F, N_F) deux $\mathbb{K}$-espaces vectoriels normés, D un sous-ensemble de E et f_1 et f_2 deux applications de D dans F. Si f_1 et f_2 sont continues en $x_0 \in D$ (resp. sur D) alors quels que soient λ_1, λ_2 dans $\mathbb{K}$ l'application $\lambda_1 \cdot f_1 + \lambda_2 \cdot f_2$ est continue en $x_0 \in D$ (resp. sur D).*

[16] Comme il s'agit d'une application linéaire, le théorème 9.3 p. 574 établira que n'étant pas continue en 0, cette application n'est continue en aucun point de E.

Afin de pouvoir considérer du point de vue de la continuité le produit de deux fonctions à valeurs dans un espace vectoriel normé, il faut pouvoir munir l'espace d'arrivée d'une loi de composition interne produit $\times$, autrement dit considérer des algèbres normées. On a la proposition suivante qui est un corollaire de la proposition 9.6.

PROPOSITION 9.11 *Soient $(E, \|\cdot\|_E)$ un $\mathbb{K}$-espace vectoriel normé, $(F, \|\cdot\|_F)$ une $\mathbb{K}$-algèbre normée, D un sous-ensemble de E et f_1 et f_2 deux applications de D dans F. Si f_1 et f_2 sont continues en $\mathbf{x}_0 \in D$ (resp. sur D) alors l'application $f_1 \times f_2$ définie par $\mathbf{x} \in E \mapsto f_1(\mathbf{x}) \times f_2(\mathbf{x}) \in F$ est continue en $\mathbf{x}_0 \in D$ (resp. sur D).*

Exemple L'espace vectoriel $(\mathcal{M}_n(\mathbb{K}), +, \cdot)$ muni de la norme de Schur est une algèbre normée pour le produit entre matrices. L'application identité $M \in \mathcal{M}_n(\mathbb{K}) \mapsto M$ étant continue sur $\mathcal{M}_n(\mathbb{K})$, on déduit de la proposition 9.11 en utilisant un raisonnement par récurrence que pour tout $k \in \mathbb{N}^*$, l'application $M \in \mathcal{M}_n(\mathbb{K}) \mapsto M^k$ est continue sur $\mathcal{M}_n(\mathbb{K})$. La proposition 9.10 permet alors d'affirmer que pour tout polynôme $P \in \mathbb{K}[X]$ l'application

$$M \in \mathcal{M}_n(\mathbb{K}) \mapsto P(M) \in \mathcal{M}_n(\mathbb{K})$$

est continue sur $\mathcal{M}_n(\mathbb{K})$.

Continuité de l'application déterminant

L'objet de cette partie est d'établir que l'application déterminant[17] d'ordre n par rapport à une base $\mathcal{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ de $\mathbb{R}^n$, qui est l'unique forme n-linéaire alternée[18] vérifiant $\det_{\mathcal{B}}(\mathbf{e}_1, \ldots, \mathbf{e}_n) = 1$, est une application continue de $\mathbb{R}^n \times \cdots \times \mathbb{R}^n$ dans $\mathbb{R}$. Rappelons que si $\mathbf{x}_1, \ldots, \mathbf{x}_n$ sont des vecteurs de $\mathbb{R}^n$ alors

$$\det_{\mathcal{B}}(\mathbf{x}_1, \ldots, \mathbf{x}_n) = \sum_{\sigma \in \mathfrak{S}_n} \left(\varepsilon(\sigma) \prod_{i=1}^{n} x_{ij} \right)$$

où $\mathfrak{S}_n$ désigne l'ensemble des permutations de $\{1, \ldots, n\}$, $\varepsilon(\sigma)$ la signature de la permutation σ et $(x_{1j}, \ldots, x_{nj})$ les composantes du vecteur $\mathbf{x}_j$ dans la base $\mathcal{B}$. Pour établir la continuité de l'application déterminant, nous allons avoir recours à l'inégalité de Hadamard.

[17] Voir la définition 11.6 p. 474 du *Cours de première année*.

[18] Voir la définition 11.5 p. 473 du *Cours de première année*.

LEMME 9.1 (Inégalité de Hadamard)
Le déterminant d'ordre n par rapport à une base $\mathcal{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ *de* $\mathbb{R}^n$ *des vecteurs* $\mathbf{x}_1, \ldots, \mathbf{x}_n$ *vérifie*

$$|\det_{\mathcal{B}}(\mathbf{x}_1, \ldots, \mathbf{x}_n)| \leqslant \prod_{j=1}^{n} \|\mathbf{x}_j\|_2$$

où $\|\mathbf{x}_j\|_2$ *désigne la norme euclidienne du vecteur* $\mathbf{x}_j$.

Démonstration Procédons par disjonction de cas : si les vecteurs de la famille $(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ sont liés alors[19] $\det_{\mathcal{B}}(\mathbf{x}_1, \ldots, \mathbf{x}_n) = 0$; la norme d'un vecteur étant un réel positif, l'inégalité est vraie. Intéressons-nous au cas où la famille $(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ est libre. S'agissant d'une famille libre de n vecteurs, elle forme[20] une base de $\mathbb{R}^n$. Nous savons par le théorème 8.1 que dans l'espace euclidien $\mathbb{R}^n$ muni du produit scalaire canonique $\langle\,,\,\rangle$ il existe une base $\mathcal{B}_o = (\mathbf{u}_1, \ldots, \mathbf{u}_n)$ de $\mathbb{R}^n$ qui est orthonormale. Cette base peut être construite par le procédé d'orthogonalisation de Schmidt ; la matrice de passage $\mathrm{P} = (p_{ij})_{i,j} \in \mathcal{M}_n(\mathbb{R})$ de la base orthogonale $\mathcal{B}_o$ à la base $\mathcal{B}_{\mathbf{x}} = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$ est une matrice triangulaire. On a[21] $\det(\mathrm{P}) = \prod_{i=1}^{n} p_{ii}$ et compte tenu de l'inégalité de Cauchy-Schwarz[22],

$$|p_{ii}| = |\langle \mathbf{x}_i, \mathbf{u}_i \rangle| \leqslant \|\mathbf{x}_i\|_2 \, \|\mathbf{u}_i\|_2 = \|\mathbf{x}_i\|_2.$$

Notons A la matrice représentant l'identité relativement aux bases $\mathcal{B}_{\mathbf{x}}$ et $\mathcal{B}$, autrement dit[23] la matrice de passage de $\mathcal{B}$ à $\mathcal{B}_{\mathbf{x}}$, et par B la matrice représentant l'identité relativement aux bases $\mathcal{B}_o$ et $\mathcal{B}$, autrement dit la matrice de passage de $\mathcal{B}$ à $\mathcal{B}_o$. La matrice B est orthogonale[24] ; son déterminant vaut donc[25] 1 ou -1. La formule de changement de base[26]

$$\mathrm{Mat}_{\mathcal{B}_{\mathbf{x}},\mathcal{B}}(\mathrm{id}) = \mathrm{Mat}_{\mathcal{B}_o,\mathcal{B}}(\mathrm{id}) \times \mathrm{Mat}_{\mathcal{B}_o,\mathcal{B}_o}(\mathrm{id}) \times \mathrm{Mat}_{\mathcal{B}_{\mathbf{x}},\mathcal{B}_o}(\mathrm{id})$$

indique que $\mathrm{A} = \mathrm{B} \times \mathrm{P}$. On en déduit que

$$|\det_{\mathcal{B}}(\mathbf{x}_1, \ldots, \mathbf{x}_n)| = |\det(\mathrm{A})| = |\det(\mathrm{B})| \times |\det(\mathrm{P})| = \prod_{i=1}^{n} |p_{ii}| \leqslant \prod_{i=1}^{n} \|\mathbf{x}_i\|_2.$$

L'inégalité d'Hadamard est démontrée. □

---

[19] Voir la proposition 11.1 p. 476 du *Cours de première année*.
[20] Voir la proposition 8.11 p. 330 du *Cours de première année*.
[21] Voir la proposition 11.5 p. 482 du *Cours de première année*.
[22] Voir le théorème 8.5 p.481.
[23] Voir la proposition 10.15 p. 435 du *Cours de première année*.
[24] Voir la proposition 8.25 p. 494.
[25] Voir la proposition 8.27 p. 497.
[26] Voir le théorème 10.1 p. 440 du *Cours de première année*.

Munissons $\mathbb{R}^n \times \cdots \times \mathbb{R}^n$ de la distance produit associée à la norme euclidienne sur $\mathbb{R}^n$, voir p. 364, et $\mathbb{R}$ de la distance usuelle. Comme le déterminant du n-uplet $(\mathbf{0}, \ldots, \mathbf{0})$ est 0, l'inégalité d'Hadamard, combinée au théorème d'encadrement, indique que l'application déterminant est continue en $(\mathbf{0}, \ldots, \mathbf{0})$. Plus généralement, on montre en utilisant l'inégalité d'Hadamard et le fait que l'application déterminant d'ordre n est une forme n-linéaire que l'application déterminant d'ordre n est une application continue sur $\mathbb{R}^n \times \cdots \times \mathbb{R}^n$.

PROPOSITION 9.12 *L'application déterminant d'ordre n est une application continue de* $\mathbb{R}^n \times \cdots \times \mathbb{R}^n$ *dans* $\mathbb{R}$.

Démonstration Pour montrer que l'application déterminant d'ordre n par rapport à une base $\mathcal{B}$ de $\mathbb{R}^n$

$$\det : (\mathbf{x}_1, \ldots, \mathbf{x}_n) \in \mathbb{R}^n \times \cdots \times \mathbb{R}^n \longmapsto \sum_{\sigma \in \mathfrak{S}_n} \left(\varepsilon(\sigma) \prod_{i=1}^{n} x_{ij} \right),$$

où $(x_{1j}, \ldots, x_{nj})$ désignent les composantes du vecteur $\mathbf{x}_j$ dans la base $\mathcal{B}$, est une application continue de $E = \mathbb{R}^n \times \cdots \times \mathbb{R}^n$ dans $\mathbb{R}$, il faut montrer que pour tout $(\mathbf{a}_1, \ldots, \mathbf{a}_n) \in E$, l'application $\det_{\mathcal{B}}$ est continue en $(\mathbf{a}_1, \ldots, \mathbf{a}_n)$, *i.e.* que la quantité $\Delta(\mathbf{x}_1, \ldots, \mathbf{x}_n) = \det(\mathbf{x}_1, \ldots, \mathbf{x}_n) - \det(\mathbf{a}_1, \ldots, \mathbf{a}_n)$ tend vers 0 quand $(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ tend vers $(\mathbf{a}_1, \ldots, \mathbf{a}_n)$. Soit $(\mathbf{a}_1, \ldots, \mathbf{a}_n) \in E$; pour tout $(\mathbf{x}_1, \ldots, \mathbf{x}_n) \in E$ on a :

$$\begin{aligned}\Delta(\mathbf{x}_1, \ldots, \mathbf{x}_n) &= \det(\mathbf{x}_1, \ldots, \mathbf{x}_n) - \det(\mathbf{a}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n) + \det(\mathbf{a}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n) \\ &\quad - \det(\mathbf{a}_1, \mathbf{a}_2, \mathbf{x}_3, \ldots, \mathbf{x}_n) + \det(\mathbf{a}_1, \mathbf{a}_2, \mathbf{x}_3, \ldots, \mathbf{x}_n) - \det(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{x}_4, \ldots, \mathbf{x}_n) \\ &\quad + \cdots + \det(\mathbf{a}_1, \ldots, \mathbf{a}_{n-1}, \mathbf{x_n}) - \det(\mathbf{a}_1, \ldots, \mathbf{a}_n).\end{aligned}$$

En utilisant le fait que le déterminant est une forme n-linéaire alternée, il vient

$$\begin{aligned}\Delta(\mathbf{x}_1, \ldots, \mathbf{x}_n) &= \det(\mathbf{x}_1 - \mathbf{a_1}, \mathbf{x_2}, \ldots, \mathbf{x}_n) + \det(\mathbf{a}_1, \mathbf{x}_2 - \mathbf{a_2}, \mathbf{x}_3, \ldots, \mathbf{x}_n) \\ &\quad + \cdots + \det(\mathbf{a}_1, \ldots, \mathbf{a}_{n-2}, \mathbf{x}_{n-1} - \mathbf{a}_{n-1}, \mathbf{x}_n) + \det(\mathbf{a}_1, \ldots, \mathbf{a}_{n-1}, \mathbf{x}_n - \mathbf{a}_n).\end{aligned}$$

En utilisant l'inégalité triangulaire puis l'inégalité d'Hadamard, on obtient :

$$\begin{aligned}|\Delta(\mathbf{x}_1, \ldots, \mathbf{x}_n)| &= \left| \sum_{i=1}^{n} \det(\mathbf{a}_1, \ldots, \mathbf{a}_{i-1}, \mathbf{x}_i - \mathbf{a}_i, \mathbf{x}_{i+1}, \ldots, \mathbf{x}_n) \right| \\ &\leqslant \sum_{i=1}^{n} \left(\|\mathbf{x}_i - \mathbf{a}_i\|_2 \prod_{j=1}^{i-1} \|\mathbf{a_j}\|_2 \prod_{j=i+1}^{n} \|\mathbf{x_j}\|_2 \right).\end{aligned}$$

Le théorème d'encadrement permet de conclure que la limite quand $(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ tend vers $(\mathbf{a}_1, \ldots, \mathbf{a}_n)$ de $\det(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ vaut $\det(\mathbf{a}_1, \ldots, \mathbf{a}_n)$, ce qui signifie que l'application det est continue en $(\mathbf{a}_1, \ldots, \mathbf{a}_n)$. La proposition est démontrée. □

Remarques

1. Plus généralement, on montre que toute forme n-linéaire[27] sur $E = E_1 \times \cdots \times E_n$ continue en $\mathbf{0}_E$ est continue sur E.

2. Le déterminant d'une matrice $A \in \mathcal{M}_n(\mathbb{R})$ étant défini comme le déterminant dans la base canonique de $\mathbb{R}^n$ de ses n vecteurs colonnes, la proposition 9.12 indique que l'application qui à une matrice associe son déterminant est continue sur $\mathcal{M}_n(\mathbb{R})$.

3. La proposition 9.12 a été démontrée pour le déterminant d'ordre n par rapport à une base $\mathcal{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ de $\mathbb{R}^n$ compte tenu du fait que l'inégalité d'Hadamard a été établie pour des vecteurs de $\mathbb{R}^n$ (on utilise des résultats concernant les espaces euclidiens). La propriété de continuité de l'application déterminant est toutefois encore vraie pour des vecteurs de $\mathbb{C}^n$ et peut être établie par une approche analogue à celle présentée ici en exploitant les propriétés des espaces dits « hermitiens ». ◇

Applications définies sur un produit d'espaces métriques

Soient (F, d_F) et $(E_1, d_1), \ldots, (E_p, d_p)$ des espaces métriques ($p \in \mathbb{N}^*$). On munit l'espace produit $E = E_1 \times \cdots \times E_p$ d'une des distances d_E définies par :

$$(x, y) \in E^2 \longmapsto \max(d_1(x_1, y_1), \ldots, d_p(x_p, y_p)), \tag{7}$$

$$(x, y) \in E^2 \longmapsto d_1(x_1, y_1) + \cdots + d_p(x_p, y_p), \tag{8}$$

$$(x, y) \in E^2 \longmapsto \sqrt{d_1(x_1, y_1)^2 + \cdots + d_p(x_p, y_p)^2}, \tag{9}$$

où $x = (x_1, \ldots, x_p)$ et $y = (y_1, \ldots, y_p)$. Ces distances sont équivalentes, voir p. 356.

PROPOSITION 9.13 *Pour tout $k \in \{1, \ldots, p\}$, la k-ème fonction de projection $p_k : (x_1, \ldots, x_p) \in E \mapsto x_k \in E_k$ est une application continue sur E.*

Démonstration Considérons sur l'espace produit $E = E_1 \times \cdots \times E_p$ la distance δ définie par la relation (7). Pour montrer que l'application p_k est continue sur E pour tout $k \in \{1, \ldots, p\}$, il faut montrer pour tout $y \in E$ l'assertion suivante :

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists \eta \in \mathbb{R}_+^* \quad \forall x \in E \quad \big(\delta(x, y) \leqslant \eta \implies d_k(p_k(x), p_k(y)) \leqslant \varepsilon\big).$$

Soit $y = (y_1, \ldots, y_p)$ un élément quelconque de E ; pour tout $x = (x_1, \ldots, x_p) \in E$ on a

$$d_k(p_k(x), p_k(y)) = d_k(x_k, y_k) \leqslant \max(d_1(x_1, y_1), \ldots, d_p(x_p, y_p)) = \delta(x, y).$$

[27] Voir la définition 11.5 p. 473 du *Cours de première année*.

On en déduit que pour tout réel $\varepsilon \in \mathbb{R}_+^*$, il existe un réel $\eta \in \mathbb{R}_+^*$ (il suffit de prendre $\eta = \varepsilon$) tel que pour tout $x \in E$ si $\delta(x,y) \leqslant \eta$ alors $d_k(p_k(x), p_k(y)) \leqslant \varepsilon$. L'application p_k est continue en y pour tout $y \in E$. On en conclut que pour tout $k \in \{1, \ldots, p\}$, l'application p_k est continue sur E. □

DÉFINITION 9.6 (Applications partielles)
Soient D un sous-ensemble de l'espace produit $E_1 \times \cdots \times E_p$, f une application de D dans F et $a = (a_1, \ldots, a_p)$ un élément de U. Pour tout entier i avec $1 \leqslant i \leqslant p$, on appelle i–ième application partielle de f en a *l'application*

$$x_i \in D_i \longmapsto f(a_1, \ldots, a_{i-1}, x_i, a_{i+1}, \ldots, a_p) \in F$$

définie sur $D_i = \{x_i \in E_i \ ; \ (a_1, \ldots, a_{i-1}, x_i, a_{i+1}, \ldots, a_n) \in D\}$.

Exemples

1. La i–ième application partielle de $f : (x_1, \ldots, x_n) \in \mathbb{R}^n \mapsto x_1 + \cdots + x_n$ en $a = (0, \ldots, 0)$ est $\phi_{a,i} : x_i \in \mathbb{R} \mapsto x_i$. La i–ième application partielle de f en $a = (1, \ldots, 1)$ est $\phi_{a,i} : x_i \in \mathbb{R} \mapsto (n-1) + x_i$.

2. L'application $f : (x_1, x_2) \in \mathbb{R}^2 \longmapsto \begin{cases} \dfrac{x_1 x_2}{x_1^2 + x_2^2} & \text{si } (x_1, x_2) \neq (0,0) \\ 1 & \text{si } (x_1, x_2) = (0,0) \end{cases}$
admet pour applications partielles en $a = (0,0)$ les applications

$$\phi_{a,1} : x_1 \in \mathbb{R} \longmapsto \begin{cases} 0 & \text{si } x_1 \neq 0 \\ 1 & \text{si } x_1 = 0 \end{cases} \quad \text{et} \quad \phi_{a,2} : x_2 \in \mathbb{R} \longmapsto \begin{cases} 0 & \text{si } x_2 \neq 0 \\ 1 & \text{si } x_2 = 0 \end{cases}.$$

3. Notons U la matrice de $\mathcal{M}_n(\mathbb{K})$ dont tous les coefficients valent 1. L'application f : $(A, B) \in \mathcal{M}_n(\mathbb{K})^2 \mapsto A \times B - B \times A$, appelée commutateur de A et B, admet pour première application partielle en (U, I) où $I \in \mathcal{M}_n(\mathbb{K})$ désigne la matrice identité, l'application $A \in \mathcal{M}_n(\mathbb{K}) \mapsto 0$ où $0 \in \mathcal{M}_n(\mathbb{K})$ désigne l'application nulle. La seconde application partielle en (U, I) est l'application $B \in \mathcal{M}_n(\mathbb{K}) \mapsto \left(\sum_{k=1}^n (b_{kj} - b_{ik})\right)_{ij}$.

PROPOSITION 9.14 *Soient (E, d_E) dans (F, d_F) deux espaces métriques, f une application de E dans F et $a = (a_1, \ldots, a_p)$ un élément de E. Si f est continue en a alors pour tout entier i avec $1 \leqslant i \leqslant p$, la i–ième application partielle de f en a est continue en a_i.*

Démonstration Supposons que f soit continue en a ; on a

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists \eta \in \mathbb{R}_+^* \quad \forall x \in E \quad \big(d_E(x,a) \leqslant \eta \implies d_F(f(x), f(a)) \leqslant \varepsilon\big). \quad (10)$$

Considérons un entier i avec $1 \leqslant i \leqslant p$. L'assertion précédente étant vraie pour tout élément $x = (x_1, \ldots, x_p)$ de E, elle est vraie pour les éléments de E de la forme $\tilde{x} = (a_1, \ldots, a_{i-1}, x_i, a_{i+1}, \ldots, a_p)$ où x_i désigne un élément de E_i. Désignons par $\phi_{a,i}$ la i–ième application partielle de f en a. La distance d_E étant l'une des distances définies par les relations (7) à (9), on a d'une part $d_E(\tilde{x}, a) = d_i(x_i, a_i)$ et d'après la définition de la i–ième application partielle de f en a, on a d'autre part :

$$f(\tilde{x}) = f(a_1, \ldots, a_{i-1}, x_i, a_{i+1}, \ldots, a_p) = \phi_{a,i}(x_i)$$

avec $f(a) = \phi_{a,i}(a_i)$. L'assertion (10) implique donc que :

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists \eta \in \mathbb{R}_+^* \quad \forall x_i \in E_i \quad \big(d_i(x_i, a_i) \leqslant \eta \implies d_F(\phi_{a,i}(x_i), \phi_{a,i}(a_i)) \leqslant \varepsilon\big).$$

D'après la définition 9.5, cela signifie que $\phi_{a,i}$ est continue en a_i. □

La proposition 9.14 n'admet pas de réciproque : toutes les applications partielles de f en a peuvent être continues sans que f soit continue en a. Par exemple, l'application

$$f : (x_1, x_2) \in \mathbb{R}^2 \longmapsto \begin{cases} \dfrac{x_1 x_2}{x_1^2 + x_2^2} & \text{si } (x_1, x_2) \neq (0,0) \\ 0 & \text{si } (x_1, x_2) = (0,0) \end{cases}$$

n'est pas continue en $(0,0)$ (voir les exemples de la p. 537) alors que les deux applications partielles en $(0,0)$ le sont (les deux applications partielles $\phi_{(0,0),1}$ et $\phi_{(0,0),2}$ sont toutes deux égales à l'application nulle).

D'un point de vue pratique, la proposition 9.14 est souvent utilisée sous sa forme contraposée pour montrer qu'une application f définie sur un espace produit $E = E_1 \times \ldots \times E_p$ n'est pas continue en $a = (a_1, \ldots, a_p) \in E$: si l'une des applications partielles de f en a, disons la i-ème, n'est pas continue en a_i alors f n'est pas continue en a. Par exemple, l'application :

$$f : (x_1, x_2) \in \mathbb{R}^2 \longmapsto \begin{cases} \dfrac{x_1 + x_2}{x_1^2 + x_2^2} & \text{si } (x_1, x_2) \neq (0,0) \\ 0 & \text{si } (x_1, x_2) = (0,0) \end{cases}$$

n'est pas continue en $(0,0)$ car la première application partielle en 0 qui est l'application $\phi_{(0,0),1}$ définie par $\phi_{(0,0),1}(x_1) = 1/x_1$ si $x_1 \neq 0$ et $\phi_{(0,0),1}(0) = 0$ n'est pas continue en $x_1 = 0$.

Applications à valeurs dans un produit d'espaces métriques

Soient (E, d_E), $(F_1, d_1), \ldots, (F_q, d_q)$ des espaces métriques et $F = F_1 \times \ldots \times F_q$ l'espace produit muni de l'une des distances d_F définies par :

$$(x, y) \in F^2 \longmapsto \max(d_1(x_1, y_1), \ldots, d_q(x_q, y_q)), \tag{11}$$

$$(x, y) \in F^2 \longmapsto d_1(x_1, y_1) + \cdots + d_q(x_q, y_q), \tag{12}$$

$$(x, y) \in F^2 \longmapsto \sqrt{d_1(x_1, y_1)^2 + \cdots + d_q(x_q, y_q)^2}, \tag{13}$$

où $x = (x_1, \ldots, x_q)$ et $y = (y_1, \ldots, y_q)$. Ces distances sont équivalentes.

PROPOSITION 9.15 *L'application $f : x \in E \mapsto (f_1(x), \ldots, f_q(x)) \in F$ est continue en $a \in E$ si et seulement si pour tout entier i avec $1 \leqslant i \leqslant p$, l'application f_i de E dans F_i est continue en a.*

Démonstration Considérons sur l'espace produit $F = F_1 \times \ldots \times F_q$ la distance d_F définie par la relation (11). L'application f est continue en a si

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists \eta \in \mathbb{R}_+^* \quad \forall x \in E \quad \big(d_E(x,a) \leqslant \eta \implies d_F(f(x), f(a)) \leqslant \varepsilon\big).$$

Puisque $d_F(f(x), f(a)) = \max(d_1(f_1(x), f_1(a)), \ldots, d_q(f_q(x), f_q(a)))$, la condition $d_F(f(x), f(a)) \leqslant \varepsilon$ implique que $d_i(f_i(x), f_i(a)) \leqslant \varepsilon$ pour tout entier i avec $1 \leqslant i \leqslant n$. On en déduit pour tout $i \in \{1, \ldots, n\}$ l'assertion

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists \eta \in \mathbb{R}_+^* \quad \forall x \in E \quad \big(d_E(x,a) \leqslant \eta \implies d_i(f_i(x), f_i(a)) \leqslant \varepsilon\big).$$

D'après la définition 9.5, on en déduit que f_i est continue en a. On établit la réciproque en remarquant que si $d_i(f_i(x), f_i(a)) \leqslant \varepsilon$ pour tout $i \in \{1, \ldots, n\}$, alors

$$d_F(f(x), f(a)) = \max(d_1(f_1(x), f_1(a)), \ldots, d_q(f_q(x), f_q(a))) \leqslant \varepsilon.$$

la proposition est démontrée. □

9.1.3 Continuité uniforme

Étant donné deux espaces métriques (E, d_E) et (F, d_F), une application f de E dans F est continue sur un sous-ensemble D de E si :

$$\forall x \in D \quad \forall \varepsilon \in \mathbb{R}_+^* \quad \exists \eta_{x,\varepsilon} \in \mathbb{R}_+^* \quad \forall y \in D \quad \big(d_E(x,y) \leqslant \eta_{x,\varepsilon} \implies d_F(f(x), f(y)) \leqslant \varepsilon\big).$$

Pour certaines applications, le réel $\eta_{x,\varepsilon}$ peut être identique pour tous les éléments x de D : il ne dépend alors pas de x. En choisissant d'imposer cette condition à une application, on définit alors une notion plus « forte » que la continuité appelée continuité uniforme.

DÉFINITION 9.7 *Soient (E, d_E) et (F, d_F) deux espaces métriques, D un sous-ensemble de E et f une application de D dans F. On dit que f est* uniformément continue sur D *si*

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists \eta \in \mathbb{R}_+^* \quad \forall (x,y) \in D^2 \quad \big(d_E(x,y) \leqslant \eta \implies d_F(f(x), f(y)) \leqslant \varepsilon\big).$$

Remarques

1. Pour une application de E dans F, la propriété d'être uniformément continue n'est pas modifiée si l'on change les distances sur E et sur F pour des distances équivalentes.

2. Il résulte des définitions 9.5 et 9.7 que toute application uniformément continue sur D est continue sur D. La réciproque n'est bien évidemment pas vraie. ◇

Exemples

1. Toute application f de (E, d_E) dans (F, d_F) constante sur $D \subset E$ est uniformément continue sur D puisqu'on a $d_F(f(x), f(y)) = 0$ pour tout $(x, y) \in D^2$.

2. Soit (E, d) un espace métrique; munissons E^2 de la distance δ_1 définie par $\delta_1(x, y) = d(x_1, y_1) + d(x_2, y_2)$ pour tous $x = (x_1, x_2)$ et $y = (y_1, y_2)$ dans E^2 et munissons $\mathbb{R}$ de la distance usuelle. Montrons que l'application d de E^2 dans $\mathbb{R}$ est uniformément continue sur E^2. En utilisant successivement les deux inégalités triangulaires[28], on établit que pour tous $x = (x_1, x_2) \in E^2$ et $y = (y_1, y_2) \in E^2$:

$$\begin{aligned} |d(x_1, x_2) - d(y_1, y_2)| &= |d(x_1, x_2) - d(y_1, x_2) + d(y_1, x_2) - d(y_1, y_2)| \\ &\leqslant |d(x_1, x_2) - d(y_1, x_2)| + |d(y_1, x_2) - d(y_1, y_2)| \\ &\leqslant d(x_1, y_1) + d(x_2, y_2) = \delta_1(x, y). \end{aligned}$$

On en déduit que pour tout $\varepsilon \in \mathbb{R}_+^*$, il existe $\eta \in \mathbb{R}_+^*$ (prendre $\eta = \varepsilon$) tel que pour tout $(x, y) \in E^2 \times E^2$, si $\delta_1(x, y) \leqslant \eta$ alors $|d(y_1, y_2) - d(x_1, x_2)| \leqslant \eta = \varepsilon$.

On a ainsi montré qu'une distance sur un ensemble E est une application uniformément continue de E^2 dans $\mathbb{R}^+$.

3. L'application $f : x \in \mathbb{R} \mapsto \frac{1}{1+|x|} \in \mathbb{R}$ est uniformément continue sur $\mathbb{R}$. En effet, puisque pour tout $(x, y) \in \mathbb{R}^2$ on a $|x| + 1 \geqslant 1$ et $|y| + 1 \geqslant 1$, on obtient :

$$|f(x) - f(y)| = \left| \frac{|y| - |x|}{(1 + |x|)(1 + |y|)} \right| \leqslant \big||y| - |x|\big| \leqslant |y - x|.$$

On en déduit que pour tout $\varepsilon \in \mathbb{R}_+^*$, il existe $\eta \in \mathbb{R}_+^*$ (prendre $\eta = \varepsilon$) tel que pour tout $(x, y) \in \mathbb{R}^2$, si $|x - y| \leqslant \eta$ alors $|f(y) - f(x)| \leqslant |y - x| = \eta = \varepsilon$.

4. Soit $(E, \|\cdot\|_E)$ un $\mathbb{K}$-espace vectoriel normé et $\mathbf{v}_1, \ldots, \mathbf{v}_n$ des vecteurs de E non tous nuls. Montrons que l'application

$$f : (x_1, \ldots, x_n) \in \mathbb{K}^n \mapsto \sum_{i=1}^{n} x_i \cdot \mathbf{v}_i \in E$$

[28] Voir la définition 7.1 p. 345 et la proposition 7.1 p. 351.

est uniformément continue sur $\mathbb{K}^n$ muni de l'une quelconque de ses normes. L'espace vectoriel $\mathbb{K}^n$ étant de dimension finie, toutes les normes y sont équivalentes ; munissons-le de la norme $\|\cdot\|_1$ définie par $\|\mathbf{x}\|_1 = \sum_{i=1}^n |x_i|$ pour tout $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{K}^n$. Pour tous $\mathbf{x} = (x_1, \ldots, x_n)$ et $\mathbf{y} = (y_1, \ldots, y_n)$, on a

$$\begin{aligned}
\|f(\mathbf{x}) - f(\mathbf{y})\|_E &= \Big\|\sum_{i=1}^n x_i \cdot \mathbf{v}_i - \sum_{i=1}^n y_i \cdot \mathbf{v}_i\Big\|_E = \Big\|\sum_{i=1}^n (x_i - y_i) \cdot \mathbf{v}_i\Big\|_E \\
&\leqslant \sum_{i=1}^n \|(x_i - y_i) \cdot \mathbf{v}_i\|_E = \sum_{i=1}^n |x_i - y_i| \, \|\mathbf{v}_i\|_E \\
&\leqslant \big(\max_{i=1,\ldots,n} \|\mathbf{v}_i\|_E\big) \sum_{i=1}^n |x_i - y_i| = C \, \|\mathbf{x} - \mathbf{y}\|_1,
\end{aligned}$$

où $C = \max_{i\in\{1\ldots,n\}} \|\mathbf{v}_i\|_E$. On en déduit que pour tout $\varepsilon \in \mathbb{R}_+^*$ il existe $\eta \in \mathbb{R}_+^*$ (prendre $\eta = \varepsilon/C$) tel que pour tout $(\mathbf{x}, \mathbf{y}) \in \mathbb{K}^n \times \mathbb{K}^n$, si $\|\mathbf{x} - \mathbf{y}\|_1 \leqslant \eta$ alors $\|f(\mathbf{x}) - f(\mathbf{y})\|_E \leqslant C\eta = \varepsilon$. L'application f est donc uniformément continue de $\mathbb{K}^n$ dans E.

9.1.4 Applications lipschitziennes

Une classe de fonctions possédant la propriété d'être uniformément continues est celle des fonctions lipschitziennes. Nous verrons ultérieurement que certaines fonctions lipschitziennes, appelées fonctions contractantes, interviennent dans l'étude des points fixes.

DÉFINITION 9.8 *Soient (E, d_E) et (F, d_F) deux espaces métriques.*

✗ *Une application f de E dans F est* lipschitzienne *si*

$$\exists k \in \mathbb{R}^+ \quad \forall (x, y) \in E^2 \quad d_F(f(x), f(y)) \leqslant k \, d_E(x, y).$$

✗ *Une application f de E dans F est qualifiée d'application* lipschitzienne de rapport k, *où k désigne un réel positif, si*

$$\forall (x, y) \in E^2 \quad d_F(f(x), f(y)) \leqslant k \, d_E(x, y).$$

✗ *Une application f de E dans F est qualifiée d'application* contractante *s'il s'agit d'une application lipschitzienne de rapport $k \in [0, 1[$.*

Exemples

1. Une application f de (E, d_E) dans (F, d_F) qui est constante est lipschitzienne de rapport 0 puisque pour tout $(x, y) \in E^2$ on a $d_F(f(x), f(y)) = 0$. Réciproquement, toute application f de (E, d_E) dans (F, d_F) qui est lipschitzienne de rapport 0 est constante car si pour tout $(x, y) \in E^2$ on a $d_F(f(x), f(y)) \leqslant 0$ alors $d_F(f(x), f(y)) = 0$, autrement dit $f(x) = f(y)$.

2. Soit (E,d) est un espace métrique ; munissons E^2 de la distance δ_1 définie par $\delta_1(x,y) = d(x_1,y_1)+d(x_2,y_2)$ pour tous $x=(x_1,x_2)$ et $y=(y_1,y_2)$ dans E^2 et munissons $\mathbb{R}$ de la distance usuelle. L'application $f : (x_1,x_2) \in E^2 \mapsto d(x_1,x_2)$ est lipschitzienne de rapport 1. En effet, en utilisant la seconde inégalité triangulaire[29], on obtient pour tous $x=(x_1,x_2)\in E^2$ et $y=(y_1,y_2)\in E^2$,

$$\begin{aligned} |f(x)-f(y)| &= |d(x_1,x_2)-d(y_1,y_2)| \\ &= |d(x_1,x_2)-d(y_1,x_2)+d(y_1,x_2)-d(y_1,y_2)| \\ &\leqslant |d(x_1,x_2)-d(y_1,x_2)|+|d(y_1,x_2)-d(y_1,y_2)| \\ &\leqslant d(x_1,y_1)+d(x_2,y_2)=\delta_1(x,y). \end{aligned}$$

3. Soient $(E_1,d_1),\dots,(E_p,d_p)$ des espaces métriques et $E=E_1\times\cdots\times E_p$ l'espace produit muni de la distance δ_1 définie par la relation (7) p. 553. Pour tout entier k avec $1\leqslant k\leqslant n$, considérons la k-ème fonction de projection $p_k : (x_1,\dots,x_p)\in E\mapsto x_k\in E_k$. Il s'agit d'une application lipschitzienne de rapport 1 car pour tous $x=(x_1,\dots,x_p)\in E$ et $y=(y_1,\dots,y_p)$ dans E, on a

$$d_k(p_k(x),p_k(y))=d_k(x_k,y_k)\leqslant \max(d_1(x_1,y_1),\dots,d_p(x_p,y_p))=\delta_1(x,y).$$

4. On munit $\mathcal{M}_n(\mathbb{K})$ de la distance subordonnée à la norme de Schur[30], et on considère une matrice U dans $\mathcal{M}_n(\mathbb{K})$. Montrons que l'application $f : \mathrm{A}\in\mathcal{M}_n(\mathbb{K})\mapsto \mathrm{A}\times\mathrm{U}-\mathrm{U}\times\mathrm{A}$ est Lipschitzienne. Pour tous $\mathrm{A},\mathrm{B}\in\mathcal{M}_n(\mathbb{K})$ on a

$$\begin{aligned} \|f(\mathrm{A})-f(\mathrm{B})\|_s &= \|\mathrm{A}\times\mathrm{U}-\mathrm{U}\times\mathrm{A}-\mathrm{B}\times\mathrm{U}+\mathrm{U}\times\mathrm{B}\|_s \\ &= \|(\mathrm{A}-\mathrm{B})\times\mathrm{U}-\mathrm{U}\times(\mathrm{A}-\mathrm{B})\|_s \leqslant \|(\mathrm{A}-\mathrm{B})\times\mathrm{U}\|_s+\|\mathrm{U}\times(\mathrm{A}-\mathrm{B})\|_s. \end{aligned}$$

Or, on a établi au cours de l'exercice 7.8 p. 366 que pour la norme de Schur, on a $\|\mathrm{A}\times\mathrm{B}\|_s\leqslant\|\mathrm{A}\|_s\,\|\mathrm{B}\|_s$ pour tous $\mathrm{A},\mathrm{B}\in\mathcal{M}_n(\mathbb{K})$. On en déduit que

$$\|f(\mathrm{A})-f(\mathrm{B})\|_s\leqslant\|(\mathrm{A}-\mathrm{B})\|_s\,\|\mathrm{U}\|_s+\|\mathrm{U}\|_s\,\|\mathrm{A}-\mathrm{B}\|_s=2\|\mathrm{U}\|_s\,\|\mathrm{A}-\mathrm{B}\|_s.$$

On en conclut que f est lipschitzienne de rapport $k=2\,\|\mathrm{U}\|_s$.

La proposition suivante est très utile ; elle indique que toute application lipschitzienne est uniformément continue.

Proposition 9.16 *Soient (E,d_E) et (F,d_F) deux espaces métriques et f une application de E dans F. Si f est lipschitzienne alors f est uniformément continue sur E.*

[29] Voir la proposition 7.1 p. 351.

[30] Voir la proposition 7.5 p. 365 et l'exercice 7.8 p. 366.

Démonstration Une application de E dans F lipschitzienne de rapport 0 est constante et par conséquent uniformément continue sur E comme cela a été mentionné p. 557. Il s'agit donc d'établir que si f est lipschitzienne de rapport $k \in \mathbb{R}_+^*$ alors f est uniformément continue sur E, autrement dit de montrer l'assertion suivante[31] :

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists \eta \in \mathbb{R}_+^* \quad \forall (x,y) \in E^2 \quad \big(d_E(x,y) \leqslant \eta \implies d_F(f(x),f(y)) \leqslant \varepsilon\big).$$

Soient $\varepsilon \in \mathbb{R}_+^*$ et $(x,y) \in E^2$; puisque f est lipschitzienne de rapport k, on a $d_F(f(x),f(y)) \leqslant k\, d_E(x,y)$ et par conséquent, en prenant $\eta = \varepsilon/k$ on obtient :

$$d_E(x,y) \leqslant \eta \implies d_F(f(x),f(y)) \leqslant \varepsilon.$$

On en conclut que pour tout $\varepsilon \in \mathbb{R}_+^*$, il existe $\eta \in \mathbb{R}_+^*$ ($\eta = \varepsilon/k$ convient) tel que pour tout $(x,y) \in E^2$ vérifiant $d_E(x,y) \leqslant \eta$ on a $d_F(f(x),f(y)) \leqslant \varepsilon$. L'application f est donc uniformément continue sur E. □

Remarque Comme toute application uniformément continue est continue, toute application lipschitzienne est continue. ◇

Exercice 9.1 Soient (E,d_E) et (F,d_F) deux espaces métriques et f une application de E dans F lipschitzienne de rapport $k \in \mathbb{R}^+$. Montrer que si on munit E et F de deux distances δ_E et δ_F équivalentes à d_E et d_F respectivement, alors f est toujours une application lipschitzienne pour ces deux nouvelles distances sur E et F.

Un cas particulier d'applications lipschitziennes de rapport 1 est celui des isométries.

DÉFINITION 9.9 (Isométrie)
*Soient (E,d_E) et (F,d_F) deux espaces métriques. Une application f de E dans F est qualifiée d'*isométrie *si :*

$$\forall (x,y) \in E^2 \quad d_F(f(x),f(y)) = d_E(x,y).$$

Remarques

1. Une isométrie est toujours injective car pour tout $(x,y) \in E^2$ si $f(x) = f(y)$ alors $d_F(f(x),f(y)) = 0$ ce qui, par définition d'une isométrie, implique que $d_E(x,y) = 0$ autrement dit que $x = y$. Par contre, une isométrie n'est pas nécessairement surjective.

2. Une isométrie de E dans F étant une application lipschitzienne de rapport 1, elle est uniformément continue sur E.

[31] Voir la définition 9.7 p. 556.

3. On dit que deux espaces métriques sont *isométriques* s'il existe une isométrie surjective de l'un sur l'autre.

4. La notion d'isométrie est étroitement liée aux distances dont sont munis E et F. Si f est une isométrie de (E, d_E) dans (F, d_F) et si δ_E et δ_F sont des distances équivalentes à d_E et d_F alors f n'est pas nécessairement une isométrie de (E, δ_E) dans (F, δ_F). ◇

La notion d'isométrie sera étudiée en détail au chapitre 8 dans le cadre des espaces euclidiens, voir p. 490.

9.1.5 Applications continues et topologie

La proposition suivante, que nous admettrons, fournit un moyen d'établir qu'un ensemble est ouvert ou fermé.

PROPOSITION 9.17 *Soient (E, d_E) et (F, d_F) deux espaces métriques et f une application de E dans F. On a les équivalences suivantes :*

1. *l'application f est continue sur E ;*
2. *l'image réciproque par f de tout ouvert de F est un ouvert de E (i.e. quel que soit l'ouvert X de F, l'ensemble $f^{-1}(X)$ est ouvert dans E) ;*
3. *l'image réciproque par f de tout fermé de F est un fermé de E (i.e. quel que soit le fermé X de F, l'ensemble $f^{-1}(X)$ est fermé dans E) ;*

Exemples

1. Comme $\mathbb{R}_+^*$ est un ouvert de $\mathbb{R}$ et que $\mathbb{R}^+$ et $\{0\}$ sont des fermés de $\mathbb{R}$, si f est une application de (E, d) dans $\mathbb{R}$ continue alors :

- l'ensemble $\{x \in E \ ; \ f(x) > 0\}$ est un ouvert de E ;
- l'ensemble $\{x \in E \ ; \ f(x) = 0\}$ est un fermé de E ;
- l'ensemble $\{x \in E \ ; \ f(x) \geqslant 0\}$ est un fermé de E.

2. L'ensemble $\mathcal{GL}_n(\mathbb{K})$ des matrices inversibles de $\mathcal{M}_n(\mathbb{K})$ est un ouvert de $\mathcal{M}_n(\mathbb{K})$ car l'ensemble des matrices inversibles coïncide avec l'ensemble des matrices de déterminant non nul[32]. On a donc $\mathcal{GL}_n(\mathbb{K}) = \complement E$ où l'ensemble E est donné par $E = \{\mathrm{A} \in \mathcal{M}_n(\mathbb{K}) \ ; \ \det(A) = 0\}$. L'application déterminant étant continue sur $\mathcal{M}_n(\mathbb{K})$ (voir la proposition 9.12), on en déduit que l'ensemble E est fermé en tant qu'image réciproque du fermé $\{0\}$ de $\mathbb{R}$ par une application continue. Puisque E est fermé, son complémentaire $\mathcal{GL}_n(\mathbb{K})$ est ouvert.

[32] Voir la proposition 11.2 p. 478 du *Cours de première année*.

3. Soient (E, d_E) et (F, d_F) deux espaces métriques, et f et g deux applications continues de E dans F. Montrons que l'ensemble $\{x \in E \ ; \ f(x) = g(x)\}$ est fermé dans E. Munissons F^2 de l'une des distances produits données p. 556; d'après la proposition 9.15, l'application $\Phi : x \in E \mapsto (f(x), g(x)) \in F^2$ est continue de E. Considérons les ensembles $B = \{(y_1, y_2) \in F^2 \ ; \ y_1 = y_2\}$ et $\Phi^{-1}(B)$ image réciproque par Φ de B. On a

$$\begin{aligned} \Phi^{-1}(B) &= \{x \in E \ ; \ \exists y \in B \ y = \Phi(x)\} \\ &= \{x \in E \ ; \ \exists (y_1, y_2) \in F^2 \ \ y_1 = y_2 \text{ et } (y_1, y_2) = (f(x), g(x))\} \\ &= \{x \in E \ ; \ f(x) = g(x)\}. \end{aligned}$$

On a établi p. 410 que B est un fermé de F^2, donc d'après la proposition 9.17, l'ensemble $\Phi^{-1}(B)$ est un fermé. L'ensemble $\{x \in E \ ; \ f(x) = g(x)\}$ est donc fermé dans E.

L'image d'un ouvert par une application continue n'est pas nécessairement un ouvert. De même l'image d'un fermé par une application continue n'est pas nécessairement un fermé. Par exemple, la fonction $f : x \in \mathbb{R} \mapsto \frac{1}{x(x-1)}$ est continue sur l'ouvert $]0, 1[$ mais $f(]0, 1[) = [4, +\infty[$ n'est pas ouvert (c'est un intervalle fermé) ; elle est également continue sur le fermé $[2, +\infty[$ mais $f([2, +\infty[) =]0, \frac{1}{2}]$ n'est pas un fermé.

Le résultat suivant indique toutefois que l'image par une application continue d'un ensemble compact[33] est un ensemble compact.

Proposition 9.18 *Soient (E, d_E) et (F, d_F) deux espaces métriques et f une application de E dans F qui est* continue. *Si K est un sous-ensemble compact de E alors $f(K)$ est un sous-ensemble compact de F.*

Démonstration Pour montrer que $f(K) = \{y \in F \ ; \ \exists x \in K \ y = f(x)\}$ est compact, montrons que[34] de toute suite $(y_n)_n$ d'éléments de $f(K)$, on peut extraire une sous-suite qui converge vers un élément de $f(K)$. Soit $(y_n)_n$ une suite d'éléments de $f(K)$; pour tout $n \in \mathbb{N}$, il existe $x_n \in K$ tel que $y_n = f(x_n)$. Comme K est compact, on peut extraire de la suite $(x_n)_n$ une sous-suite $(x_{h(n)})_n$ (h étant l'application de $\mathbb{N}$ dans $\mathbb{N}$ strictement croissante désignant l'extractrice) qui converge vers un élément $x \in K$. Comme f est une application continue sur E, elle est continue en x et par conséquent[35] la suite de terme général $y'_n = f(x_{h(n)})$ converge vers $y = f(x)$. La suite $(y'_n)_n$ est par construction une suite extraite de la suite $(y_n)_n$ et elle converge vers un élément $y \in f(K)$ puisque $x \in K$. On peut donc extraire de la suite $(y_n)_n$ une sous-suite qui converge vers un élément de $f(K)$. Le résultat est démontré. □

[33] Voir la définition 7.23 p. 413.
[34] Voir la définition 7.23 p. 413.
[35] Voir la proposition 9.8 p. 549.

La notion de densité est d'un point de vue pratique très intéressante en analyse. Elle permet lorsque l'on souhaite montrer certaines propriétés pour les éléments d'un ensemble B de ne faire la vérification que pour les éléments d'un sous-ensemble A de B. On conclut alors en invoquant des résultats généraux de densité tel que celui donné dans la proposition suivante.

PROPOSITION 9.19 *Soient (E, d_E) et (F, d_F) deux espaces métriques, f et g deux applications continues de E dans F et A un sous-ensemble dense dans E. Si pour tout $x \in A$ on a $f(x) = g(x)$ alors pour tout $x \in E$ on a $f(x) = g(x)$.*

Démonstration Montrons que l'ensemble $B = \{x \in E \ ; \ f(x) = g(x)\}$ est égal à l'ensemble E. Par hypothèse, on a $A \subset B$ donc $\overline{A} \subset \overline{B}$. Par ailleurs, puisque f et g sont continues, l'ensemble B est fermé, voir p. 562. On a donc $\overline{B} = B$. On en déduit que $\overline{A} \subset B \subset E$. Mais comme A est dense dans E on a $\overline{A} = E$, ce qui implique que $B = E$. □

Exercice 9.2 Soient (E, d) un espace métrique et f, g deux applications de E dans $\mathbb{R}$. On munit $\mathbb{R}$ de la distance usuelle.
1 - Soient $A = \{x \in E \ ; \ f(x) > g(x)\}$, a un élément de A et α le réel $\frac{1}{2}(f(a) + g(a))$. On note $U = f^{-1}(]\alpha, +\infty[)$ et $V = g^{-1}(]-\infty, \alpha[)$.
a) Montrer que $U \cap V$ est un sous-ensemble de A et que $U \cap V$ est ouvert.
b) En déduire que A est un ouvert de (E, d).
2 - Soit D un sous-ensemble de E, dense dans E. Montrer que si $f(x) \leqslant g(x)$ pour tout $x \in D$ alors $f(x) \leqslant g(x)$ pour tout $x \in E$.

PROPOSITION 9.20 *Soient (E, d_E) et (F, d_F) deux espaces métriques et f une application continue de E dans F. Si A est un sous-ensemble de E qui est dense dans E alors $f(A)$ est dense dans $f(E)$.*

Démonstration Pour établir que $f(A)$ est dense dans $f(E)$, montrons que[36] pour tout $y \in f(E)$, il existe une suite $(y_n)_n$ d'éléments de $f(A)$ qui converge vers y. Soit y un élément de $f(E)$: il existe $x \in E$ tel que $y = f(x)$. Comme A est dense dans E, x est la limite d'une suite $(x_n)_n$ d'éléments de A. Puisque f est continue sur E, elle est continue en x et par conséquent[37] la suite de terme général $f(x_n)$ converge vers $f(x)$ (autrement dit vers y). Par ailleurs, comme pour tout $n \in \mathbb{N}$ on a $x_n \in A$, on en déduit que la suite de terme général $y_n = f(x_n)$ est une suite d'éléments de $f(A)$ qui converge vers y. □

[36] Voir le corollaire 7.2 p. 411.
[37] Voir la proposition 9.8, p. 549.

Le théorème suivant concerne les fonctions continues de $\mathbb{R}^n$ dans $\mathbb{R}$ et généralise un résultat connu pour les fonctions réelles de la variable réelle[38]. L'espace vectoriel $\mathbb{R}^n$ est muni de l'une quelconque de ses normes, notée $\|\cdot\|$ et $\mathbb{R}$ est muni de la distance usuelle.

THÉORÈME 9.2 *Soient X un ensemble non vide, fermé et borné dans $(\mathbb{R}^n, \|\cdot\|)$, et f une application continue de X dans $\mathbb{R}$. L'application f est bornée sur X et atteint ses bornes, i.e. :*

$$\text{il existe } (a,b) \in X^2 \text{ tel que } f(a) = \inf_{x\in X} f(x) \text{ et } f(b) = \sup_{x\in X} f(x).$$

Démonstration L'ensemble X qui est un ensemble fermé et borné de $\mathbb{R}^n$ est un ensemble compact, voir le théorème 7.7 p. 415. Comme f est une application continue de X dans $\mathbb{R}$, d'après la proposition 9.18, l'ensemble $f(X)$ est un compact de $\mathbb{R}$. C'est donc[39] un ensemble non vide et borné de $\mathbb{R}$; il admet par conséquent[40] une borne supérieure M et une borne inférieure m dans $\mathbb{R}$ qui vérifient :

$$\begin{array}{cccc} \big(\forall y \in f(X) \quad y \leqslant M\big) & \text{et} & \big(\forall \varepsilon \in \mathbb{R}_+^* \quad \exists y_\varepsilon \in f(X) \quad y_\varepsilon > M - \varepsilon\big) \\ \text{et} \quad \big(\forall y' \in f(X) \quad y \geqslant m\big) & \text{et} & \big(\forall \varepsilon \in \mathbb{R}_+^* \quad \exists y'_\varepsilon \in f(X) \quad y'_\varepsilon < m + \varepsilon\big). \end{array}$$

En considérant des réels $\varepsilon = 1/n$ avec $n \in \mathbb{N}^*$, on en déduit d'une part qu'il existe une suite $(y_n)_n$ d'éléments de $f(X)$ tel que pour tout $n \in \mathbb{N}^*$ on ait $M - 1/n < y_n \leqslant M$ et d'autre part qu'il existe une suite $(y'_n)_n$ d'éléments de $f(X)$ tel que pour tout $n \in \mathbb{N}^*$ on ait $m \leqslant y'_n < m + 1/n$. Le théorème d'encadrement indique que la suite $(y_n)_n$ d'éléments de $f(X)$ converge vers M et que la suite $(y'_n)_n$ d'éléments de $f(X)$ converge vers m. Comme $f(X)$ est compact, c'est un ensemble fermé[39] ce qui implique que $M \in f(X)$ et que $m \in f(X)$ (voir le théorème 7.6 p. 410), autrement dit qu'il existe $(a,b) \in X^2$ tel que $f(a) = \inf_{x\in X} f(x)$ et $f(b) = \sup_{x\in X} f(x)$. □

Remarque On sera attentif aux hypothèses du théorème 9.2. Les exemples suivants précisent l'importance de chacune des hypothèses.

1. Soit $X = \{(x_1,x_2) \in \mathbb{R}^2 \ ; \ 0 < x_1^2 + x_2^2 \leqslant 1\}$; considérons l'application $f : (x_1,x_2) \in X \mapsto (x_1^2+x_2^2)^{-\frac{1}{2}}$ qui donne l'inverse de la distance euclidienne du point de coordonnées (x_1,x_2) à l'origine. Cette application est continue sur l'ensemble X qui est borné mais pas fermé. Elle n'est pas bornée sur X puisque pour tout $M \in \mathbb{R}^+$ le point $(\frac{1}{M+1}, 0)$ a pour image par f le réel $\sqrt{(M+1)^2}$ qui est strictement supérieur à M.

[38] Voir le théorème 13.4 p. 596 du *Cours de première année*.

[39] Voir la proposition 7.20 p. 414.

[40] Voir la proposition 3.2 p. 95 du *Cours de première année*.

2. L'application $f : (x_1, x_2) \in \mathbb{R}^2 \mapsto \sqrt{x_1^2 + x_2^2}$ qui donne la distance euclidienne du point de coordonnées (x_1, x_2) à l'origine est continue sur $\mathbb{R}^2$ qui est un ensemble fermé mais non borné. L'application f n'est pas bornée.
3. L'application $f : (x_1, x_2) \in B(0,1) \mapsto \sqrt{x_1^2 + x_2^2}$ où $B(0,1)$ désigne la boule ouverte $\{(x_1, x_2) \in \mathbb{R}^2 \ ; \ x_1^2 + x_2^2 < 1\}$, est continue sur $B(0,1)$ qui est un ensemble borné mais non fermé. Cette application est bornée, mais n'atteint pas sa borne supérieure 1 (elle atteint toutefois sa borne inférieure qui est 0).

On sera par ailleurs attentif au fait que le théorème 9.2 fournit une condition suffisante pour qu'une application atteigne ses bornes. Cette condition n'est pas nécessaire : par exemple l'application $f : (x_1, x_2) \in X \mapsto \sqrt{x_1^2 + x_2^2}$ où $X = B(0,1) \cup \{(1,0)\}$ est bornée et atteint ses bornes bien que X ne soit pas fermé. ◇

9.2 Limite et continuité des fonctions de $\mathbb{R}^n$ dans $\mathbb{R}^p$

Considérons une application f définie sur un ouvert U de $\mathbb{R}^n$ ($n \in \mathbb{N}^*$) à valeurs dans $\mathbb{R}^p$ ($p \in \mathbb{N}^*$) :

$$f : \begin{array}{ccc} U \subset \mathbb{R}^n & \longrightarrow & \mathbb{R}^p \\ x = (x_1, \ldots, x_n) & \longmapsto & (f_1(x), \ldots, f_p(x)) \end{array}$$

où pour tout $i \in \mathbb{N}$ avec $1 \leqslant i \leqslant p$, f_i désigne la i–ième application composante de f.

✗ Il résulte de la proposition 9.15 que f est continue en $a \in U$ (resp. admet une limite en a) si et seulement si pour tout entier i avec $1 \leqslant i \leqslant p$ l'application f_i est continue en $a \in U$ (resp. admet une limite en a). En conséquence, l'étude de la continuité (resp. le calcul d'une limite) pour une fonction de $\mathbb{R}^n$ dans $\mathbb{R}^p$ consiste à étudier la continuité (resp. à calculer la limite) de p fonctions de $\mathbb{R}^n$ dans $\mathbb{R}$.

✗ Comme $\mathbb{R}^n$ est un $\mathbb{R}$-espace vectoriel de dimension finie, pour l'étude de la continuité de f (resp. l'étude de l'existence d'une limite pour f), on peut prendre dans la définition 9.5 (resp. la définition 9.1) pour distance d_E n'importe quelle distance subordonnée à une norme ; elles sont toutes équivalentes. D'un point de vue pratique, on choisit celle pour laquelle les calculs sont les plus aisés.

Exemples

1. Pour montrer que l'application $f : (x_1, x_2) \in \mathbb{R}^2_* \mapsto \dfrac{x_1 x_2}{\sqrt{x_1^2 + x_2^2}}$ admet pour limite 0 en $(0,0)$, le raisonnement est facilité par le choix de la distance euclidienne. Compte tenu de l'inégalité $|x_1 x_2| \leqslant 2|x_1|\,|x_2| \leqslant x_1^2 + x_2^2 = \|(x_1, x_2)\|_2^2$, on a pour tout $(x_1, x_2) \in \mathbb{R}^2_*$,

$$|f(x_1, x_2) - 0| = \left| \frac{x_1 x_2}{\sqrt{x_1^2 + x_2^2}} \right| = \frac{|x_1 x_2|}{\|(x_1, x_2)\|_2} \leqslant \|(x_1, x_2)\|_2.$$

On en déduit que pour tout $\varepsilon \in \mathbb{R}_+^*$, on a $|f(x_1,x_2)-0| \leqslant \varepsilon$ pour tout $(x_1,x_2) \in \mathbb{R}_*^2$ tel que $\|(x_1,x_2)-(0,0)\|_2 \leqslant \varepsilon$. D'après la définition 9.1, cela indique que f admet pour limite 0 en $(0,0)$.

2. Pour montrer que l'application $f : (x_1,x_2) \in \mathbb{R}_*^2 \mapsto \dfrac{(x_1+x_2)^2}{|x_1|+|x_2|}$ admet pour limite 0 en $(0,0)$, il est préférable de considérer la distance subordonnée à la norme 1. Pour tout $(x_1,x_2) \in \mathbb{R}_*^2$ on a :

$$|f(x_1,x_2)-0| = \left|\frac{(x_1+x_2)^2}{|x_1|+|x_2|}\right| = \frac{(x_1+x_2)^2}{\|(x_1,x_2)\|_1} \leqslant \frac{(|x_1|+|x_2|)^2}{\|(x_1,x_2)\|_1} = \|(x_1,x_2)\|_1.$$

On en déduit que pour tout $\varepsilon \in \mathbb{R}_+^*$, on a $|f(x_1,x_2)-0| \leqslant \varepsilon$ pour tout $(x_1,x_2) \in \mathbb{R}_*^2$ tel que $\|(x_1,x_2)-(0,0)\|_1 \leqslant \varepsilon$. D'après la définition 9.1, cela indique que f admet pour limite 0 en $(0,0)$.

Dans le cas des applications définies sur un ouvert de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}$, on a les résultats suivants qui généralisent les propriétés liées à la continuité des applications réelles de la variable réelle[41].

Proposition 9.21 *Soient λ un réel, U un ouvert de $\mathbb{R}^n$ et f et g deux applications de U dans $\mathbb{R}$. Si f et g sont continues en $\mathbf{a} \in U$ (resp. sur U) alors : :*

1. *$|f|$ est continue en $\mathbf{a}$ (resp. sur U) ;*
2. *$f+g$ est continue en $\mathbf{a}$ (resp. sur U) ;*
3. *$f \times g$ est continue en $\mathbf{a}$ (resp. sur U) ;*
4. *$\lambda \cdot f$ est continue en $\mathbf{a}$ (resp. sur U) ;*
5. *si de plus $g(\mathbf{a}) \neq 0$ (resp. g ne s'annule pas sur U) alors f/g est continue en $\mathbf{a}$ (resp. sur U).*

Rappelons qu'un tel résultat ne peut être énoncé pour des applications d'un espace métrique E dans un espace métrique F de façon générale, faute de disposer des lois de composition interne $+$ et $\times$ sur F.

Un moyen pratique d'établir qu'une fonction de $\mathbb{R}^n$ dans $\mathbb{R}$ n'a pas de limite

La méthode présentée p. 181 pour montrer qu'une fonction complexe de la variable complexe n'avait pas de limite se généralise, comme nous allons le voir, aux fonctions de $\mathbb{R}^n$ dans $\mathbb{R}$.

Il résulte de la proposition 9.14 qu'une condition nécessaire pour qu'une application définie sur un ouvert U de $\mathbb{R}^n$ et à valeurs réelles admette en $\mathbf{a} \in U$ pour limite le réel ℓ est que les fonctions partielles

$$f_i : t \in \mathbb{R} \mapsto f(a_1,\ldots,a_{i-1},t,a_{i+1},\ldots,a_n) \in \mathbb{R}$$

[41] Voir la proposition 13.18 p. 593 du *Cours de première année*.

admettent toutes pour limite le réel ℓ lorsque t tend vers a_i. Ainsi, pour montrer qu'une fonction f de $\mathbb{R}^n$ dans $\mathbb{R}$ n'a pas de limite en $\mathbf{a} = (a_1, \ldots, a_n)$, il suffit :

- de montrer qu'une des fonctions partielles f_i n'a pas de limite en a_i ;
- ou bien que deux fonctions partielles f_i et f_j n'ont pas le même réel pour limite en a_i et a_j respectivement.

Exemples

1. L'application $f : (x_1, x_2) \in \mathbb{R}^2_* \mapsto \dfrac{\operatorname{ch}(x_1)\operatorname{sh}(x_2)}{x_1 + x_2}$ n'a pas de limite en $(0,0)$ car les applications partielles en $(0,0)$ qui sont :

$$f_1 : x_1 \in \mathbb{R}^* \longmapsto 0 \quad \text{et} \quad f_2 : x_2 \in \mathbb{R}^* \longmapsto \frac{\operatorname{sh}(x_2)}{x_2}$$

ont pour limites 0 et 1 respectivement en $x_1 = 0$ et $x_2 = 0$.

2. L'application $f : (x_1, x_2) \in \mathbb{R}^2_* \mapsto \dfrac{\cos(x_1)\sin(x_2)}{x_1^2 + x_2^2}$ n'a pas de limite en $(0,0)$ car la deuxième application partielle $f_2 : x_2 \in \mathbb{R}^* \mapsto \sin(x_2)/x_2^2$ n'a pas de limite en $x_2 = 0$.

Même si toutes les fonctions partielles ont la même limite, on ne peut rien conclure quant à l'existence d'une limite pour f.

On a toutefois l'information suivante : si f admet une limite en $\mathbf{a}$, sa valeur est la valeur de la limite des fonctions partielles. Par exemple, l'application $f : (x_1, x_2) \in \mathbb{R}^2_* \mapsto \dfrac{x_1 x_2}{x_1^2 + x_2^2}$ admet pour applications partielles en $(0,0)$

$$f_1 : x_1 \in \mathbb{R}^* \longmapsto f(x_1, 0) = 0 \quad \text{et} \quad f_2 : x_2 \in \mathbb{R}^* \longmapsto f(0, x_2) = 0$$

qui ont une limite nulle en $x_1 = 0$ et $x_2 = 0$ respectivement. On ne peut pas dire si f a une limite ou non en $(0,0)$, mais on peut affirmer que si f a une limite en $(0,0)$ cette limite est obligatoirement 0.

Nous venons de voir que des informations sur l'existence d'une limite pour une fonction f de $\mathbb{R}^n$ dans $\mathbb{R}$ pouvaient être obtenues en considérant les applications partielles. S'intéresser à la limite des applications partielles revient à considérer le limite de la fonction f en restriction à des droites parallèles aux axes de coordonnées ou, en d'autres termes, à considérer la limite de la quantité $f(\mathbf{x})$ lorsque $\mathbf{x}$ tend vers $\mathbf{a}$ selon des droites passant par $\mathbf{a}$ et parallèles aux axes de coordonnées.

Plus généralement, si f admet pour limite le réel ℓ en $\mathbf{a}$ et si C désigne une courbe paramétrée[42] de $\mathbb{R}^n$ de classe $\mathcal{C}^1$ contenant $\mathbf{a}$, alors la restriction à la courbe paramétrée C de la fonction f admet aussi pour limite le réel ℓ en $\mathbf{a}$. Autrement dit, si $\varphi \in \mathcal{C}^1(I, \mathbb{R}^n)$ est une paramétrisation de C telle que $\mathbf{a} = \varphi(t_0), t_0 \in I$ alors la limite lorsque t tend vers t_0 de $f(\varphi(t))$ est égale[43] à ℓ.

---

[42] L'étude détaillée des courbes paramétrées dans $\mathbb{R}^n$ sera effectuée au chap. 11, mais aucune connaissance précise en ce qui concerne les courbes paramétrées n'est requise à ce niveau.

[43] Voir la proposition 9.4 p. 542.

On dispose ainsi d'un moyen supplémentaire pour montrer qu'une fonction f de $\mathbb{R}^n$ dans $\mathbb{R}$ n'a pas de limite en $\mathbf{a}$.

1. On peut chercher un arc paramétré φ ayant $\mathbf{a}$ dans son image (*i.e* $\exists t_0 \in \mathbb{R}\ \mathbf{a} = \varphi(t_0)$) tel que la fonction $f \circ \varphi$ n'ait pas de limite en t_0.
2. On peut chercher deux arcs paramétrés distincts φ et ψ ayant $\mathbf{a}$ dans leur image (*i.e* $\exists (s_0, t_0) \in \mathbb{R}^2\ \varphi(t_0) = \psi(s_0) = \mathbf{a}$) tels que la limite de la fonction $f \circ \varphi$ en t_0 et la limite de la fonction $f \circ \psi$ en s_0 soient distinctes.

Exemples

1. L'application $f : (x_1, x_2) \in \mathbb{R}^2_* \mapsto \dfrac{x_1 x_2}{x_1^2 + x_2^2}$ n'a pas de limite en $(0,0)$. Considérons l'arc paramétré $\varphi : t \in \mathbb{R} \mapsto (t,t) \in \mathbb{R}^2$ dont l'image est la première bissectrice ; on a $(0,0) = \varphi(t_0)$ avec $t_0 = 0$ et pour tout $t \in \mathbb{R}$, on a $f(\varphi(t)) = \frac{1}{2}$. On en déduit que l'application $f \circ \varphi$ admet pour limite $\frac{1}{2}$ lorsque t tend vers $t_0 = 0$ alors que l'on a vu que les applications partielles ont pour limite 0.

2. L'application $f : (x_1, x_2) \in \mathbb{R}^2_* \mapsto \dfrac{x_1 x_2^2}{x_1^4 + x_2^3}$ n'a pas de limite en $(0,0)$. Considérons l'arc paramétré $\varphi : t \in \mathbb{R} \mapsto (t^3, t^4) \in \mathbb{R}^2$; on a $(0,0) = \varphi(0)$ et pour tout $t \in \mathbb{R}^*$, on a $f(\varphi(t)) = 1/t$. Comme $f \circ \varphi$ n'a pas de limite en 0, f n'a pas non plus de limite en $(0,0)$.

D'un point de vue pratique, il est commode de considérer simultanément toutes les droites de $\mathbb{R}^n$ concourantes en $\mathbf{a}$ dont les paramétrisations sont $\varphi : t \in \mathbb{R} \mapsto \mathbf{a} + t\,\mathbf{u}$ où $\mathbf{u}$ désigne le vecteur dont toutes les composantes sont égales à un réel λ fixé à l'exception de l'une d'entre elle qui vaut 1. On a $\varphi(0) = \mathbf{a}$ et si la limite lorsque t tend vers 0 de $f(\varphi(t))$ dépend de λ alors f n'a pas de limite en $\mathbf{a}$.

Exemples

1. Intéressons-nous à l'existence d'une limite en $(0,0)$ à l'application

$$f : (x_1, x_2) \in \mathbb{R}^2_* \longmapsto \frac{|x_1| + |x_2|}{\sqrt{x_1^2 + x_2^2}}.$$

Considérons l'arc paramétré $\varphi : t \in \mathbb{R} \mapsto t\,\mathbf{u}$ avec $\mathbf{u} = (1, \lambda)$ où λ désigne un réel fixé. On a $(0,0) = \varphi(0)$ et

$$\forall t \in \mathbb{R} \qquad f(\varphi(t))) = \frac{(1 + |\lambda|)|t|}{\sqrt{(1+\lambda^2)t^2}} = \frac{1 + |\lambda|}{\sqrt{1+\lambda^2}}.$$

Comme la limite en 0 de $f(\varphi(t)))$ dépend de la valeur de λ, l'application f n'a pas de limite en $(0,0)$.

2. L'application

$$f : (x_1, x_2, x_3) \in \mathbb{R}^3_* \longmapsto \frac{|x_1 x_2 x_3|}{(x_1^2 + x_2^2 + x_3^2)^{\frac{3}{2}}}$$

n'a pas de limite en 0. Considérons l'arc paramétré $\varphi : t \in \mathbb{R} \mapsto t\,\mathbf{u}$ avec $\mathbf{u} = (1, \lambda, \lambda)$ où λ désigne un réel fixé ; on a $(0,0,0) = \varphi(0)$ et

$$\forall t \in \mathbb{R} \qquad f(\varphi(t))) = \frac{\lambda^2}{\sqrt{1+2\lambda^2}}.$$

Comme la limite de $f \circ \varphi$ en 0 dépend de λ, l'application f n'a pas de limite en $(0,0,0)$.

Comment établir qu'une fonction de $\mathbb{R}^n$ dans $\mathbb{R}$ a une limite ?

Considérons une application f définie sur un ensemble D de $\mathbb{R}^n$ et à valeurs dans $\mathbb{R}$; munissons $\mathbb{R}^n$ d'une quelconque de ses normes $\|\cdot\|$ et $\mathbb{R}$ de la distance usuelle.

D'après la définition 9.1, l'application f admet pour limite en $\mathbf{a} \in \overline{D}$ le réel ℓ si

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists \eta \in \mathbb{R}_+^* \quad \forall \mathbf{x} \in D \quad \big(\|\mathbf{x}-\mathbf{a}\| \leqslant \eta \Longrightarrow |f(\mathbf{x}) - \ell| \leqslant \varepsilon\big).$$

Cette assertion indique que pour montrer que f admet ℓ pour limite en $\mathbf{a}$, il faut et il suffit de montrer que l'application $g : \mathbf{x} \in D \mapsto |f(\mathbf{x}) - \ell| \in \mathbb{R}^+$ admet pour limite 0 en $\mathbf{a}$. Il résulte par ailleurs du théorème 9.1 que si h est une fonction à valeurs réelles qui admet 0 pour limite en $\mathbf{a}$ et telle qu'au voisinage de $\mathbf{a}$ on ait $0 \leqslant g(\mathbf{x}) \leqslant h(\mathbf{x})$ alors g admet pour limite 0 en $\mathbf{a}$.

D'un point de vue pratique, pour établir qu'une fonction f de $\mathbb{R}^n$ dans $\mathbb{R}$ a une limite en $\mathbf{a}$ on peut procéder ainsi.

1. On cherche la valeur ℓ de la limite de f en $\mathbf{a}$ en considérant un arc paramétré φ ayant $\mathbf{a}$ dans son image (*i.e.* $\exists t_0 \in \mathbb{R}\ \mathbf{a} = \varphi(t_0)$) et on calcule la limite de la restriction de f à la courbe paramétrée associée à φ, c'est-à-dire, on calcule la limite de $f \circ \varphi$ en t_0.
2. On majore au voisinage de $\mathbf{a}$ la fonction $g : \mathbf{x} \in \mathbb{R}^n \mapsto |f(\mathbf{x}) - \ell| \in \mathbb{R}$ par une fonction h dont on sait que la limite en $\mathbf{a}$ est 0.

Exemples

1. Considérons l'application $f : (x_1, x_2) \in \mathbb{R}_*^2 \mapsto \dfrac{\sin(x_1 x_2)}{|x_1| + |x_2|}$. Ses applications partielles en $(0,0)$ correspondent à l'application nulle. On en déduit que si f a une limite en $(0,0)$, cette limite est nécessairement 0. En utilisant les inégalités $|\sin(t)| \leqslant |t|$ pour tout $t \in \mathbb{R}$ et $|x_1 x_2| \leqslant (|x_1| + |x_2|)^2$, on obtient que pour tout $(x_1, x_2) \in \mathbb{R}^2$

$$|f(x_1,x_2) - 0| = \frac{|\sin(x_1 x_2)|}{|x_1| + |x_2|} \leqslant \frac{|x_1 x_2|}{|x_1| + |x_2|} \leqslant |x_1| + |x_2| = \|(x_1,x_2)\|_1.$$

Puisque $\lim_{(x_1,x_2)\to(0,0)} \|(x_1,x_2)\|_1 = 0$, on en déduit que f admet pour limite 0 en $(0,0)$.

2. Considérons l'application $f : (x_1, x_2, x_3) \in \mathbb{R}^3_* \mapsto \dfrac{x_1 x_2 x_3}{\sqrt{x_1^2 + x_2^2 + x_3^2}}$ en $(0, 0, 0)$. Ses applications partielles à l'origine coïncidant avec l'application nulle, si f a une limite en $(0, 0, 0)$ ce ne peut être que 0. Pour tout $\mathbf{x} = (x_1, x_2, x_3) \in \mathbb{R}^3_*$, on a[(44)]

$$|f(x_1, x_2, x_3)| = \frac{|x_1|\,|x_2|\,|x_3|}{\sqrt{x_1^2 + x_2^2 + x_3^2}} \leqslant \frac{1}{6}\,\frac{(|x_1| + |x_2\| + |x_3|)^3}{\sqrt{x_1^2 + x_2^2 + x_3^2}} = \frac{1}{6}\frac{\|\mathbf{x}\|_1^3}{\|\mathbf{x}\|_2}.$$

D'après la proposition 7.4 p. 362, on a $\|\mathbf{x}\|_1 \leqslant \sqrt{3}\,\|\mathbf{x}\|_2$. On en déduit que $|f(x_1, x_2, x_3)| \leqslant \frac{\sqrt{3}}{2}\,\|\mathbf{x}\|_2^2$. Le théorème d'encadrement permet de conclure que f admet pour limite 0 en $(0, 0, 0)$.

3. Déterminons la limite de l'application $f : (x_1, x_2) \in \mathbb{R}^2_* \mapsto x_1 x_2 \ln(x_1^2 + x_2^2)$ en $(0, 0)$. Ses applications partielles à l'origine coïncidant avec l'application nulle, si f a une limite en $(0, 0)$ ce ne peut être que 0. Pour tout $(x_1, x_2) \in \mathbb{R}^2_*$, on a $2|x_1 x_2| \leqslant x_1^2 + x_2^2$ d'où

$$\begin{aligned} |f(x_1, x_2)| &= |x_1 x_2 \ln(x_1^2 + x_2^2)| \leqslant (x_1^2 + x_2^2) \ln(\sqrt{x_1^2 + x_2^2}) \\ &= \|(x_1, x_2)\|_2^2 \ln(\|(x_1, x_2)\|_2). \end{aligned}$$

Comme $\lim_{t \to 0^+} t^2 \ln(t) = 0$, on en conclut que f admet pour limite 0 en $(0, 0)$.

Exercice 9.3 Déterminer, si elles existent, les limites des fonctions suivantes en $(0, 0)$:

1 - $(x_1, x_2) \in \mathbb{R}^2 \mapsto \dfrac{x_1}{\sqrt{x_1^2 + x_2^2}}$ 2 - $(x_1, x_2) \in \mathbb{R}^2 \mapsto \dfrac{|\sin(x_1 + x_2)|}{\sqrt{x_1^2 + x_2^2}}$

3 - $(x_1, x_2) \in \mathbb{R}^2 \mapsto \dfrac{\sin(x_1^2 + x_2^2)}{|x_1| + |x_2|}$ 4 - $f : (x_1, x_2) \in \mathbb{R}^2 \mapsto \dfrac{\ln(1 + x_1 x_2)}{\sqrt{x_1^2 + x_2^2}}$

5 - $(x_1, x_2) \in \mathbb{R}^2 \mapsto \dfrac{x_1 x_2^2}{x_1^4 + x_2^3}$ 6 - $(x_1, x_2) \in \mathbb{R}^2 \mapsto \dfrac{x_1 x_2^2}{x_1^4 + x_2^2}$

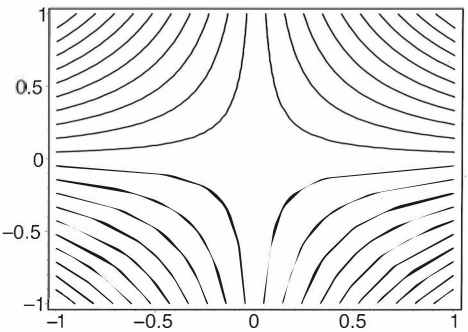
Représentation graphique

La visualisation de la représentation graphique d'une fonction de $\mathbb{R}^n$ dans $\mathbb{R}$ n'est pas aisée et n'est guère envisageable que lorsque $n = 2$. Le logiciel MAPLE dispose de plusieurs commandes permettant de représenter graphiquement des fonctions de deux variables réelles soit sous la forme de lignes de niveau, soit sous la forme de surfaces. L'étude de telles courbes et surfaces sera présentée au chapitre 11 ; nous nous contenterons dans cette partie de présenter les différents outils de visualisation des fonctions de $\mathbb{R}^2$ dans $\mathbb{R}$ disponibles avec MAPLE.

(44) On utilise ici l'inégalité $6abc \leqslant (a + b + c)^3$ qui est vraie pour tous réels positifs a, b, c ; on peut le vérifier en développant le membre de droite de l'inégalité.

Considérons la fonction $f : (x_1, x_2) \in \mathbb{R}^2 \mapsto x_1 x_2 / \sqrt{x_1^2 + x_2^2}$. On peut visualiser cette fonction en dessinant les courbes de niveaux de f, c'est-à-dire les courbes d'équation $f(x, y) = C$ où C est un paramètre réel. Pour ce faire, on dispose avec MAPLE de la commande `contourplot` (elle requiert l'utilisation de la librairie `plots`) dont l'utilisation est illustrée sur l'exemple suivant.

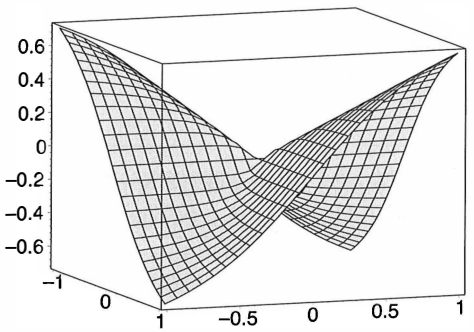
```
> f:=(x,y)-> x*y/sqrt(x^2+y^2):
> with(plots):
> contourplot(f(x,y), x=-1..1, y=-1..1,contours=20);
```



Le nombre de lignes de niveau est spécifié par l'option `contours`.

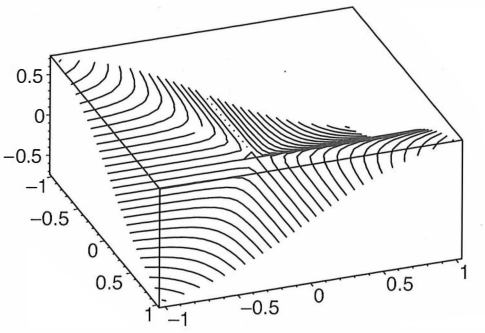
La commande `plot3d` permet de visualiser la représentation graphique d'une fonction f de $\mathbb{R}^2$ dans $\mathbb{R}$ sous forme de surface dans l'espace.

```
plot3d(f(x,y), x=-1..1, y=-1..1);
```



Il existe également la commande `contourplot3d` qui permet de concilier les deux types de représentation précédents en visualisant les lignes de niveau de f sur la surface dont l'équation est $z = f(x, y)$.

```
> contourplot3d(f(x,y), x=-1..1, y=-1..1,contours=20);
```



Signalons que l'on peut visualiser une fonction de $\mathbb{R}^2$ dans $\mathbb{R}^p$ en représentant séparément chacune des p fonctions composantes.

9.3 Continuité des applications linéaires

Si (E, N_E) et (F, N_F) sont deux $\mathbb{K}$-espaces vectoriels normés, on dit que l'application f de E dans F est continue si :

$$\forall \mathbf{x}_0 \in E \quad \forall \varepsilon \in \mathbb{R}_+^* \quad \exists \eta \in \mathbb{R}_+^* \quad \forall \mathbf{x} \in E \\ \big(N_E(\mathbf{x} - \mathbf{x}_0) \leqslant \eta \implies N_F(f(\mathbf{x}) - f(\mathbf{x}_0)) \leqslant \varepsilon\big).$$

Cette assertion correspond à la définition 9.5 où les distances considérées sur E et F sont celles subordonnées aux normes N_E et N_F respectivement.

Dans le cas où E (resp. F) est un $\mathbb{K}$-espace vectoriel de dimension finie, toutes les normes sur E (resp. F) sont équivalentes. Comme la notion de continuité est inchangée par le choix de distances équivalentes, il n'est pas nécessaire en dimension finie de préciser la norme qui a été choisie. On prendra garde au fait qu'une application linéaire entre $\mathbb{K}$-espaces vectoriels de dimension infinie peut être continue pour un certain choix de normes et non continue pour un autre choix de normes sur ces espaces.

Proposition 9.22 *Soit (E, N) un $\mathbb{K}$-espace vectoriel normé ; l'application $N : \mathbf{x} \in E \mapsto N(\mathbf{x})$ est continue de (E, N) dans $(\mathbb{R}, |\ |)$.*

Démonstration Si $\mathbf{x}$ et $\mathbf{y}$ désignent deux éléments de E, on a

$$N(\mathbf{x}) \leqslant N(\mathbf{y}) + N(\mathbf{x} - \mathbf{y}) \quad \text{et} \quad N(\mathbf{y}) \leqslant N(\mathbf{x}) + N(\mathbf{x} - \mathbf{y}).$$

On en déduit que

$$-N(\mathbf{x} - \mathbf{y}) \leqslant N(\mathbf{x}) - N(\mathbf{y}) \leqslant N(\mathbf{x} - \mathbf{y}),$$

autrement dit, que $|N(\mathbf{x}) - N(\mathbf{y})| \leqslant N(\mathbf{x} - \mathbf{y})$. Ainsi, N est une application lipschitzienne de rapport 1 ; d'après la proposition 9.16, elle est continue. □

Rappelons[45] qu'une application linéaire d'un $\mathbb{K}$-espace vectoriel E vers un $\mathbb{K}$-espace vectoriel F est une application de E dans F vérifiant :

1. $\forall \mathbf{x} \in E \quad \forall \mathbf{y} \in E \quad f(\mathbf{x}+\mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y})$,
2. $\forall \mathbf{x} \in E \quad \forall \alpha \in \mathbb{K} \quad f(\alpha \cdot \mathbf{x}) = \alpha \cdot f(\mathbf{x})$.

On note $\mathcal{L}_K(E,F)$ (ou plus simplement $\mathcal{L}(E,F)$) l'ensemble des applications linéaires de E dans F. Cet ensemble muni des lois $+$ (somme d'applications) et $\cdot$ (multiplication d'une application par un scalaire) possède une structure d'espace vectoriel sur $\mathbb{K}$.

Rappelons aussi qu'une condition nécessaire et suffisante pour qu'une application f d'un $\mathbb{K}$-espace vectoriel E vers un $\mathbb{K}$-espace vectoriel F soit une application linéaire est que pour tout $(\mathbf{x},\mathbf{y}) \in E^2$ et pour tout $(\alpha,\beta) \in \mathbb{K}^2$ on ait :

$$f(\alpha \cdot \mathbf{x} + \beta \cdot \mathbf{y}) = \alpha \cdot f(\mathbf{x}) + \beta \cdot f(\mathbf{y}).$$

Exercice 9.4 Soit f une application linéaire d'un $\mathbb{K}$-espace vectoriel normé $(E, \|\cdot\|_E)$ vers un $\mathbb{K}$-espace vectoriel normé $(F, \|\cdot\|_F)$.
1 - Montrer que $f(\mathbf{0}_E) = \mathbf{0}_F$.
2 - Soit U un ouvert non vide de E ; montrer que si pour tout $\mathbf{x} \in U$ on a $f(\mathbf{x}) = \mathbf{0}_F$ alors f est l'application nulle.

Terminons par quelques rappels de terminologie concernant les applications linéaires.

✗ On appelle *isomorphisme* d'espaces vectoriels une application linéaire de E dans F qui est bijective. On dit alors que E et F sont isomorphes par f.

✗ On appelle *endomorphisme* de E une application linéaire de l'espace E dans lui-même. On note $\mathcal{L}(E)$ l'ensemble des endomorphismes de E. Cet ensemble muni des lois $+$ (somme d'applications), $\cdot$ (multiplication d'une application par un scalaire) et $\circ$ (composition d'applications) possède une structure de $\mathbb{K}$-algèbre.

✗ En particulier, si un endomorphisme f est bijectif alors on dit que f est un *automorphisme* de l'espace E. On note $\mathcal{GL}(E)$ l'ensemble des automorphismes de E.

Le théorème suivant est d'une grande importance pratique puisqu'il fournit plusieurs moyens de montrer qu'une application linéaire est continue. Il indique en particulier que dès qu'une application linéaire est continue en 0, elle est continue dans tout l'espace.

[45] Voir le chap. 9 du *Cours de première année*.

THÉORÈME 9.3 *Soient* $(E, \|\cdot\|_E)$ *et* $(F, \|\cdot\|_F)$ *deux* $\mathbb{K}$*-espaces vectoriels normés et* f *une* application linéaire *de* E *dans* F. *Les assertions suivantes sont équivalentes :*

1. *l'application* f *est continue sur* E ;
2. *l'application* f *est continue en* $\mathbf{0}_E$ *(*$\mathbf{0}_E$ *désignant le vecteur nul dans* E*)* ;
3. *l'application* f *est bornée sur la boule fermée* $\overline{B}(\mathbf{0}_E, 1)$;
4. *il existe un réel positif* C *tel que pour tout* $\mathbf{x} \in E$: $\|f(\mathbf{x})\|_F \leqslant C\|\mathbf{x}\|_E$.

Démonstration Comme une application est continue sur un ensemble si elle est continue en chaque point de l'ensemble, l'assertion (1) implique l'assertion (2). Nous allons montrer que l'assertion (2) implique l'assertion (3) et que l'assertion (3) implique l'assertion (4). Par transitivité de la relation d'implication l'assertion (1) impliquera l'assertion (3) et l'assertion (4), et l'assertion (2) impliquera l'assertion (4). Pour établir l'équivalence des quatre assertions, il suffira ensuite de montrer que l'assertion (4) implique l'assertion (1). La transitivité de la relation d'implication établira l'équivalence entre toutes les relations.

$\trianglerighteq$ Supposons que f est continue en $\mathbf{0}_E$; puisque f est linéaire, on a $f(\mathbf{0}_E) = \mathbf{0}_F$ et l'hypothèse de continuité de f se traduit par

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists \eta \in \mathbb{R}_+^* \quad \forall \mathbf{x} \in E \quad (\|\mathbf{x}\|_E \leqslant \eta \implies \|f(\mathbf{x})\|_F \leqslant \varepsilon).$$

On en déduit qu'il existe $\eta \in \mathbb{R}_+^*$ tel que

$$\forall \mathbf{x} \in E \quad (\|\mathbf{x}\|_E \leqslant \eta_1 \implies \|f(\mathbf{x})\|_F \leqslant 1). \tag{14}$$

Montrer que f est bornée sur la boule fermée $\overline{B}(\mathbf{0}_E, 1) = \{\mathbf{x} \in E \ ; \ \|\mathbf{x}\|_E \leqslant 1\}$ revient à montrer que[46]

$$\exists M \in \mathbb{R}^+ \quad \forall \mathbf{x}' \in \overline{B}(\mathbf{0}_E, 1) \quad \|f(\mathbf{x}')\|_F \leqslant M. \tag{15}$$

Soit $\mathbf{x}'$ un élément de $\overline{B}(\mathbf{0}_E, 1)$; on a $\|\eta_1 \mathbf{x}'\|_E = \eta_1 \|\mathbf{x}'\|_E \leqslant \eta_1$. On déduit de l'assertion (14) que $\|f(\eta_1 \mathbf{x}')\|_F \leqslant 1$. Comme f est une application linéaire,

$$\|f(\eta_1 \mathbf{x}')\|_F = \|\eta_1 \ f(\mathbf{x}')\|_F = |\eta_1| \ \|f(\mathbf{x}')\|_F = \eta_1 \ \|f(\mathbf{x}')\|_F.$$

On en déduit que $\|f(\mathbf{x}')\|_F \leqslant 1/\eta_1$. L'assertion (15) est démontrée (il suffit de prendre $M = 1/\eta_1$).

[46] Voir la définition 9.4 p. 546.

$\trianglerighteq$ Supposons maintenant que f est bornée par $M \in \mathbb{R}^+$ sur la boule fermée $\overline{B}(\mathbf{0}_E, 1)$ et montrons qu'il existe un réel positif C tel que pour tout $\mathbf{x} \in E$ on ait $\|f(\mathbf{x})\|_F \leqslant C\|\mathbf{x}\|_E$. Remarquons que si $\mathbf{x} = \mathbf{0}_E$ alors la relation est vraie puisque f étant linéaire $f(\mathbf{0}_E) = \mathbf{0}_F$. Considérons un élément $\mathbf{x}$ de E qui est non nul ; l'élément $\mathbf{x}' = \mathbf{x}/\|\mathbf{x}\|_E$ appartient à la boule fermée $\overline{B}(\mathbf{0}_E, 1)$ puisque $\|\mathbf{x}'\|_E = 1$. Par hypothèse, on a $\|f(\mathbf{x}')\|_F \leqslant M$, voir l'assertion (15). Or, puisque f est linéaire,

$$\|f(\mathbf{x}')\|_F = \left\| \frac{1}{\|\mathbf{x}\|_E} f(\mathbf{x}) \right\|_F = \frac{1}{\|\mathbf{x}\|_E} \|f(\mathbf{x})\|_F.$$

On en déduit que $\|f(\mathbf{x})\|_F \leqslant M\|\mathbf{x}\|_E$. La quatrième assertion du théorème 9.3 est démontrée.

$\trianglerighteq$ Supposons qu'il existe un réel positif C tel que pour tout $\mathbf{x} \in E$ on ait $\|f(\mathbf{x})\|_F \leqslant C\|\mathbf{x}\|_E$ et montrons que f est continue sur E. Soit $(\mathbf{x}', \mathbf{x}'') \in E^2$; en utilisant la linéarité de f et l'hypothèse, on obtient :

$$\|f(\mathbf{x}') - f(\mathbf{x}'')\|_F = \|f(\mathbf{x}' - \mathbf{x}'')\|_F \leqslant C\|\mathbf{x}' - \mathbf{x}''\|_E.$$

Cette relation indique que f est lipschitzienne sur E de rapport C. La proposition 9.16 permet de conclure que f est uniformément continue sur E et par conséquent continue sur E. □

On sera attentif au fait que les différentes manières d'établir qu'une application linéaire est continue données par le théorème 9.3 ne sont valables que pour les applications qui sont effectivement linéaires. La démonstration du théorème 9.3 permet de bien comprendre l'importance du caractère linéaire de l'application pour que les équivalences indiquées soient vraies. Il ne saurait être question d'utiliser la même approche pour montrer la continuité d'une application qui n'est pas linéaire.

Remarque Dans l'énoncé du théorème 9.3, on peut remplacer dans l'assertion (2) l'élément $\mathbf{0}_E$ par n'importe quel autre élément de E. Dans l'assertion (3), on peut remplacer la boule fermée de rayon 1 par n'importe quelle autre boule de rayon non nul. ◇

En complément au théorème 9.3, l'exercice suivant permet d'établir d'autres propriétés des applications linéaires.

Exercice 9.5 Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux $\mathbb{K}$-espaces vectoriels normés et f une application linéaire de E dans F. Montrer que les assertions suivantes sont équivalentes :

1. f est continue sur E ;
2. f est lipschitzienne sur E ;
3. f est uniformément continue sur E.

Exemples

1. Pour $n \in \mathbb{N}^*$ fixé, montrons que l'application trace, qui à une matrice A de $\mathcal{M}_n(\mathbb{K})$ associe la somme de ses éléments diagonaux, est une forme linéaire continue. Munissons $\mathcal{M}_n(\mathbb{K})$ de la norme de Schur[47] ; en utilisant l'inégalité de Cauchy-Schwarz, on obtient pour tout $A = (a_{ij})_{(i,j)\in\{1,\dots,n\}^2}$:

$$\begin{aligned} |\mathrm{Tr}(A)| &= \left|\sum_{i=1}^{n} a_{ii}\right| \leqslant \sum_{i=1}^{n} |a_{ii}| \leqslant \sqrt{n}\left(\sum_{i=1}^{n} |a_{ii}|^2\right)^{\frac{1}{2}} \\ &\leqslant \sqrt{n}\left(\sum_{i=1}^{n}\sum_{j=1}^{n} |a_{ij}|^2\right)^{\frac{1}{2}} = \sqrt{n}\,\|A\|_s. \end{aligned}$$

D'après l'assertion (4) du théorème 9.3, on en conclut que l'application trace est une application linéaire continue de $(\mathcal{M}_n(\mathbb{K}), \|\cdot\|_s)$ dans $(\mathbb{R}, |\cdot|)$.

2. Considérons l'application Φ qui à une application f continue sur $[a,b]$ à valeurs réelles associe sont intégrale indéfinie $F : x \in [a,b] \mapsto \int_a^x f(t)\,\mathrm{d}t$. Cette application est linéaire car pour tous $f_1, f_2 \in \mathcal{C}^0([a,b],\mathbb{R})$, pour tout $\lambda_1, \lambda_2 \in \mathbb{R}$ et pour tout $x \in [a,b]$ on a

$$\begin{aligned} \Phi(\lambda_1 f_1 + \lambda_2 f_2)(x) &= \int_a^x (\lambda_1 f_1 + \lambda_2 f_2)(t)\,\mathrm{d}t \\ &= \lambda_1 \int_a^x f_1(t)\,\mathrm{d}t + \lambda_2 \int_a^x f_2(t)\,\mathrm{d}t = \lambda_1 \Phi(f_1)(x) + \lambda_2 \Phi(f_2)(x). \end{aligned}$$

Par ailleurs, si f est continue sur $[a,b]$ alors F est également continue[48] sur $[a,b]$; Φ est donc un endomorphisme de $\mathcal{C}^0([a,b],\mathbb{R})$. Munissons $\mathcal{C}^0([a,b],\mathbb{R})$ de la norme infinie définie pour $f \in \mathcal{C}^0([a,b],\mathbb{R})$ par $\|f\|_\infty = \sup_{x\in[a,b]} |f(x)|$, et montrons que Φ est une application continue de $(\mathcal{C}^0([a,b],\mathbb{R}), \|\cdot\|_\infty)$ dans $(\mathcal{C}^0([a,b],\mathbb{R}), \|\cdot\|_\infty)$. Pour tout $x \in [a,b]$, on a

$$|\Phi(f)(x)| = \left|\int_a^x f(t)\,\mathrm{d}t\right| \leqslant \int_a^x |f(t)|\,\mathrm{d}t \leqslant \int_a^x \Big(\sup_{s\in[a,b]} |f(s)|\Big)\,\mathrm{d}t = (b-a)\|f\|_\infty.$$

On en déduit que $\|\Phi(f)\|_\infty = \sup_{x\in[a,b]} |\Phi(f)(x)| \leqslant (b-a)\|f\|_\infty$, ce qui d'après l'assertion (4) du théorème 9.3 permet d'affirmer que Φ est continue.

---

[47] Voir la proposition 7.5 p. 365. Pour tout $A = (a_{ij})_{(i,j)\in\{1,\dots,n\}^2}$, la norme de Schur de A est $\|A\|_s = \sqrt{\sum_{i=1}^n \sum_{j=1}^n |a_{ij}|^2}$. Rappelons que l'espace vectoriel $\mathcal{M}_n(\mathbb{K})$ étant de dimension finie, toutes les normes y sont équivalentes ; on choisit celle qui semble la plus appropriée au calcul.

[48] Voir la proposition 18.7 p. 847 du *Cours de première année*.

Exercice 9.6
1 - Montrer que l'application f qui à un polynôme P de $\mathbb{K}[X]$ associe son polynôme dérivé P' est une application linéaire.
2 - Montrer que l'application N qui à un polynôme $P = \sum_{k=0}^{n} a_k X^k$ de $\mathbb{K}[X]$ associe le réel $N(P) = |a_0| + \max_{k\in\{1,\dots,n\}} |k\, a_k|$ est une norme sur $\mathbb{K}[X]$.
3 - En utilisant le théorème 9.3, montrer que f est une application continue de $\mathbb{K}[X]$ muni de la norme N dans $\mathbb{K}[X]$ muni de la norme $\|\cdot\|_\infty$ définie pour tout $P = \sum_{k=0}^{n} a_k X^k$ par $\|P\|_\infty = \max_{k\in\{0,\dots,n\}} |a_k|$, voir p. 350.

Remarque Pour montrer qu'une application linéaire f de $(E, \|\cdot\|_E)$ dans $(F, \|\cdot\|_F)$ n'est pas continue, il suffit de montrer que l'assertion suivante, qui constitue la négation de la quatrième assertion du théorème 9.3, est vraie :

$$\forall C \in \mathbb{R}_+ \quad \exists \mathbf{x} \in E \quad \|f(\mathbf{x})\|_F > C\, \|\mathbf{x}\|_E.$$

L'exemple suivant illustre la remarque. ◇

Exemple L'exercice 9.6 a permis d'établir que l'application f qui à un polynôme P de $\mathbb{K}[X]$ associe son polynôme dérivé P' est une application linéaire continue de $(\mathbb{K}(X), N)$ dans $(\mathbb{K}(X), \|\cdot\|_\infty)$. Montrons que f n'est pas continue de $(\mathbb{K}(X), \|\cdot\|_\infty)$ dans $(\mathbb{K}(X), \|\cdot\|_\infty)$. Soient C un réel positif et $n = \mathrm{E}(C) + 1$; le polynôme $P = X^n$ est tel que $\|P\|_\infty = 1$ et $P' = n\, X^{n-1}$ de sorte que

$$\|f(P)\|_\infty = \|P'\|_\infty = n = \mathrm{E}(C) + 1 > C = C\, \|P\|_\infty.$$

L'application f n'est donc pas une application linéaire continue de $\mathbb{K}(X)$ dans $\mathbb{K}[X]$ lorsque l'espace vectoriel de départ et d'arrivée sont munis de la norme $\|\cdot\|_\infty$ car pour tout réel C positif, le polynôme $P = X^{\mathrm{E}(C)+1}$ est tel que $\|f(P)\|_\infty > C\, \|P\|_\infty$.

La proposition suivante précise la structure de l'ensemble des applications linéaires continues.

Proposition 9.23 *Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux $\mathbb{K}$-espaces vectoriels normés. L'ensemble $\mathcal{L}_c(E, F)$ des applications linéaires de E dans F qui sont continues est un sous-espace vectoriel de $\mathcal{L}(E, F)$. De plus, l'application*

$$f \in \mathcal{L}_c(E, F) \longmapsto \sup_{\substack{\mathbf{x}\in E \\ \mathbf{x}\neq 0}} \frac{\|f(\mathbf{x})\|_F}{\|\mathbf{x}\|_E}$$

est une norme sur $\mathcal{L}_c(E, F)$ appelée norme linéaire *et notée $|||\cdot|||$.*

Démonstration ⊵ L'ensemble $\mathcal{L}_c(E,F)$ des applications linéaires de E dans F continues est non vide car l'application nulle $\mathbf{x} \in E \mapsto \mathbf{0}_F \in F$ est une application linéaire continue. Pour montrer que l'ensemble $\mathcal{L}_c(E,F)$ est un sous-espace vectoriel de $\mathcal{L}(E,F)$, il faut montrer[49] que :

$$\forall(\alpha,\beta) \in \mathbb{K}^2 \quad \forall(f,g) \in \mathcal{L}_c(E,F) \qquad \alpha \cdot f + \beta \cdot g \in \mathcal{L}_c(E,F).$$

Soient $(\alpha,\beta) \in \mathbb{K}^2$ et $(f,g) \in \mathcal{L}_c(E,F)$; l'application $h = \alpha \cdot f + \beta \cdot g$ est une application linéaire car $\mathcal{L}(E,F)$ est un espace vectoriel. Il s'agit donc de montrer que cette application linéaire h est continue sur E. Par hypothèse, f et g sont continues sur E donc d'après l'assertion (4) du théorème 9.3,

- il existe $C_1 \in \mathbb{R}^+$ tel que pour tout $\mathbf{x} \in E$: $\|f(\mathbf{x})\|_F \leqslant C_1 \|\mathbf{x}\|_E$;
- il existe $C_2 \in \mathbb{R}^+$ tel que pour tout $\mathbf{x} \in E$: $\|g(\mathbf{x})\|_F \leqslant C_2 \|\mathbf{x}\|_E$.

On en déduit que pour tout $\mathbf{x} \in E$,

$$\|h(\mathbf{x})\|_F = |\alpha| \, \|f(\mathbf{x})\|_F + |\beta| \, \|g(\mathbf{x})\|_F \leqslant (|\alpha| C_1 + |\beta| C_2) \, \|\mathbf{x}\|_E.$$

On a ainsi montré qu'il existe $C \in \mathbb{R}^+$ (prendre $C = |\alpha| C_1 + |\beta| C_2$) tel que pour tout $\mathbf{x} \in E$: $\|h(\mathbf{x})\|_F \leqslant C \|\mathbf{x}\|_E$. D'après le théorème 9.3 ceci implique que h est continue sur E.

⊵ Montrons que l'application $N : f \in \mathcal{L}_c(E,F) \longmapsto \sup_{\substack{\mathbf{x} \in E \\ \mathbf{x} \neq 0}} \frac{\|f(\mathbf{x})\|_F}{\|\mathbf{x}\|_E}$ est une norme sur $\mathcal{L}_c(E,F)$.

▷ Notons tout d'abord que cette application est bien définie car d'après le théorème 9.3, pour toute application $f \in \mathcal{L}_c(E,F)$ l'ensemble

$$U = \left\{ \frac{\|f(\mathbf{x})\|_F}{\|\mathbf{x}\|_E} \; ; \; \mathbf{x} \in E \quad \mathbf{x} \neq 0 \right\}$$

est borné ; de plus, cette application est à valeurs positives car les éléments de l'ensemble U sont des réels positifs.

Vérifions que l'application N satisfait aux quatre conditions de la définition 7.3.
▷ Les propriétés de la borne supérieure impliquent que $N(f) = 0$ si et seulement si pour tout $\mathbf{x} \in E$ avec $\mathbf{x} \neq \mathbf{0}_E$ on a $\|f(\mathbf{x})\|_F = 0$. Or $\|f(\mathbf{x})\|_F = 0$ équivaut à $f(\mathbf{x}) = \mathbf{0}_F$; comme par ailleurs $f(\mathbf{0}_E) = \mathbf{0}_F$, on en déduit que $N(f) = 0$ si et seulement si pour tout $\mathbf{x} \in E$ on a $f(\mathbf{x}) = \mathbf{0}_F$, autrement dit si f est l'application nulle.

▷ Soient $f \in \mathcal{L}_c(E,F)$ et $\lambda \in \mathbb{K}$; pour tout $\mathbf{x} \in E$ avec $\mathbf{x} \neq \mathbf{0}_E$ on a

$$\frac{\|\lambda f(\mathbf{x})\|_F}{\|\mathbf{x}\|_E} = |\lambda| \, \frac{\|f(\mathbf{x})\|_F}{\|\mathbf{x}\|_E}.$$

Deux fonctions égales sur un ensemble ayant même borne supérieure sur cet ensemble, on en déduit que

$$N(\lambda f) = \sup_{\substack{\mathbf{x} \in E \\ \mathbf{x} \neq 0}} \frac{\|\lambda f(\mathbf{x})\|_F}{\|\mathbf{x}\|_E} = |\lambda| \sup_{\substack{\mathbf{x} \in E \\ \mathbf{x} \neq 0}} \frac{\|f(\mathbf{x})\|_F}{\|\mathbf{x}\|_E} = |\lambda| \, N(f).$$

[49] Voir la définition 8.5 p. 306 du *Cours de première année*.

▷ Pour tout $(f,g) \in \mathcal{L}_c(E,F)^2$ et pour tout $\mathbf{x} \in E$ avec $\mathbf{x} \neq \mathbf{0}_E$, on a $\|(f+g)(\mathbf{x})\|_F \leqslant \|f(\mathbf{x})\|_F + \|g(\mathbf{x})\|_F$. Par conséquent,

$$\frac{\|(f+g)(\mathbf{x})\|_F}{\|\mathbf{x}\|_E} \leqslant \frac{\|f(\mathbf{x})\|_F}{\|\mathbf{x}\|_E} + \frac{\|g(\mathbf{x})\|_F}{\|\mathbf{x}\|_E} \leqslant \sup_{\substack{\mathbf{x}' \in E \\ \mathbf{x}' \neq 0}} \frac{\|f(\mathbf{x}')\|_F}{\|\mathbf{x}'\|_E} + \sup_{\substack{\mathbf{x}' \in E \\ \mathbf{x}' \neq 0}} \frac{\|g(\mathbf{x}')\|_F}{\|\mathbf{x}'\|_E}$$
$$= N(f) + N(g).$$

On en déduit que :

$$N(f+g) = \sup_{\substack{\mathbf{x} \in E \\ \mathbf{x} \neq 0}} \frac{\|(f+g)(\mathbf{x})\|_F}{\|\mathbf{x}\|_E} \leqslant N(f) + N(g).$$

On en conclut que l'application N est bien une norme sur $\mathcal{L}_c(E,F)$. □

Remarques

1. Pour tout $\mathbf{x} \in E$ on a $|||f||| = \sup_{\substack{\mathbf{x}' \in E \\ \mathbf{x}' \neq 0}} \frac{\|f(\mathbf{x}')\|_F}{\|\mathbf{x}'\|_E} \geqslant \frac{\|f(\mathbf{x})\|_F}{\|\mathbf{x}\|_E}$ et par conséquent : $\|f(\mathbf{x})\|_F \leqslant |||f|||\ \|\mathbf{x}\|_E$.

2. En exploitant les propriétés de la norme et le fait que f est une application linéaire, on obtient :

$$|||f||| = \sup_{\substack{\mathbf{x} \in E \\ \mathbf{x} \neq 0}} \frac{\|f(\mathbf{x})\|_F}{\|\mathbf{x}\|_E} = \sup_{\substack{\mathbf{x} \in E \\ \mathbf{x} \neq 0}} \|f(\mathbf{x}/\|\mathbf{x}\|_E)\|_F = \sup_{\substack{\mathbf{x}' \in E \\ \|\mathbf{x}'\| = 1}} \|f(\mathbf{x}')\|_F.$$

Par ailleurs, comme $\{\mathbf{x} \in \overline{B}(0,1)\ ;\ \mathbf{x} \neq 0\} \subset \{\mathbf{x} \in E\ ;\ \mathbf{x} \neq 0\}$ et que pour $\mathbf{x} \in \overline{B}(0,1)$ on a $1/\|\mathbf{x}\|_E \geqslant 1$ et par conséquent,

$$|||f||| = \sup_{\substack{\mathbf{x} \in E \\ \mathbf{x} \neq 0}} \frac{\|f(\mathbf{x})\|_F}{\|\mathbf{x}\|_E} \geqslant \sup_{\substack{\mathbf{x} \in \overline{B}(0,1) \\ \mathbf{x} \neq 0}} \frac{\|f(\mathbf{x})\|_F}{\|\mathbf{x}\|_E} \geqslant \sup_{\substack{\mathbf{x} \in \overline{B}(0,1) \\ \mathbf{x} \neq 0}} \|f(\mathbf{x})\|_F = \sup_{\mathbf{x} \in \overline{B}(0,1)} \|f(\mathbf{x})\|_F.$$

De plus, comme pour tout $\mathbf{x} \in E$ avec $\mathbf{x} \neq \mathbf{0}_E$, le vecteur $\mathbf{y} = \mathbf{x}/\|\mathbf{x}\|_E$ est de norme 1, on a $\mathbf{y} \in \overline{B}(\mathbf{0}_E, 1)$ et

$$\frac{\|f(\mathbf{x})\|_F}{\|\mathbf{x}\|_E} = \|f(\mathbf{x}/\|\mathbf{x}\|_E)\|_F = \|f(\mathbf{y})\|_F \leqslant \sup_{\mathbf{x}' \in \overline{B}(0,1)} \|f(\mathbf{x}')\|_F$$

ce qui entraîne que $\sup_{\substack{\mathbf{x} \in E \\ \mathbf{x} \neq 0}} \frac{\|f(\mathbf{x})\|_F}{\|\mathbf{x}\|_E} \leqslant \sup_{\mathbf{x} \in \overline{B}(0,1)} \|f(\mathbf{x})\|_F$ et finalement que $|||f||| = \sup_{\mathbf{x} \in \overline{B}(0,1)} \|f(\mathbf{x})\|_F$. On a donc

$$\boxed{|||f||| = \sup_{\substack{\mathbf{x} \in E \\ \mathbf{x} \neq 0}} \frac{\|f(\mathbf{x})\|_F}{\|\mathbf{x}\|_E} = \sup_{\mathbf{x} \in \overline{B}(0,1)} \|f(\mathbf{x})\|_F = \sup_{\substack{\mathbf{x}' \in E \\ \|\mathbf{x}'\| = 1}} \|f(\mathbf{x}')\|_F}$$

3. L'endomorphisme identité de $(E, \|\cdot\|_E)$ dans $(E, \|\cdot\|_E)$ est de norme 1 puisque pour tout $\mathbf{x} \in E$ on a $\|\mathrm{id}(\mathbf{x})\|_E = \|\mathbf{x}\|_E$. ◇

Exemples

1. Nous avons établi que pour $n \in \mathbb{N}^*$ fixé, l'application trace qui à une matrice de $\mathcal{M}_n(\mathbb{K})$ associe la somme de ses éléments diagonaux, est une forme linéaire continue et que pour tout $A = (a_{ij})_{i,j}$ on a $|\mathrm{Tr}(A)| \leqslant \sqrt{n}\, \|A\|_s$. On en déduit que si $\mathcal{M}_n(\mathbb{K})$ est muni de la norme de Schur, alors

$$|||\mathrm{Tr}||| = \sup_{\substack{A \in \mathcal{M}_n(\mathbb{K}) \\ A \neq 0}} \frac{|\mathrm{Tr}(A)|}{\|A\|_s} \leqslant \sqrt{n}.$$

Par ailleurs, la trace de la matrice identité I valant n et sa norme de Schur valant $\sqrt{n}$, on a

$$\sqrt{n} = \frac{|\mathrm{Tr}(I)|}{\|I\|_s} \leqslant \sup_{\substack{A \in \mathcal{M}_n(\mathbb{K}) \\ A \neq 0}} \frac{|\mathrm{Tr}(A)|}{\|A\|_s} = |||\mathrm{Tr}|||,$$

et on en conclut que $|||\mathrm{Tr}||| = \sqrt{n}$.

2. Notons E l'espace vectoriel $\mathcal{C}^0([a,b],\mathbb{R})$ et considérons l'endomorphisme Φ de E qui à une application associe son intégrale indéfinie. On a établi p. 576 que Φ est une application continue de $(E, \|\cdot\|_\infty)$ dans $(E, \|\cdot\|_\infty)$ et que pour tout $f \in E$, on a $\|\Phi(f)\|_\infty \leqslant (b-a)\|f\|_\infty$. On en déduit que

$$|||\Phi||| = \sup_{\substack{f \in E \\ f \neq 0}} \frac{\|\Phi(f)\|_\infty}{\|f\|_\infty} \leqslant (b-a).$$

Soit g l'application constante égale à 1 sur $[a,b]$; on a $\|g\|_\infty = 1$ et

$$\|\Phi(g)\|_\infty = \sup_{x \in [a,b]} \left| \int_a^x 1 \ \mathrm{d}x \right| = \sup_{x \in [a,b]} |x-a| = b-a.$$

On en déduit que

$$b-a = \frac{\|\Phi(g)\|_\infty}{\|g\|_\infty} \leqslant \sup_{\substack{f \in E \\ f \neq 0}} \frac{\|\Phi(f)\|_\infty}{\|f\|_\infty} = |||\Phi|||$$

et par conséquent que $|||\Phi||| = b-a$ pour le choix de la norme de la convergence uniforme sur E.

Exercice 9.7 Soit Φ l'endomorphisme de $\mathcal{C}^0([a,b],\mathbb{R})$ qui à une application f de $\mathcal{C}^0([a,b],\mathbb{R})$ associe son intégrale indéfinie. On munit l'espace vectoriel $\mathcal{C}^0([a,b],\mathbb{R})$ de la norme de la convergence en moyenne $f \in \mathcal{C}^0([a,b],\mathbb{R}) \mapsto \int_a^b |f(x)|\,\mathrm{d}x$, notée $\|\cdot\|_1$.

1 - Montrer que Φ est continue.

2 - Pour $n \in \mathbb{N}$ fixé, on note $f_n : x \in [a,b] \mapsto n^2 \mathrm{e}^{n(a-x)}$.

a) Calculer $\|f_n\|_1$ et $\|\Phi(f_n)\|_1$.

b) En déduire que la norme linéaire de Φ vaut $b-a$ pour le choix de la norme de la convergence en moyenne sur $\mathcal{C}^0([a,b],\mathbb{R})$.

Noyau d'une application linéaire continue

Si f est une application linéaire continue d'un espace vectoriel normé $(E, \|\cdot\|_E)$ dans un espace vectoriel normé $(F, \|\cdot\|_F)$ alors son noyau est un sous-espace vectoriel fermé de E car par définition $\mathrm{Ker}(f) = \{x \in E \ ; \ f(x) = 0_F\}$ est l'image réciproque par f de $\{0_F\}$ qui est fermé[50] dans F, voir la proposition 9.17.

Dans la cas où $F = \mathbb{R}$, on peut montrer que l'on a une équivalence : une forme linéaire sur E est continue si et seulement si son noyau est fermé dans E.

Exercice 9.8 Soit $(E, \|\cdot\|_E)$ un $\mathbb{K}$-espace vectoriel normé et p un endomorphisme de E, continu sur E et vérifiant $p \circ p = p$ (p est un projecteur). Montrer que le noyau et l'image de p sont des fermés de E.

Interprétation géométrique de $|||f|||$

Soit f une forme linéaire (non nulle) continue sur un $\mathbb{K}$-espace vectoriel normé $(E, \|\cdot\|_E)$; le réel $|||f|||$ correspond à l'inverse de la distance de l'origine $\mathbf{0} \in E$ à l'hyperplan d'équation $f(\mathbf{x}) = 1$ pour la distance subordonnée à la norme $\|\cdot\|_E$.

Pour établir cette propriété, notons $H = \{\mathbf{x} \in E \ ; \ f(\mathbf{x}) = 1\}$; d'après la définition 7.14 p. 386, la distance de $\mathbf{0} \in E$ à H est :

$$d(\mathbf{0}, H) = \inf \{\|\mathbf{y}\|_E \ ; \ \mathbf{y} \in H\}.$$

Commençons par remarquer que pour tout $\mathbf{y} \in H$ on a $f(\mathbf{y}) = 1$ et par conséquent que

$$\frac{1}{\|\mathbf{y}\|_E} = \frac{|f(\mathbf{y})|}{\|\mathbf{y}\|_E} \leqslant \sup_{\substack{\mathbf{x} \in E \\ \mathbf{x} \neq 0}} \frac{|f(\mathbf{x})|}{\|\mathbf{x}\|_E} = |||f|||.$$

On en déduit que pour tout $\mathbf{y} \in H$ on a $\|\mathbf{y}\|_E \geqslant 1/|||f|||$ et par conséquent que que $d(\mathbf{0}, H) \geqslant 1/|||f|||$. Comme

$$|||f||| = \sup_{\substack{\mathbf{x} \in E \\ \mathbf{x} \neq 0}} |f(\mathbf{x})|/\|\mathbf{x}\|_E = \sup \{|f(\mathbf{x})|/\|\mathbf{x}\|_E \ ; \ \mathbf{x} \in E \ \mathbf{x} \neq 0\},$$

on en déduit des propriétés de la borne supérieure[51] que pour tout $\varepsilon \in \mathbb{R}_+^*$, il existe $\mathbf{x} \in E \setminus \{\mathbf{0}\}$ tel que

$$|||f||| - \varepsilon \leqslant \frac{|f(\mathbf{x})|}{\|\mathbf{x}\|_E} \leqslant |||f|||.$$

Puisque f n'est pas la forme linéaire nulle, on a $|||f||| \neq 0$; on en déduit que pour tout réel ε strictement positif tel que $|||f||| - \varepsilon > 0$, il existe $\mathbf{x} \in E \setminus \{\mathbf{0}\}$ avec $f(\mathbf{x}) \neq 0$ et

$$\frac{1}{|||f|||} \leqslant \frac{\|\mathbf{x}\|_E}{|f(\mathbf{x})|} \leqslant \frac{1}{|||f||| - \varepsilon}.$$

[50] Cette affirmation a été établie p. 372.

[51] Voir la proposition 3.2 p. 95 du *Cours de première année*.

Le vecteur $\mathbf{y} = \mathbf{x}/f(\mathbf{x})$ appartient à H car f étant une forme linéaire, on a $f(\mathbf{y}) = f(\mathbf{x}/f(\mathbf{x})) = \frac{1}{f(\mathbf{x})}\, f(\mathbf{x}) = 1$. Par conséquent

$$d(\mathbf{0}_E, H) \leqslant d(\mathbf{0}_E, \mathbf{y}) = \|\mathbf{y}\|_E = \frac{\|\mathbf{x}\|_E}{|f(\mathbf{x})|} \leqslant \frac{1}{|||f|||\,-\varepsilon}.$$

Cette inégalité étant vraie pour tout $\varepsilon \in \mathbb{R}_+^*$ tel que $\varepsilon < |||f|||$, on en déduit en considérant la limite à droite lorsque ε tend vers 0 que :

$$d(\mathbf{0}_E, H) = \lim_{\varepsilon \to 0^+} d(\mathbf{0}_E, H) \leqslant \lim_{\varepsilon \to 0^+} \frac{1}{|||f|||\,-\varepsilon} = \frac{1}{|||f|||}.$$

Puisque l'on a simultanément $d(\mathbf{0}, H) \leqslant 1/\,|||\,f|||$ et $d(\mathbf{0}, H) \geqslant 1/\,|||\,f|||$, on en conclut que $d(\mathbf{0}, H) = 1/\,|||\,f|||$.

Considérons la forme linéaire $f : (x_1, x_2, x_3) \in \mathbb{R}^3 \mapsto ax_1 + bx_2 + cx_3 \in \mathbb{R}$ où a, b, c désignent trois réels. Dans $\mathbb{R}^3$, la distance du plan d'équation $ax_1 + bx_2 + cx_3 = 1$ à l'origine valant[52] $1/\sqrt{a^2+b^2+c^2}$, on en déduit que $|||f||| = \sqrt{a^2+b^2+c^2}$.

Exercice 9.9 ***Complétude de*** $\mathcal{L}_c(E)$
On désigne par $(E, \|\cdot\|)$ un espace de Banach et par $\mathcal{L}_c(E)$ l'espace vectoriel des endomorphismes de E qui sont continus, muni de la norme linéaire $|||\ |||$.
Soit $(f_n)_n$ une suite de Cauchy dans $(\mathcal{L}_c(E), |||\ |||)$.
1 - Justifier que pour tout $\mathbf{x} \in E$, la suite $(f_n(\mathbf{x}))_n$ converge vers un élément $f(\mathbf{x})$ de E. On note f l'application $\mathbf{x} \in E \mapsto f(\mathbf{x}) \in E$.
2 - Montrer que f est un endomorphisme de E et qu'il est continu.
3 - Montrer que la suite $(f_n)_n$ converge vers f dans $(\mathcal{L}_c(E), |||\ |||)$ et conclure.

THÉORÈME 9.4 *Soient (F, N_F) un $\mathbb{K}$-espace vectoriel normé et (E, N_E) un $\mathbb{K}$-espace vectoriel normé de* dimension finie. *Toute application linéaire de E dans F est continue.*

Démonstration Notons $\mathcal{B} = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ et considérons une application linéaire f de E dans F. Pour montrer que f est continue sur E, d'après le théorème 9.3, il faut et il suffit de montrer que f est continue en $\mathbf{0}_E$, c'est-à-dire montrer que $\lim_{\mathbf{x} \to \mathbf{0}_E} \|f(\mathbf{x})\|_F = 0$ car $f(\mathbf{0}_E) = \mathbf{0}_F$.

Soit $\mathbf{x}$ un vecteur de E de composantes $(x_1, \ldots, x_n)$ dans la base $\mathcal{B}$; on a $\mathbf{x} = \sum_{k=1}^n x_k\, \mathbf{e}_k$ et f étant linéaire, $f(\mathbf{x}) = \sum_{k=1}^n x_k\, f(\mathbf{e}_k)$. On en déduit, en utilisant les propriétés de la norme, que

$$\|f(\mathbf{x})\|_F = \Big\| \sum_{k=1}^n x_k\, f(\mathbf{e}_k) \Big\|_F \leqslant \sum_{k=1}^n \|x_k\, f(\mathbf{e}_k)\|_F = \sum_{k=1}^n |x_k|\, \|f(\mathbf{e}_k)\|_F.$$

[52] Voir p. 386.

Pour tout $k \in \{1, \ldots, n\}$, on a $|x_k| \leqslant \max_{1 \leqslant i \leqslant n} |x_i| = \|\mathbf{x}\|_\infty$ d'où

$$0 \leqslant \|f(\mathbf{x})\|_F \leqslant C \|\mathbf{x}\|_\infty \quad \text{où} \quad C = \sum_{k=1}^{n} \|f(\mathbf{e}_k)\|_F.$$

Toute norme sur E étant équivalente[53] à la norme $\|\cdot\|_\infty$, on en déduit par le théorème d'encadrement que $\lim_{\mathbf{x} \to \mathbf{0}_E} \|f(\mathbf{x})\|_F = 0$. □

On sera attentif au fait que le théorème 9.4 n'est valide que dans le cas où E un $\mathbb{K}$-espace vectoriel de *dimension finie*. Il existe des applications linéaires qui ne sont pas continues. C'est le cas, voir l'exercice 9.12, de la forme linéaire $f : h \in \mathcal{C}^0([0,1], \mathbb{R}) \longmapsto h(1)$ qui n'est pas continue sur $\mathcal{C}^0([0,1], \mathbb{R})$ muni de la norme $\|\cdot\|_1$ définie par $\|h\|_1 = \int_0^1 |h(x)| \, \mathrm{d}x$.

Continuité des applications bilinéaires

En complément au théorème 9.3 concernant la continuité des applications linéaires, le théorème suivant permet de caractériser la continuité des applications bilinéaires (voir la définition 8.1 p. 438). La vérification de ces propriétés reprend le principe de la démonstration du théorème 9.3 et est laissée en exercice.

THÉORÈME 9.5 *Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux $\mathbb{R}$-espaces vectoriels normés et f une* application bilinéaire *de $E \times E$ dans F. Les assertions suivantes sont équivalentes :*

1. *l'application f est continue sur $E \times E$;*
2. *l'application f est continue en $(\mathbf{0}_E, \mathbf{0}_E)$ (où $\mathbf{0}_E$ est le vecteur nul dans E) ;*
3. *l'application f est bornée sur $\overline{B}(\mathbf{0}_E, 1) \times \overline{B}(\mathbf{0}_E, 1)$ (où $\overline{B}(\mathbf{0}_E, 1)$ désigne la boule fermée de centre $\mathbf{0}_E$ et de rayon unité dans E) ;*
4. *il existe un réel positif C tel que pour tout $(\mathbf{x}, \mathbf{y}) \in E^2$:* $\|f(\mathbf{x}, \mathbf{y})\|_F \leqslant C \|\mathbf{x}\|_E \, \|\mathbf{y}\|_E$.

Exemples

1. Soit E un espace préhilbertien (*i.e.* un $\mathbb{R}$-espace vectoriel muni d'un produit scalaire φ). La forme bilinéaire symétrique f qui à deux vecteurs $\mathbf{x}$ et $\mathbf{y}$ de E associe le produit scalaire $\varphi(\mathbf{x}, \mathbf{y})$ est continue sur $E \times E$ lorsque E est muni de la norme $\|\cdot\|$ associée au produit scalaire[54]. En effet, d'après l'inégalité de Cauchy-Schwarz[55], pour tous $\mathbf{x}$ et $\mathbf{y}$ dans E on a

$$|f(\mathbf{x}, \mathbf{y})| = |\varphi(\mathbf{x}, \mathbf{y})| \leqslant \sqrt{\varphi(\mathbf{x}, \mathbf{x}) \, \varphi(\mathbf{y}, \mathbf{y})} = \|\mathbf{x}\| \cdot \|\mathbf{y}\|.$$

[53] Voir le corollaire 7.1 p. 368.
[54] Voir la proposition 8.19 p. 482.
[55] Voir le théorème 8.5 p. 481.

La quatrième assertion du théorème 9.5 est établie et permet de conclure que f est continue sur $E \times E$.

2. Dans l'algèbre $\mathbb{K}[X]$ des polynômes à coefficients dans $\mathbb{K}$, considérons l'application bilinéaire symétrique $f : (P,Q) \in \mathbb{K}[X] \times \mathbb{K}[X] \mapsto P \times Q \in \mathbb{K}[X]$ qui à deux polynômes associe leur produit. Munissons $\mathbb{K}[X]$ de la norme $\|\cdot\|_1$ définie pour tout $P = \sum_{k=0}^{p} a_k X^k$ par $\|P\|_1 = \sum_{k=0}^{p} |a_k|$. Pour tous polynômes $P = \sum_{k=0}^{p} a_k X^k$ et $Q = \sum_{k=0}^{q} b_k X^k$ de $\mathbb{K}[X]$, on a $P \times Q = \sum_{k=0}^{p+q} c_k X^k$ où $c_k = \sum_{\ell=0}^{k} a_\ell\, b_{k-\ell}$ et

$$\begin{aligned}\|f(P,Q)\|_1 &= \|P \times Q\|_1 = \sum_{k=0}^{p+q} |c_k| = \sum_{k=0}^{p+q} \left| \sum_{\ell=0}^{k} a_\ell\, b_{k-\ell} \right| \\ &\leqslant \sum_{k=0}^{p+q} \sum_{\ell=0}^{k} |a_\ell\, b_{k-\ell}| = \sum_{i=0}^{p} \sum_{j=0}^{q} |a_i|\,|b_j| = \sum_{i=0}^{p} |a_i| \sum_{j=0}^{q} |b_j| = \|P\|_1\ \|Q\|_1.\end{aligned}$$

La quatrième assertion du théorème 9.5 permet de conclure que f est continue sur $\mathbb{K}[X] \times \mathbb{K}[X]$ lorsque $\mathbb{K}[X]$ est muni de la norme $\|\cdot\|_1$.

3. Si on munit $\mathbb{K}[X]$ de la norme $\|\cdot\|_\infty$ définie pour tout $P = \sum_{k=0}^{p} a_k X^k$ par $\|P\|_\infty = \max_{k \in \{0,\dots,p} |a_k|$ alors l'application bilinéaire symétrique f qui à deux polynômes associe leur produit n'est pas continue. Pour l'établir, nous allons montrer que f n'est pas bornée sur la boule unité fermée dans $\mathbb{K}[X] \times \mathbb{K}[X]$ pour la norme $\|\cdot\|_\infty$ que l'on note $\overline{B}(0,1)$ (on a recours à la troisième assertion du théorème 9.5). Pour $n \in \mathbb{N}$, soit $P_n = \sum_{k=0}^{n} X^k$; on a $\|P_n\|_\infty = 1$ donc P_n appartient à la boule fermée unité. De plus, $f(P,P) = P \times P = \sum_{k=0}^{2n} c_k X^k$ où $c_k = \min(k+1, 2n-k+1)$ donc $\|f(P,P)\|_\infty = \max_{k \in \{0,\dots,2n\}} |c_k| = n+1$. Comme $\lim_{n \to +\infty} \|f(P,P)\|_\infty = +\infty$, on en déduit que f n'est pas bornée sur $\overline{B}(0,1)$ donc que f n'est pas continue pour la norme $\|\cdot\|_\infty$.

9.4 Le théorème du point fixe

Le théorème du point fixe est un résultat essentiel en mathématiques pour la résolution d'équations. Il indique qu'une fonction f admet, sous certaines conditions, un unique point fixe, c'est-à-dire qu'il existe un unique élément ζ solution de l'équation $f(x) = x$. De plus, le théorème du point fixe fournit une méthode permettant de calculer une approximation de la solution lorsque celle-ci ne peut être déterminée explicitement. Signalons que la recherche des zéros d'une fonction numérique g donnée, *i.e.* la résolution de l'équation $g(x) = 0$ se ramène à la recherche des points fixes de la fonction $f : x \in \mathbb{K} \mapsto g(x) - x \in \mathbb{K}$. Outre la résolution d'équations d'inconnue réelle ou complexe, le théorème du point fixe s'applique notamment à l'étude des solutions d'équations dites « fonctionnelles » (l'inconnue est une fonction) dont les équations différentielles.

Définition 9.10 (Point fixe)
Soit f une application d'un ensemble E dans lui-même. On appelle point fixe *de f tout élément x de E tel que $f(x) = x$.*

Remarque Si $f(E) \cap E = \varnothing$ alors f ne peut pas avoir de point fixe dans E. $\diamond$

THÉORÈME 9.6 (Théorème du point fixe de Banach)
Soit (E, d) un espace métrique complet et f une application de E dans E. Si f est contractante[56] *alors f admet un unique point fixe. De plus, quel que soit l'élément α de E, la suite $(u_n)_n$ définie par $u_0 = \alpha$ et $u_{n+1} = f(u_n)$ converge vers ce point fixe.*

Démonstration $\trianglerighteq$ Soit $\alpha \in E$; comme E est complet, pour montrer que la suite $(u_n)_n$ définie par

$$\begin{cases} u_0 = \alpha, \\ u_{n+1} = f(u_n) \quad \forall n \in \mathbb{N} \end{cases}$$

converge, il suffit de montrer qu'il s'agit d'une suite de Cauchy dans E. L'application f étant contractante de rapport $k \in [0,1[$, un raisonnement par récurrence permet d'établir que pour tout $i \in \mathbb{N}$, on a

$$d(u_{i+1}, u_i) \leqslant k^i \, d(u_1, u_0).$$

Soient m et n deux entiers tels que $m \geqslant n \geqslant 0$; en utilisant la première inégalité triangulaire et les propriétés des progressions géométriques[57], on obtient :

$$\begin{aligned} d(u_m, u_n) &\leqslant \sum_{i=n}^{m-1} d(u_{i+1}, u_i) \leqslant \sum_{i=n}^{m-1} k^i d(u_1, u_0) = d(u_1, u_0)\, k^n \sum_{j=0}^{m-n-1} k^j \\ &= d(u_1, u_0)\, k^n \frac{1 - k^{m-n}}{1-k} \leqslant \frac{k^n}{1-k}\, d(u_1, u_0). \end{aligned}$$

Puisque $k \in [0,1[$, la suite réelle positive de terme général $v_n = \dfrac{k^n}{1-k} d(u_1, u_0)$ converge vers 0 ; on a donc

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad \left(n \geqslant N \Longrightarrow \frac{k^n}{1-k}\, d(u_1, u_0) \leqslant \varepsilon \right).$$

On en déduit que

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N \in \mathbb{N} \quad \forall (m,n) \in \mathbb{N}^2 \quad \big((n \geqslant N \text{ et } m \geqslant N) \Longrightarrow d(u_m, u_n) \leqslant \varepsilon \big),$$

ce qui signifie que la suite $(u_n)_n$ est une suite de Cauchy dans E. Comme E est complet, la suite $(u_n)_n$ converge dans E vers un élément ζ de E.

---

[56] Voir la définition 9.8 p. 558.

[57] Voir la proposition 5.21 p. 195 du *Cours de première année*.

D'après la proposition 9.16 p. 559, puisque f est une application contractante sur E, elle est continue sur E. On déduit de la proposition 9.8 que la suite de terme général $f(u_n)$ converge vers $f(\zeta)$. La relation $u_{n+1} = f(u_n)$ pour tout $n \in \mathbb{N}$ exprime le fait que les deux suites de terme général $f(u_n)$ et u_{n+1} sont égales et donc qu'elles ont même limite. Comme la première converge vers $f(\zeta)$ et que la seconde converge vers ζ, on en déduit que $\zeta = f(\zeta)$, autrement dit que ζ est un point fixe de f.

$\trianglerighteq$ Montrons à présent que f ne peut pas admettre deux points fixes ζ et η distincts. Puisque f est une application contractante sur E, il existe un réel $k \in [0,1[$ tel que : $\forall (x,y) \in E^2 \quad d(f(x), f(y)) \leqslant k\, d(x,y)$. On a alors,

$$d(\zeta, \eta) = d(f(\zeta), f(\eta)) \leqslant k\ d(\zeta, \eta)$$

et par conséquent : $(k-1)\ d(\zeta, \eta) \geqslant 0$. Puisque $k \in [0,1[$ on a $k - 1 < 0$ ce qui impose que $d(\zeta, \eta) = 0$ c'est-à-dire $\eta = \zeta$. □

Remarques

1. L'hypothèse 1-lipschitzienne n'est pas suffisante pour assurer l'existence et l'unicité du point fixe. L'exemple de la fonction identité qui est lipschitzienne de rapport 1 pour la distance euclidienne sur $\mathbb{R}$ montre qu'il n'y a pas forcement unicité. L'exemple de la translation $f : \mathbf{x} \in E \mapsto \mathbf{x} + \mathbf{a}$ (où $\mathbf{a} \in E$, $\mathbf{a} \neq \mathbf{0}_E$) qui est aussi lipschitzienne de rapport 1 montre qu'il n'y a pas toujours existence.

2. La condition $d(f(x), f(y)) < d(x,y)$ pour tout $(x,y) \in E^2$ n'est pas suffisante pour assurer l'existence et l'unicité du point fixe. Par exemple, l'application $f : x \in \mathbb{R} \mapsto \sqrt{1+x^2}$ n'a pas de point fixe bien que pour tous réels x et y on ait[58] $|f(x) - f(y)| < |x - y|$.

3. Il se peut que f ne soit pas contractante sur l'espace vectoriel E tout entier mais seulement sur une partie $A \subset E$. Dans ce cas, le théorème s'applique encore à condition que $f(A) \subset A$ et que A soit fermé[59]. ◇

Exercice 9.10 Soient $(E, \|\cdot\|)$ un espace de Banach et f une application continue de E dans E. Montrer que s'il existe $n \in \mathbb{N}^*$ tel que l'application $f^n = f \circ f \circ \cdots \circ f$ (n fois) est contractante alors l'application f admet un point fixe et que celui-ci est unique.

Le théorème du point fixe de Banach indique qu'une application contractante d'un espace métrique complet dans lui-même admet un unique point fixe et fournit une suite qui converge vers cet unique point fixe. Cette suite permet d'un point de vue pratique de calculer une valeur approchée du point fixe lorsque celui-ci ne peut pas être calculé explicitement. La proposition suivante nous donne une majoration de l'erreur commise si l'on considère le n-ème terme de la suite comme valeur approchée du point fixe.

---

[58] On établit ce résultat en remarquant que $|f'(x)| = |x|(1+x^2)^{-1/2}$ est borné par 1 et en utilisant l'inégalité des accroissements finis.

[59] Dans ce cas A est complet car un sous-ensemble fermé d'un espace complet est complet, voir la proposition 7.19 p. 412.

Proposition 9.24 *Soit (E,d) un espace métrique complet et f une application de E dans E contractante de rapport $k \in [0,1[$. Pour tout $\alpha \in E$, la suite $(u_n)_n$ définie par $u_0 = \alpha$ et $u_{n+1} = f(u_n)$, qui converge vers l'unique point fixe ζ de f, vérifie*[60] *pour tout $n \in \mathbb{N}$:*

$$d(u_n,\zeta) \leqslant \frac{k^n}{1-k}\, d(u_1,u_0) \quad \text{et} \quad d(u_n,\zeta) \leqslant k^n\, d(u_0,\zeta).$$

Démonstration ⊵ Pour tous $N \in \mathbb{N}^*$ et $n \in \mathbb{N}$, on obtient en utilisant l'inégalité triangulaire :

$$\begin{aligned} d(u_n,\zeta) &\leqslant d(u_n,u_{n+1}) + d(u_{n+1},u_{n+2}) + \dots \\ &\qquad \dots + d(u_{n+N},u_{n+N+1}) + d(u_{n+N+1},\zeta) \\ &= \sum_{j=n}^{N+n} d(u_j,u_{j+1}) + d(u_{n+N+1},\zeta). \end{aligned} \tag{16}$$

Or, comme f est contractante de rapport $k \in [0,1[$, on a pour tout $j \in \mathbb{N}^*$:

$$d(u_j,u_{j+1}) = d(f(u_{j-1}),f(u_j)) \leqslant k\, d(u_{j-1},u_j).$$

On montre par récurrence que cela implique que pour tout $j \in \mathbb{N}$, on a : $d(u_j,u_{j+1}) \leqslant k^j\, d(u_1,u_0)$. On déduit alors de l'inégalité (16) que pour tous $N \in \mathbb{N}^*$ et $n \in \mathbb{N}$ on a[61]

$$\begin{aligned} d(u_n,\zeta) &\leqslant \sum_{j=n}^{N+n} k^j\, d(u_1,u_0) + d(u_{n+N+1},\zeta) \\ &= k^n\, \frac{1-k^{N+1}}{1-k}\, d(u_1,u_0) + d(u_{n+N+1},\zeta). \end{aligned}$$

D'après le théorème 9.6, la suite $(u_n)_n$ converge vers ζ ; en considérant la limite quand N tend vers $+\infty$ dans l'inégalité précédente, on obtient[62] :

$$\forall n \in \mathbb{N} \qquad d(u_n,\zeta) \leqslant \frac{k^n}{1-k}\, d(u_1,u_0).$$

---

[60] Ces majorations indiquent que la vitesse de convergence de la suite $(u_n)_n$ vers le point fixe ζ est exponentielle (donc très rapide) et d'autant plus rapide que k est proche de 0.

[61] On rappelle, voir la proposition 5.21 p. 195 du *Cours de première année*, que la somme des $n+1$ premiers termes d'une suite géométrique $(x_n)_n$ de raison r vaut : $\sum_{i=0}^{n} x_i = x_0\, \frac{1-r^{n+1}}{1-r}$.

[62] Ce « passage à la limite » se justifie de la manière suivante : considérons pour $n \in \mathbb{N}$ fixé, les suites numérique réelles $(X_N)_N$ et $(Y_N)_N$ de termes généraux $X_N = d(u_n,\zeta)$ (la suite $(X_N)_N$ est donc constante) et $Y_N = k^n\, \frac{1-k^{N+1}}{1-k}\, d(u_1,u_0) + d(u_{n+N+1} - \zeta)$. Pour tout $N \in \mathbb{N}^*$ on a $X_N \leqslant Y_N$. Comme la suite $(X_N)_N$ converge vers $\ell_X = d(u_n,\zeta)$ et que la suite $(Y_N)_N$ converge vers $\ell_Y = \frac{k^n}{1-k} d(u_1,u_0)$, le théorème d'encadrement (théorème 5.1 p. 179 du *Cours de première année*) indique que $\ell_X \leqslant \ell_Y$.

$\trianglerighteq$ La seconde partie de la proposition se démontre par récurrence sur n. Pour $n = 0$, l'inégalité qui s'écrit $d(u_1, u_0) \leqslant d(u_1, u_0)$ est vraie. Supposons la relation vraie pour un entier n donné. En utilisant successivement le fait que ζ est point fixe de f, que f est contractante de rapport k puis l'hypothèse de récurrence, on obtient

$$d(u_{n+1}, \zeta) = d(f(u_n), f(\zeta)) \leqslant k\, d(u_n, \zeta) \leqslant kk^n\, d(u_1, u_0) = k^{n+1}\, d(u_1, u_0).$$

La relation est donc encore vraie pour l'entier $n + 1$. On en conclut que la relation est vraie pour tout $n \in \mathbb{N}$. □

Application du théorème du point fixe à la résolution de systèmes d'équations non linéaires

Le théorème du point fixe joue un rôle essentiel dans l'étude et la résolution de systèmes d'équations non linéaires. Intéressons-nous par exemple à la résolution du système

$$(\mathcal{S}) \qquad \left\{ \begin{array}{rcl} x_1 & = & \lambda_1 \cos(x_2) \\ x_2 & = & \lambda_2 \sin(x_1) \end{array} \right.$$

où λ_1 et λ_2 désignent deux réels appartenant à $]-1, 1[$. Les solutions de ce système sont les points fixes de l'application

$$\Phi : (x_1, x_2) \in \mathbb{R}^2 \longmapsto (\lambda_1 \cos(x_2), \lambda_2 \sin(x_1)) \in \mathbb{R}^2.$$

On peut remarquer que $\Phi(\mathbb{R}^2) =]-\lambda_1, \lambda_1[\times]-\lambda_2, \lambda_2[\subset [-1, 1]^2$ ce qui implique que les points fixes de Φ appartiennent nécessairement à $D = [-1, 1]^2$. L'espace vectoriel $\mathbb{R}^2$ muni de l'une quelconque de ses normes est un espace de Banach, voir la proposition 7.18 p. 408. Par ailleurs, D est un sous-ensemble fermé de $\mathbb{R}^2$ donc il est complet, voir la proposition 7.19 p. 412. Pour pouvoir utiliser le théorème du point fixe, il reste à montrer que Φ est contractante dans $\mathbb{R}^2$ que l'on munit de la norme infinie. Soient $\mathbf{x} = (x_1, x_2)$ et $\mathbf{y} = (y_1, y_2)$ deux éléments de $\mathbb{R}^2$; on a

$$\|\Phi(\mathbf{x}) - \Phi(\mathbf{y})\|_\infty = \max\left(\lambda_1 \left|\cos(x_2) - \cos(y_2)\right|, \lambda_2 \left|\sin(x_1) - \sin(y_1)\right|\right).$$

Or,

$$\begin{array}{rcl} \left|\cos(x_2) - \cos(y_2)\right| & = & 2\left|\sin\left(\frac{1}{2}(x_2 + y_2)\right) \sin\left(\frac{1}{2}(x_2 - y_2)\right)\right|, \\ \left|\sin(x_1) - \sin(y_1)\right| & = & 2\left|\cos\left(\frac{1}{2}(x_1 + y_1)\right) \sin\left(\frac{1}{2}(x_1 - y_1)\right)\right|, \end{array}$$

et pour tout $(\mathbf{x}, \mathbf{y}) \in D^2$ on a $\left|\sin\left(\frac{1}{2}(x_2 + y_2)\right)\right| \leqslant 1$ et $\left|\cos\left(\frac{1}{2}(x_1 + y_1)\right)\right| \leqslant 1$. Par ailleurs, comme pour tout $u \in \mathbb{R}$ on a $|\sin(u)| \leqslant |u|$, on en déduit que

$$\|\Phi(\mathbf{x}) - \Phi(\mathbf{y})\|_\infty \leqslant \max\left(|\lambda_1|\ |x_1 - y_1|, |\lambda_2|\ |x_2 - y_2|\right) \leqslant \max(|\lambda_1|, |\lambda_2|)\ \|\mathbf{x}-\mathbf{y}\|_\infty.$$

L'application Φ est donc une application contractante de D dans D de rapport $k = \max(|\lambda_1|, |\lambda_2|) < 1$ pour la norme infinie. Le théorème du point fixe permet de conclure que Φ admet un unique point fixe dans D. Cela signifie que le système $(\mathcal{S})$ admet une unique solution $\tilde{\mathbf{x}}$ appartenant à D.

On peut obtenir une valeur approchée de cette solution en considérant la suite $(\mathbf{u}_n)_n$ définie par la relation de récurrence $\mathbf{u}_{n+1} = \Phi(\mathbf{u}_n)$ à partir d'une valeur $\mathbf{u}_0 = (u_0^{(1)}, u_0^{(2)})$ quelconque. D'après la proposition 9.24, le n–ième terme de cette suite est une approximation du point fixe recherché et l'erreur commise est inférieure à

$$\frac{k^n}{1-k}\,\|\mathbf{u}_1 - \mathbf{u}_0\|_\infty = \frac{k^n}{1-k}\,\max\left(|\lambda_1\,\cos(u_0^{(2)}) - u_0^{(1)}|, |\lambda_2\,\sin(u_0^{(1)}) - u_0^{(2)}|\right).$$

L'erreur est donc majorée par $M_n = 2k^n/(1-k)$.

Considérons le cas où $\lambda_1 = \lambda_2 = \frac{1}{2}$. On obtiendra une valeur approchée de la solution avec une erreur inférieure à 10^{-3} en prenant n tel que $M_n \leqslant 10^{-3}$. Cette valeur correspond à $n = 12$. Le programme MAPLE ci-dessous met en œuvre l'approche qui vient d'être décrite pour calculer une valeur approchée de $\tilde{\mathbf{x}}$.

```
> lambda1:=1/2: lambda2:=1/2: k:=max(abs(lambda1),abs(lambda2)):
> Phi:= x->[lambda1*cos(x[2]),lambda2*sin(x[1])]:
> err:=10^(-3): n:=ceil(solve(2*k^x/(1-k)=err,x)):
> u:=array(0..n): u[0]:=[0,0]:
> for i from 1 to n do
    u[i]:=evalf(Phi(u[i-1]),4):
  od;
```

```
u[1] := [.5000, 0.]
u[2] := [.5000, .2397]
u[3] := [.4857, .2397]
u[4] := [.4857, .2334]
u[5] := [.4864, .2334]
u[6] := [.4864, .2337]
u[7] := [.4864, .2337]
u[8] := [.4864, .2337]
u[9] := [.4864, .2337]
u[10] := [.4864, .2337]
u[11] := [.4864, .2337]
u[12] := [.4864, .2337]
```

On peut constater qu'à partir de $\mathbf{u}_5$ la valeur approchée de la solution n'évolue plus si l'on s'intéresse à une précision de 10^{-3}. Cela est à mettre en relation avec la fait que la valeur $n = 12$ a été obtenue en effectuant une majoration de l'erreur commise ; l'erreur réelle étant en général inférieure au majorant obtenu.

Le théorème du point fixe a également un rôle dans l'étude des équations dites « intégrales » comme l'illustre l'exercice suivant.

Exercice 9.11 Soient $f \in \mathcal{C}^0([0,1],\mathbb{R})$ et $k \in \mathcal{C}^0([0,1]^2,\mathbb{R})$ telle que $\sup_{(x,y)\in[0,1]^2} |k(x,y)| = \frac{1}{2}$. On considère l'application ϕ qui à une application $g \in \mathcal{C}^0([0,1],\mathbb{R})$ associe l'application $\phi(g)$ est définie par

$$\forall x \in [0,1] \quad \phi(g)(x) = f(x) + \int_0^1 k(x,y)\ g(y)\ \mathrm{d}y.$$

En utilisant le théorème du point fixe, montrer qu'il existe une unique application $g \in \mathcal{C}^0([0,1],\mathbb{R})$ solution de l'équation de Volterra[63] :

$$\forall x \in [0,1] \quad g(x) = f(x) + \int_0^1 k(x,y)\ g(y)\ \mathrm{d}y.$$

Application du théorème du point fixe à l'étude d'équations différentielles

Le théorème du point fixe est également un outil important pour l'étude et la résolution de certaines équations différentielles. Nous allons l'illustrer au travers d'exemples.

Considérons l'équation différentielle linéaire du premier ordre

$$y'(t) + t\,y(t) = t^2 \tag{17}$$

sur l'intervalle $I = [0,1]$ sous la condition initiale $y(0) = 1$. L'équation homogène associée à (17) admet pour solution générale $y_0 : x \in I \mapsto C\,\mathrm{e}^{-x^2/2}$ où C désigne une constante réelle. On peut penser avoir recours à la méthode de la variation de la constante pour déterminer une solution particulière de l'équation (17) : une solution particulière est de la forme $y_1 : x \in I \mapsto C(x)\mathrm{e}^{-x^2/2}$ où $C(x) = \int x^2\,\mathrm{e}^{-x^2/2}\ \mathrm{d}x$. Malheureusement, cette primitive n'est pas exprimable à l'aide des fonctions usuelles.

Le théorème du point fixe permet toutefois de construire une suite de fonctions qui converge vers la solution y de notre équation. On peut remarquer que cette solution y vérifie pour tout $t \in I$,

$$y(t) = y(0) + \int_0^t y'(s)\ \mathrm{d}s = 1 + \int_0^t \big(s^2 - s y(s)\big)\ \mathrm{d}s;$$

elle est donc point fixe de l'application $\Phi : \varphi \in \mathcal{C}^0(I,\mathbb{R}) \mapsto \Phi(\varphi) \in \mathcal{C}^0(I,\mathbb{R})$ où pour tout $x \in I$,

$$\Phi(\varphi)(x) = 1 + \int_0^x \big(s^2 - s\varphi(s)\big)\ \mathrm{d}s.$$

[63] Volterra, Vito (1860, Ancône - 1940, Rome). Mathématicien italien.

Munissons l'espace vectoriel $\mathcal{C}^0(I,\mathbb{R})$ de la norme infinie et montrons que Φ est contractante dans $(\mathcal{C}^0(I,\mathbb{R}),\|\cdot\|_\infty)$. Soient φ_1 et φ_2 deux applications continues sur I et $x\in I$; on a

$$\begin{aligned}|\Phi(\varphi_1)(x)-\Phi(\varphi_2(x))| &= \left|\int_0^x \big(s^2-s\varphi_1(s)\big)\,\mathrm{d}s-\int_0^x \big(s^2-s\varphi_2(s)\big)\,\mathrm{d}s\right| \\ &= \left|\int_0^x s\big(\varphi_1(s)-\varphi_2(s)\big)\,\mathrm{d}s\right| \leqslant \int_0^x |s|\big|\varphi_1(s)-\varphi_2(s)\big|\,\mathrm{d}s.\end{aligned}$$

Or, pour tout $s\in[0,x]$ on a

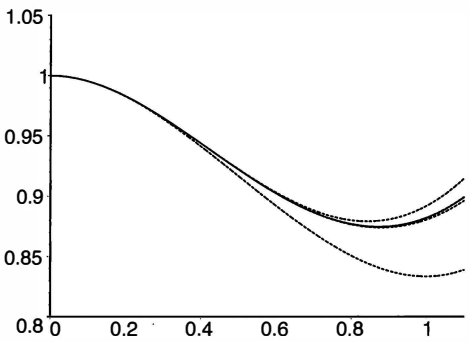
$$\big|\varphi_1(s)-\varphi_2(s)\big| \leqslant \sup_{t\in[0,x]}\big|\varphi_1(t)-\varphi_2(t)\big| \leqslant \sup_{t\in I}\big|\varphi_1(t)-\varphi_2(t)\big|$$

et $\int_0^x |s|\,\mathrm{d}s=\int_0^x s\,\mathrm{d}s=\frac{1}{2}x^2\leqslant\frac{1}{2}$. On en déduit que pour tout $x\in I$ on a

$$|\Phi(\varphi_1)(x)-\Phi(\varphi_2(x))| \leqslant \tfrac{1}{2}\,\|\varphi_1-\varphi_2\|_\infty,$$

ce qui implique que $\|\Phi(\varphi_1)-\Phi(\varphi_2)\|_\infty\leqslant\frac{1}{2}\,\|\varphi_1-\varphi_2\|_\infty$. Ainsi, l'application Φ est contractante dans $(\mathcal{C}^0(I,\mathbb{R}),\|\cdot\|_\infty)$ qui est un espace de Banach (voir p. 403) et le théorème du point fixe indique que la suite de fonctions $(f_n)_n$ définie par la relation de récurrence $f_{n+1}=\Phi(f_n)$ converge quel que soit le choix de la fonction f_0 vers l'unique point fixe de Φ. Ce point fixe coïncide sur I avec la solution de l'équation différentielle (17) sous la condition initiale $y(0)=1$. On peut établir en utilisant un raisonnement par récurrence que si f_0 est une fonction polynomiale de degré p, alors f_n est une fonction polynomiale de degré $p+n$ pour tout $n\in\mathbb{N}^*$. Cela permet de construire très facilement une suite de fonctions qui converge vers la solution recherchée. Illustrons cette étude avec MAPLE.

```
> N:=10: f:=array(0..N);
> Phi := phi -> unapply(y0 +int(g(t)-t*phi(t),t=0..x),x):
> f[0]:= x-> 1;
> for n from 1 to N do
    f[n] := Phi(f[n-1]);
  od;
> plot([seq(f[i](x),i=1..N)],x=0..1.1);
```



Graphiquement, on remarque que la suite de fonctions polynomiales $(f_n)_n$ converge très rapidement. Le tracé plus sombre, correspondant à la superposition des représentations graphiques des fonctions f_4 à f_{10}, donne la forme de la représentation graphique de notre solution. D'après la proposition 9.24, l'écart entre le n–ième terme de la suite $(f_n)_n$ et la solution exacte mesuré par la distance subordonnée à la norme infinie est majoré par

$$2^{1-n}\|f_1 - f_0\|_\infty = 2^{1-n} \sup_{x\in[0,1]} \left|\tfrac{1}{3}x^3 - \tfrac{1}{2}x^2\right| = \frac{2^{-n}}{3}.$$

Ainsi, f_n sera une approximation de la solution avec une erreur inférieure à 1 pour cent pour n tel que $2^{-n}/3 \leqslant 10^{-2}$ soit pour $n \geqslant -\ln(3\,10^{-2})/\ln(2)$. La fonction polynomiale f_6 convient ; son expression est

$$f_6(x) = -\frac{1}{135135}x^{13} + \frac{1}{46080}x^{12} + \frac{1}{10395}x^{11} - \frac{1}{3840}x^{10} - \frac{1}{945}x^9 + \frac{1}{384}x^8 + \frac{1}{105}x^7 - \frac{1}{48}x^6 - \frac{1}{15}x^5 + \frac{1}{8}x^4 + \frac{1}{3}x^3 - \frac{1}{2}x^2 + 1.$$

Insistons sur le fait que $2^{-n}/3$ est un majorant de l'erreur maximale commise sur l'intervalle $[0, 1]$; l'erreur réellement commise peut être très inférieure à ce majorant.

9.5 Exercices de synthèse

Exercice 9.12

1 - Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux $\mathbb{K}$-espaces vectoriels normés et f une application linéaire de E dans F vérifiant la propriété suivante : pour toute suite $(\mathbf{x}_n)_n$ d'éléments de E convergeant pour $\|\cdot\|_E$ vers $\mathbf{0}_E$, la suite de terme général $f(\mathbf{x}_n)$ est bornée dans $(F, \|\cdot\|_F)$.

a) Montrer que[64] si une suite $(\mathbf{y}_n)_n$ d'éléments de E converge vers $\mathbf{0}_E$ alors la suite de terme général $f(\mathbf{y}_n)$ converge vers $\mathbf{0}_F$.

b) En déduire que f est une application continue sur E.

2 - a) Montrer, en utilisant le résultat établi à la question précédente, que la forme linéaire $f : h \in \mathcal{C}^0([0,1],\mathbb{R}) \mapsto h(1)$ n'est pas continue si $\mathcal{C}^0([0,1],\mathbb{R})$ est muni de la norme : $h \in \mathcal{C}^0([0,1],\mathbb{R}) \mapsto \int_0^1 |h(x)|\,dx$ (norme de la convergence en moyenne, notée $\|\cdot\|_1$).

b) Montrer que f est une application continue de $(\mathcal{C}^0([0,1],\mathbb{R}), \|\cdot\|_\infty)$ dans $\mathbb{R}$.

---

[64] On pourra considérer une suite $(u_n)_n$ d'éléments de E convergeant vers $\mathbf{0}_E$ telle que pour tout $n \in \mathbb{N}$ on ait $\mathbf{x}_n = \sqrt{\|\mathbf{x}_n\|_E}\,\mathbf{u}_n$.

Exercice 9.13

Soit b une fonction positive et continue sur $[0,T]$ (où $T \in \mathbb{R}_+^*$ est fixé). On s'intéresse au problème suivant[65] : trouver $y \in \mathcal{C}^1([0,T],\mathbb{R})$ telle que

$$(\mathcal{P}) \qquad \begin{cases} y'(t) &= b(t) & \forall t \in [0,T] \\ y(0) &= \sqrt{1+y(T)} \end{cases}.$$

1 - Justifier que résoudre le problème $(\mathcal{P})$ revient à trouver $y \in \mathcal{C}^0([0,T],\mathbb{R})$ telle que

$$\forall t \in [0,T] \qquad y(t) = \sqrt{1+y(T)} + \int_0^t b(s)\,\mathrm{d}s.$$

2 - Dans $\mathcal{C}^0([0,T],\mathbb{R})$ muni de la norme uniforme, montrer que l'ensemble A des fonctions positives est complet.

3 - En utilisant le théorème du point fixe, montrer que le problème $(\mathcal{P})$ admet une unique solution.

Exercice 9.14

Soit f une forme linéaire sur $\mathcal{C}^0([0,1],\mathbb{R})$ qui vérifie la propriété $(\mathcal{P})$ suivante :

« l'image par f de toute application continue et positive sur $[0,1]$ est un réel positif. »

Le but de cet exercice est de montrer que f est une application continue de $(\mathcal{C}^0([0,1],\mathbb{R}), \|\cdot\|_\infty)$ dans $(\mathbb{R}, |\cdot|)$.

1 - Donner un exemple d'application de $\mathcal{L}((\mathcal{C}^0([0,1],\mathbb{R}),\mathbb{R})$ vérifiant la propriété $(\mathcal{P})$ et calculer l'image par cette application de la fonction constante égale à 1.

2 - a) Montrer que si f est une forme linéaire sur $\mathcal{C}^0([0,1],\mathbb{R})$ vérifiant la propriété $(\mathcal{P})$ alors f est une application croissante.

b) En déduire que pour tout $h \in \mathcal{C}^0([0,1],\mathbb{R})$ on a $|f(h)| \leqslant f(|h|)$ puis que $f(|h|) \leqslant \|h\|_\infty f(1)$ où 1 désigne la fonction constante égale à 1.

3 - Montrer que f est une application continue de $(\mathcal{C}^0([0,1],\mathbb{R}), \|\cdot\|_\infty)$ dans $(\mathbb{R}, |\cdot|)$ et calculer la norme linéaire de f.

---

[65] La fonction y représente l'évolution au cours du temps d'une certaine quantité dont on souhaite relier la valeur initiale et la valeur finale. Cette problématique intervient par exemple en sciences financières où il est essentiel de pouvoir relier valeur initiale et valeur finale d'un investissement.

9.6 Solutions des exercices

Solution de l'exercice 9.1

Puisque δ_E est une distance sur E équivalente à d_E, il existe $C_1, C_2 \in \mathbb{R}_+^*$ tels que $C_1\ d_E \leqslant \delta_E \leqslant C_2\ d_E$ et puisque δ_F est une distance sur F équivalente à d_F, il existe $C_1', C_2' \in \mathbb{R}_+^*$ tels que $C_1'\ d_F \leqslant \delta_F \leqslant C_2'\ d_F$. On en déduit que pour tout $(x, y) \in E^2$ on a

$$\delta_F(f(x), f(y)) \leqslant C_2'\ d_F(f(x), f(y)) \leqslant C_2'\, k\ d_E(x, y) \leqslant C_2'\, k\, C_1^{-1}\ \delta_E(x, y).$$

L'application f est donc une application lipschitzienne de (E, δ_E) dans (F, δ_F) de rapport $C_2'\, C_1^{-1}\, k$. On notera toutefois que le rapport dépend en général des distances choisies sur E et F.

Solution de l'exercice 9.2

1 - a) On a

$$\begin{aligned} U &= f^{-1}(]\alpha, +\infty[) = \{x \in E\ ;\ f(x) > \alpha\}, \\ V &= g^{-1}(]-\infty, \alpha[) = \{x \in E\ ;\ g(x) < \alpha\}, \end{aligned}$$

et par conséquent si $x \in U \cap V$ alors $f(x) > \alpha$ et $g(x) < \alpha$ donc $f(x) > g(x)$ ce qui implique que $x \in A$. On en déduit que $U \cap V \subset A$.

Pour montrer que A est ouvert, montrons que[(66)] pour tout $a \in A$, il existe $r \in \mathbb{R}_+^*$ tel que $B(a, r) \subset A$. Soit $a \in A$; on a $f(a) > g(a)$ et par conséquent $g(a) < \alpha < f(a)$ ce qui indique que $a \in U$ et $a \in V$. Les ensembles U et V sont des ensembles ouverts (ils sont images réciproques d'intervalles ouverts de $\mathbb{R}$ par des fonctions continues) donc $U \cap V$ est un ensemble ouvert[(67)].

Comme a est un élément de $U \cap V$ et que $U \cap V$ est ouvert, on en déduit qu'il existe $r \in \mathbb{R}_+^*$ tel que $B(a, r) \subset (U \cap V)$. Pour conclure que A est ouvert, il suffit d'utiliser l'inclusion $U \cap V \subset A$.

2 - Notons $B = \{x \in E\ ;\ f(x) \leqslant g(x)\}$; montrer que $f(x) \leqslant g(x)$ pour tout $x \in E$ revient à montrer que $B = E$. Par hypothèse, $D \subset B \subset E$ ce qui implique que[(68)] $\overline{D} \subset \overline{B} \subset E$. Or, comme D est dense dans E, on a $\overline{D} = E$ et par conséquent $\overline{B} = E$. Ainsi, montrer que $B = E$, revient à montrer que B est fermé (car on a alors $\overline{B} = B$). L'ensemble B est bien un fermé de E car son complémentaire est

$$\complement_E B = \{x \in E\ ;\ f(x) > g(x)\} = A$$

et on a montré que A est un ouvert.

---

(66) Voir la définition 7.7 p. 371.

(67) Voir la proposition 7.9 p. 375.

(68) Voir la proposition 7.11 p. 379.

Solution de l'exercice 9.3

1 - La fonction $f : (x_1, x_2) \in \mathbb{R}^2 \mapsto x_1/\sqrt{x_1^2 + x_2^2}$ est définie sur $\mathbb{R}^2_*$. Sa première application partielle en $(0,0)$ est l'application $f_{(0,0),1} : x_1 \in \mathbb{R}^* \mapsto x_1/|x_1|$. Comme celle-ci n'admet pas de limite en 0, on en déduit que f n'a pas de limite en $(0,0)$.

2 - La fonction $f : (x_1, x_2) \in \mathbb{R}^2 \mapsto |\sin(x_1 + x_2)|/\sqrt{x_1^2 + x_2^2}$ est définie sur $\mathbb{R}^2_*$. Comme $\sin(u) \underset{0}{\sim} u$, on a pour tout $\lambda \in \mathbb{R}$

$$f(x_1, \lambda x_1) = \frac{|\sin((\lambda+1)x_1)|}{\sqrt{(1+\lambda^2)x_1^2}} \underset{0}{\sim} \frac{|\lambda+1|}{\sqrt{1+\lambda^2}}.$$

On en déduit que f n'a pas de limite en $(0,0)$.

3 - La fonction $f : (x_1, x_2) \in \mathbb{R}^2 \mapsto \sin(x_1^2 + x_2^2)/(|x_1| + |x_2|)$ est définie sur $\mathbb{R}^2_*$. Pour tout $\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2_*$, on a

$$f(x_1, x_2) = \frac{\sin(x_1^2 + x_2^2)}{x_1^2 + x_2^2} \times \frac{x_1^2 + x_2^2}{|x_1| + |x_2|}.$$

Puisque $\sin(u) \underset{0}{\sim} u$, on a $\lim_{(x_1,x_2)\to(0,0)} \frac{\sin(x_1^2+x_2^2)}{x_1^2+x_2^2} = 1$. Par ailleurs,

$$0 \leqslant \frac{x_1^2 + x_2^2}{|x_1| + |x_2|} = \frac{\|\mathbf{x}\|_2^2}{\|\mathbf{x}\|_1}.$$

Comme sur $\mathbb{R}^2$, toutes les normes sont équivalentes, il existe $C \in \mathbb{R}^*_+$ tel que $\|\mathbf{x}\|_2^2/\|\mathbf{x}\|_1 \leqslant C\|\mathbf{x}\|^2$. On en déduit que $\lim_{(x_1,x_2)\to(0,0)} \frac{x_1^2+x_2^2}{|x_1|+|x_2|} = 0$ et par conséquent que f a pour limite 0 en $(0,0)$.

4 - Pour tout $(x_1, x_2) \in \mathbb{R}^2$ avec $x_1 \neq 0$ et $x_2 \neq 0$ on a

$$f(x_1, x_2) = \frac{\ln(1 + x_1x_2)}{x_1x_2} \times \frac{x_1x_2}{\sqrt{x_1^2 + x_2^2}}$$

Comme $\lim_{(x_1,x_2)\to(0,0)} x_1x_2 = 0$ et que $\lim_{u\to 0} \frac{\ln(1+u)}{u} = 1$ on a $\lim_{(x_1,x_2)\to(0,0)} \frac{\ln(1+x_1x_2)}{x_1x_2} = 1$. Par ailleurs, comme $2|x_1x_2| \leqslant x_1^2 + x_2^2$ on a

$$0 \leqslant \frac{|x_1x_2|}{\sqrt{x_1^2 + x_2^2}} \leqslant \tfrac{1}{2}\sqrt{x_1^2 + x_2^2} = \tfrac{1}{2}\|(x_1, x_2)\|_2$$

et d'après le théorème d'encadrement $\lim_{(x_1,x_2)\to(0,0)} \frac{|x_1x_2|}{\sqrt{x_1^2+x_2^2}} = 0$. Enfin, en restriction aux droites d'équations $x_1 = 0$ et $x_2 = 0$, f est l'application nulle donc elle admet également 0 pour limite en $(0,0)$. On en conclut que $\lim_{(x_1,x_2)\to(0,0)} f(x_1, x_2) = 0$.

5 - La fonction $f : (x_1, x_2) \in \mathbb{R}^2 \mapsto x_1x_2^2/(x_1^4 + x_2^3)$ est définie sur $\mathbb{R}^2_*$. Considérons l'arc paramétré $\varphi : t \in \mathbb{R} \mapsto t\,\mathbf{u}$ avec $\mathbf{u} = (1, \lambda)$ où λ désigne un réel fixé dont l'image est une droite passant par l'origine. On a $(0,0) = \varphi(0)$ et

$$\forall t \in \mathbb{R} \qquad f(\varphi(t))) = \frac{\lambda^2 t^3}{t^4 + \lambda^3 t^3} = \frac{\lambda^2}{t + \lambda^3}.$$

Comme la limite en 0 de $f(\varphi(t))$, qui vaut $1/\lambda$, dépend de la valeur de λ, l'application f n'a pas de limite à l'origine.

6 - La fonction $f : (x_1, x_2) \in \mathbb{R}^2 \mapsto x_1 x_2^2/(x_1^4 + x_2^2)$ est définie sur $\mathbb{R}^2_*$. Considérons l'arc paramétré $\varphi : t \in \mathbb{R} \mapsto t\,\mathbf{u}$ avec $\mathbf{u} = (1, \lambda)$ où λ désigne un réel fixé dont l'image est une droite passant par l'origine. On a $(0,0) = \varphi(0)$ et

$$\forall t \in \mathbb{R} \qquad f(\varphi(t))) = \frac{\lambda^2 t^3}{t^4 + \lambda^2 t^2} = \frac{t\lambda^2}{t^2 + \lambda^2}.$$

La limite en 0 de $f(\varphi(t)))$ vaut 0. Contrairement à la fonction précédente, on n'a pas pu établir que f n'avait pas de limite en $(0,0)$ par cette méthode. Par contre, on peut déduire de cette étude que si la limite de f en $(0,0)$ existe, elle vaut 0. Essayons d'établir que f a pour limite 0 en $(0,0)$. Pour tout $(x_1, x_2) \in \mathbb{R}^2_*$, on a $x_1^4 + x_2^2 \geqslant x_2^2 \geqslant 0$ donc

$$0 \leqslant |f(x_1, x_2)| \leqslant \frac{|x_1 x_2^2|}{x_2^2} = |x_1| \leqslant \|(x_1, x_2)\|_1.$$

Le théorème d'encadrement permet de conclure que f a pour limite 0 en $(0,0)$.

Solution de l'exercice 9.4

1 - L'application f étant une application linéaire de E dans F, on a pour tous $\mathbf{x}$ et $\mathbf{y}$ dans E : $f(\mathbf{x}+\mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y})$. En particulier en prenant $\mathbf{y} = -\mathbf{x}$, on obtient d'une part

$$f(\mathbf{x}+\mathbf{y}) = f(\mathbf{x}-\mathbf{x}) = f(\mathbf{0}_E)$$

et d'autre part,

$$f(\mathbf{x}) + f(\mathbf{y}) = f(\mathbf{x}) + f(-\mathbf{x}) = f(\mathbf{x}) - f(\mathbf{x}) = \mathbf{0}_F.$$

On peut donc en conclure que $f(\mathbf{0}_E) = \mathbf{0}_F$.

2 - Pour montrer que f est l'application nulle, considérons un élément $\mathbf{y}$ quelconque de E et montons que $f(\mathbf{y}) = \mathbf{0}_F$.

L'ensemble U étant non vide, il existe $\mathbf{a} \in U$ tel que $f(\mathbf{a}) = \mathbf{0}_F$. De plus, U étant un ouvert[69] il existe $r \in \mathbb{R}_+^*$ tel que la boule ouverte $B(\mathbf{a}, r)$ soit incluse dans U.

Soit $\mathbf{y}$ un élément quelconque de E différent de $\mathbf{0}_E$ et $\mathbf{z} = \mathbf{a} + \frac{r\,\mathbf{y}}{2\|y\|_E}$. On a

$$\|\mathbf{z} - \mathbf{a}\|_E = \left\| \frac{r\,\mathbf{y}}{2\,\|y\|_E} \right\|_E = \frac{1}{2} r \left\| \frac{\mathbf{y}}{\|y\|_E} \right\|_E = \frac{1}{2} r < r.$$

On en déduit que $\mathbf{z}$ appartient à $B(\mathbf{a}, r)$, donc que $\mathbf{z}$ appartient à U. D'après l'hypothèse, cela implique que $f(\mathbf{z}) = \mathbf{0}_F$. Or, f étant une application linéaire, on a

$$f(\mathbf{z}) = \underbrace{f(\mathbf{a})}_{=\mathbf{0}_F} + \underbrace{\frac{r}{2\|y\|_E}}_{\neq 0} f(\mathbf{y}).$$

---

[69] Voir la définition 7.7 p. 371.

On en déduit que $f(\mathbf{y}) = \mathbf{0}_F$, ce qui permet de conclure que f est l'application nulle.

SOLUTION DE L'EXERCICE 9.5

Commençons par remarquer que les implications suivantes sont vraies, pas uniquement pour des applications linéaires, mais pour toute application de E dans F (voir la proposition 9.16 p. 559) :

1. f est lipschitzienne sur E implique f est continue sur E ;
2. f est lipschitzienne sur E implique f est uniformément continue sur E ;
3. f est uniformément continue sur E implique f est continue sur E.

Pour établir l'équivalence des trois assertions proposées, il reste à établir que pour les applications linéaires, on a les implications suivantes :

1. si f est continue sur E alors f est uniformément continue sur E ;
2. si f est continue sur E alors f est lipschitzienne sur E.

La troisième implication (si f est uniformément continue sur E alors. f est lipschitzienne sur E) s'obtiendra par transitivité.

1 - Soit f une application linéaire de E dans F continue sur E. On a $f(\mathbf{0}_E) = \mathbf{0}_F$ et f étant continue en 0_E on a

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists \eta \in \mathbb{R}_+^* \quad \forall \mathbf{z} \in E \quad \big(\|\mathbf{z}\|_E \leqslant \eta \implies \|f(\mathbf{z})\|_F \leqslant \varepsilon\big). \tag{18}$$

Montrons que f est uniformément continue sur E, *i.e.* montrons que

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists \eta \in \mathbb{R}_+^* \quad \forall (\mathbf{x}, \mathbf{y}) \in E^2 \quad \big(\|\mathbf{x} - \mathbf{y}\|_E \leqslant \eta \implies \|f(\mathbf{x}) - f(\mathbf{y})\|_F \leqslant \varepsilon\big). \tag{19}$$

Comme f est une application linéaire, pour tout $(\mathbf{x}, \mathbf{y}) \in E^2$, on a $\|f(\mathbf{x}) - f(\mathbf{y})\|_F = \|f(\mathbf{x} - \mathbf{y})\|_F$ et l'assertion (19) est une conséquence de l'assertion (18) considérée avec $\mathbf{z} = \mathbf{x} - \mathbf{y}$.

2 - Soit f une application linéaire de E dans F continue sur E ; d'après le théorème 9.3, il existe un réel positif C tel que pour tout $\mathbf{z} \in E$ on ait $\|f(\mathbf{z})\|_F \leqslant C\|\mathbf{z}\|_E$. Montrons que f est lipschitzienne, *i.e.* montrons qu'il existe un réel positif K tel que :

$$\forall (\mathbf{x}, \mathbf{y}) \in E^2 \quad \|f(\mathbf{x}) - f(\mathbf{y})\|_F \leqslant K \|\mathbf{x} - \mathbf{y}\|_E.$$

Comme f est une application linéaire, pour tout $(\mathbf{x}, \mathbf{y}) \in E^2$, on a $\|f(\mathbf{x}) - f(\mathbf{y})\|_F = \|f(\mathbf{x} - \mathbf{y})\|_F$. L'hypothèse, qui se traduit par l'assertion (18), considérée avec $\mathbf{z} = \mathbf{x} - \mathbf{y}$ indique que $\|f(\mathbf{x}) - f(\mathbf{y})\|_F \leqslant C \|\mathbf{x} - \mathbf{y}\|_E$. On en déduit que f est lipschitzienne de rapport $K = C$. On peut bien entendu démontrer ce résultat sans avoir recours au théorème 9.3 en utilisant directement l'approche qui a permis d'établir dans la démonstration du théorème que si f est une application linéaire de E dans F continue sur E alors il existe un réel positif C tel que pour tout $\mathbf{z} \in E$ on ait $\|f(\mathbf{z})\|_F \leqslant C\|\mathbf{z}\|_E$.

Solution de l'exercice 9.6

1 - Le fait que l'application f est une application linéaire découle des propriétés de la dérivation de polynôme, voir la proposition 6.4 p. 234 du *Cours de première année* ; pour tous $P, Q \in \mathbb{K}[X]$ et pour tous $\lambda \in \mathbb{K}$, on a

$$(P+Q)' = P' + Q' \quad \text{et} \quad (\lambda P)' = \lambda P'.$$

2 - Pour montrer que l'application N qui à un polynôme $P = \sum_{k=0}^{n} a_k X^k$ de $\mathbb{K}[X]$ associe le réel $N(P) = |a_0| + \max_{k\in\{1,\dots,n\}} |k\,a_k|$ est une norme, il s'agit d'établir les quatre conditions de la définition 7.3 p. 358.

L'application N est bien à valeurs positives. De plus, si P désigne le polynôme nul alors tous les coefficients de P sont nuls de sorte que $N(P) = 0$. Réciproquement, si un polynôme $P = \sum_{k=0}^{n} a_k X^k$ vérifie $N(P) = 0$ alors $|a_0| = 0$ et pour tout $k \in \{1, \dots, n\}$ on a $|k\,a_k| = 0$ car une somme de réels positifs est nulle si et seulement si chacun des termes de la somme est égal à 0. On a donc $a_k = 0$ pour tout $k \in \{0, \dots, n\}$ ce qui implique que P est le polynôme nul. Si $P = \sum_{k=0}^{n} a_k X^k$ est un polynôme de $\mathbb{K}[X]$ et λ un réel, le polynôme $\lambda \cdot P$ est le polynôme de degré n dont les coefficients sont $\lambda\,a_k$ pour $k \in \{0, \dots, n\}$. On en déduit que

$$N(\lambda \cdot P) = |\lambda\,a_0| + \max_{k\in\{1,\dots,n\}} |k\,\lambda\,a_k| = |\lambda| \left(|a_0| + \max_{k\in\{1,\dots,n\}} |k\,a_k| \right) = |\lambda|\,N(P).$$

Pour pouvoir affirmer que N est une norme sur $\mathbb{K}[X]$, il reste à établir l'inégalité triangulaire :

$$\forall (P, Q) \in \mathbb{K}[X] \times \mathbb{K}[X] \qquad N(P+Q) \leqslant N(P) + N(Q).$$

Considérons deux polynômes $P = \sum_{k=0}^{n} a_k X^k$ et $Q = \sum_{k=0}^{m} b_k X^k$ et supposons[70] que le degré n de P soit supérieur au degré m de Q. De plus, si $n > m$, on peut poser $b_k = 0$ pour tout $k \in \{m+1, \dots, n\}$. Le polynôme $P + Q$ s'écrit alors $\sum_{k\in\{0,\dots,n\}} (a_k + b_k) X^k$ et on a

$$N(P+Q) = |a_0 + b_0| + \max_{k\in\{1,\dots,n\}} |k\,(a_k + b_k)|.$$

En utilisant la première inégalité triangulaire pour la valeur absolue (ou le module si $\mathbb{K} = \mathbb{C}$), on en déduit que

$$N(P+Q) \leqslant |a_0| + |b_0| + \max_{k\in\{1,\dots,n\}} \{|k\,a_k| + |k\,b_k)|\}.$$

Comme pour tous réels $\alpha_1, \alpha_2, \beta_1, \beta_2$, on a l'inégalité

$$\max\{(\alpha_1 + \beta_1), (\alpha_2 + \beta_2)\} \leqslant \max\{\alpha_1, \alpha_2\} + \max\{\beta_1, \beta_2\},$$

[70] Cette hypothèse ne limite en rien la portée du raisonnement : étant donné deux polynômes, on peut toujours choisir d'appeler P celui de degré le plus élevé et Q l'autre.

on en conclut que :

$$N(P+Q) \leqslant |a_0|+|b_0|+\max_{k\in\{1,\dots,n\}}\{|k\,a_k|\}+\max_{k\in\{1,\dots,n\}}\{|k\,b_k|\} = N(P)+N(Q).$$

L'application N définit bien une norme sur $\mathbb{K}[X]$.

3 - Montrons que l'application linéaire f qui à un polynôme $P=\sum_{k=0}^{n} a_k\,X^k$ associe son polynôme dérivé $P'=\sum_{k=1}^{n} k\,a_k\,X^{k-1}=\sum_{\ell=0}^{n-1}(\ell+1)\,a_{\ell+1}\,X^{\ell}$ est continue de $(\mathbb{K}(X),N)$ dans $(\mathbb{K}(X),\|\cdot\|_\infty)$. Pour tout $P\in\mathbb{K}[X]$ on a

$$\begin{aligned}\|f(P)\|_\infty &= \|P'\|_\infty = \max_{\ell\in\{0,\dots,n-1\}}|(\ell+1)\,a_{\ell+1}| = \max_{k\in\{1,\dots,n\}}|k\,a_k| \\ &\leqslant |a_0|+\max_{k\in\{1,\dots,n\}}|k\,a_k| = N(P).\end{aligned}$$

La quatrième assertion du théorème 9.3 permet de conclure à la continuité de l'application f (on a $C=1$).

Solution de l'exercice 9.7

1 - Pour tout $f\in\mathcal{C}^0([a,b],\mathbb{R})$ et pour tout $x\in[a,b]$, on a

$$|\Phi(f)(x)| = \left|\int_a^x f(t)\,\mathrm{d}t\right| \leqslant \int_a^x |f(t)|\,\mathrm{d}t \leqslant \int_a^b |f(t)|\,\mathrm{d}t = \|f\|_1.$$

On en déduit que

$$\|\Phi(f)\|_1 = \int_a^b |\Phi(f)(x)|\,\mathrm{d}x \leqslant \int_a^b \|f\|_1\,\mathrm{d}x = (b-a)\|f\|_1,$$

ce qui d'après le théorème 9.3 indique que Φ est continue sur $(\mathcal{C}^0([a,b],\mathbb{R}),\|\cdot\|_1)$.

2 - Comme pour tout $f\in\mathcal{C}^0([a,b],\mathbb{R})$, on a $\|\Phi(f)\|_1\leqslant(b-a)\|f\|_1$, on en déduit que

$$|||\Phi||| = \sup_{\substack{f\in\mathcal{C}^0([a,b],\mathbb{R})\\ f\neq 0}} \frac{\|\Phi(f)\|_1}{\|f\|_1} \leqslant (b-a).$$

Considérons la suite $(f_n)_n$ de terme général $f_n : x\in[a,b]\mapsto n^2\mathrm{e}^{n(a-x)}$. Pour tout $n\in\mathbb{N}^*$, on a

$$\frac{\|\Phi(f_n)\|_1}{\|f_n\|_1} \leqslant |||\Phi||| \leqslant b-a. \tag{20}$$

Pour tous $n\in\mathbb{N}$ et $x\in[a,b]$, on a $\Phi(f_n)(x)=\int_a^x f_n(t)\,\mathrm{d}t = n(1-\mathrm{e}^{n(a-x)})$. On en déduit que

$$\|f_n\|_1 = \int_a^b |f_n(x)|\,\mathrm{d}x = \int_a^b f_n(x)\,\mathrm{d}x = \Phi(f_n)(b) = n(1-\mathrm{e}^{-n(b-a)})$$

$$\|\Phi(f_n)\|_1 = \int_a^b \Phi(f_n)(x)\,\mathrm{d}x = \mathrm{e}^{-n(b-a)}+n(b-a)-1$$

puis que $\dfrac{\|\Phi(f_n)\|_1}{\|f_n\|_1} = -\dfrac{1}{n} + \dfrac{b-a}{1-\mathrm{e}^{-n(b-a)}}$. Comme $\lim\limits_{n\to+\infty} \|\Phi(f_n)\|_1 = b-a$, le théorème d'encadrement appliqué à la relation (20) permet de conclure que $|||\Phi||| = b-a$.

Solution de l'exercice 9.8

Le fait que le noyau de p est fermé est une propriété qui a été établie p. 581 pour toute application linéaire continue. Intéressons-nous à l'image de p :

$$\mathrm{Im}(p) = \{\mathbf{y} \in E \ ; \ \exists \mathbf{x} \in E \quad \mathbf{y} = p(\mathbf{x})\}.$$

Soit $\mathbf{y} \in \mathrm{Im}(p)$; il existe $\mathbf{x} \in E$ tel que $p(\mathbf{y}) = p(p(\mathbf{x})) = p(\mathbf{x}) = \mathbf{y}$ donc tout élément de $\mathrm{Im}(p)$ vérifie $p(\mathbf{y}) = \mathbf{y}$. Inversement, si $\mathbf{y} = p(\mathbf{y})$ alors $\mathbf{y}$ appartient à $\mathrm{Im}(p)$ (prendre $\mathbf{x} = \mathbf{y}$). On a donc

$$\mathrm{Im}(p) = \{\mathbf{y} \in E \ ; \ p(\mathbf{y}) = \mathbf{y}\} = \{\mathbf{y} \in E \ ; \ p(\mathbf{y}) - \mathbf{y} = 0\} = \mathrm{Ker}(p - \mathrm{id}_E).$$

Comme p est continue sur E, l'application $f = p - \mathrm{id}_E$ est un endomorphisme de E continu. Son noyau est donc fermé dans E ; cela implique que $\mathrm{Im}(p)$ est fermé dans E (voir p. 581).

Solution de l'exercice 9.9

1 - La suite $(f_n)_n$ étant une suite de Cauchy dans $\mathcal{L}_c(E)$, on a[71]

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N \in \mathbb{N} \quad \forall (m,n) \in \mathbb{N}^2 \quad \big(n \geqslant m \geqslant N \implies |||f_n - f_m||| \leqslant \varepsilon\big). \quad (21)$$

Or, f étant un endomorphisme continu, pour tous $\mathbf{x} \in E$ et $m, n \in \mathbb{N}$, on a[72]

$$\|f_n(\mathbf{x}) - f_m(\mathbf{x})\| = \|(f_n - f_m)(\mathbf{x})\| \leqslant |||f_n - f_m|||\ \|\mathbf{x}\|.$$

Soient ε' un réel strictement positif fixé et $\mathbf{x} \in E$ avec $\mathbf{x} \neq \mathbf{0}_E$; d'après l'assertion (21) considérée avec $\varepsilon = \varepsilon'/\|\mathbf{x}\|$, il existe $N \in \mathbb{N}$ tel que

$$\forall (m,n) \in \mathbb{N}^2 \quad \big(n \geqslant m \geqslant N \implies |||f_n - f_m||| \leqslant \varepsilon'/\|\mathbf{x}\|\big).$$

On déduit des deux dernières assertions que pour tout $\mathbf{x} \in E$ avec $\mathbf{x} \neq \mathbf{0}_E$ on a :

$$\forall \varepsilon' \in \mathbb{R}_+^* \quad \exists N \in \mathbb{N} \quad \forall (m,n) \in \mathbb{N}^2 \quad \big(n \geqslant m \geqslant N \ \Rightarrow \ \|f_n(\mathbf{x}) - f_m(\mathbf{x})\| \leqslant \varepsilon'\big). \quad (22)$$

Si $\mathbf{x} = \mathbf{0}_E$ alors l'assertion est encore vraie car on a dans ce cas $\|f_n(\mathbf{x}) - f_m(\mathbf{x})\| = 0 \leqslant \varepsilon'$.

D'après l'assertion 22, pour tout $\mathbf{x} \in E$, la suite $(f_n(\mathbf{x}))_n$ est une suite de Cauchy dans $(E, \|\cdot\|)$. L'espace vectoriel normé $(E, \|\cdot\|)$ étant complet, la suite $(f_n(\mathbf{x}))_n$ converge dans $(E, \|\cdot\|)$ vers un élément $\ell_{\mathbf{x}}$ de E.

---

[71] Voir la définition 7.20 p. 398.

[72] Voir p. 579.

2 - a) Montrons que l'application $f : \mathbf{x} \in E \mapsto \ell_{\mathbf{x}} \in E$ est un endomorphisme. Soient $\mathbf{y}$ et $\mathbf{z}$ deux éléments de E et λ, μ deux scalaires. Pour tout $n \in \mathbb{N}$, l'application f_n étant un endomorphisme de E on a

$$f_n(\lambda\mathbf{y} + \mu\mathbf{z}) = \lambda f_n(\mathbf{y}) + \mu f_n(\mathbf{z})$$

Comme pour tout $\mathbf{x} \in E$, la suite $(f_n(\mathbf{x}))_n$ converge dans $(E, \|\cdot\|)$ vers $\ell_{\mathbf{x}} = f(\mathbf{x})$, la suite de terme général $f_n(\lambda\mathbf{y} + \mu\mathbf{z})$ converge vers $f(\lambda\mathbf{y} + \mu\mathbf{z})$. On déduit par ailleurs des propriétés de la limite que la suite de terme général $\lambda f_n(\mathbf{y}) + \mu f_n(\mathbf{z})$ converge vers $\lambda f(\mathbf{y}) + \mu f(\mathbf{z})$. Par unicité de la limite d'une suite on obtient

$$f(\lambda\mathbf{y} + \mu\mathbf{z}) = \lambda f(\mathbf{y}) + \mu f(\mathbf{z})$$

ce qui permet de conclure que f est une application linéaire de E dans E.

b) Montrons maintenant que f est un endomorphisme continu. Pour tout $\mathbf{x} \in E$ on obtient d'après (22) considérée avec $\varepsilon' = \|\mathbf{x}\|$,

$$\exists N \in \mathbb{N} \quad \forall (m,n) \in \mathbb{N}^2 \quad \big(n \geqslant m \geqslant N \implies \|f_n(\mathbf{x}) - f_m(\mathbf{x})\| \leqslant \|\mathbf{x}\|\big). \tag{23}$$

Puisque la suite de terme général $f_n(\mathbf{x})$ converge vers $f(\mathbf{x})$, pour tout $m \in \mathbb{N}$, la suite de terme général $f_n(\mathbf{x}) - f_m(\mathbf{x})$ converge vers $f(\mathbf{x}) - f_m(\mathbf{x})$. La norme $\|\cdot\|$ sur E étant une application continue[73] sur E, on a

$$\lim_{n\to+\infty} \|f_n(\mathbf{x}) - f_m(\mathbf{x})\| = \|f(\mathbf{x}) - f_m(\mathbf{x})\|.$$

On déduit de (23) (en considérant $N = \mathrm{E}(N) + 1$) que

$$\exists N \in \mathbb{N}^* \quad \forall m \in \mathbb{N} \quad \big(m \geqslant N \implies \|f(\mathbf{x}) - f_m(\mathbf{x})\| \leqslant \|\mathbf{x}\|\big) \tag{24}$$

puis en utilisant successivement l'inégalité triangulaire, la relation (24) et le fait que f_N est une application linéaire continue, que

$$\begin{aligned} \|f(\mathbf{x})\| &\leqslant \|f(\mathbf{x}) - f_N(\mathbf{x})\| + \|f_N(\mathbf{x})\| \leqslant \|\mathbf{x}\| + \|f_N(\mathbf{x})\| \\ &\leqslant \|\mathbf{x}\| + |||f_N||| \, \|\mathbf{x}\| = (1 + |||f_N|||)\,\|\mathbf{x}\|. \end{aligned}$$

La quatrième assertion du théorème 9.3 permet de conclure à la continuité de f sur E.

3 - Puisque pour tout $n \in \mathbb{N}$, f_n est une application linéaire, on a $\|f_n(\mathbf{x}) - f_m(\mathbf{x})\| = \|(f_n - f_m)(\mathbf{x})\| \leqslant |||f_n - f_m||| \, \|\mathbf{x}\|$ ce qui d'après l'assertion (21) implique que

$$\begin{aligned} \forall \varepsilon \in \mathbb{R}_+^* \quad \exists N \in \mathbb{N} \quad \forall (m,n) \in \mathbb{N}^2 \quad \forall \mathbf{x} \in E \\ \big(n \geqslant m \geqslant N \implies \|f_n(\mathbf{x}) - f_m(\mathbf{x})\| \leqslant \varepsilon \|\mathbf{x}\|\big). \end{aligned} \tag{25}$$

On a établi à la question précédente que pour tout $\mathbf{x} \in E$,

$$\lim_{n\to+\infty} \|f_n(\mathbf{x}) - f_m(\mathbf{x})\| = \|f(\mathbf{x}) - f_m(\mathbf{x})\|.$$

[73] Voir la proposition 9.22 p. 572.

En considérant la limite quand n tend vers $+\infty$ dans l'assertion (25) on obtient :

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N \in \mathbb{N} \quad \forall m \in \mathbb{N} \quad \forall \mathbf{x} \in E \quad \big(m \geqslant N \implies \|f(\mathbf{x}) - f_m(\mathbf{x})\| \leqslant \varepsilon \|\mathbf{x}\|\big).$$

Puisque $|||f - f_m||| = \sup\limits_{\substack{\mathbf{x}\in E \\ \mathbf{x}\neq 0_E}} \dfrac{\|f(\mathbf{x}) - f_m(\mathbf{x})\|}{\|\mathbf{x}\|}$, on en déduit que

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N \in \mathbb{N} \quad \forall m \in \mathbb{N} \quad \big(m \geqslant N \implies |||f - f_m||| \leqslant \varepsilon\big),$$

ce qui signifie que la suite $(f_n)_n$ converge vers f dans $(\mathcal{L}_c(E), |||\cdot|||)$.

Solution de l'exercice 9.10

L'application f^n étant une application de E dans E contractante, d'après le théorème 9.6 elle admet un unique point fixe ζ dans E. Puisque $f^n(\zeta) = \zeta$, en composant avec l'application f, on en déduit que $f^{n+1}(\zeta) = f(\zeta)$. Or $f^{n+1}(\zeta) = f^n(f(\zeta))$, ce qui implique que $f^n(f(\zeta)) = f(\zeta)$, autrement dit que $f(\zeta)$ est un point fixe de f^n. Comme ζ est l'unique point fixe de f^n, on a $f(\zeta) = \zeta$. Cette égalité indique que ζ est un point fixe de f.

Supposons à présent que f possède un second point fixe ζ'. La relation $f(\zeta') = \zeta'$ implique que $f^2(\zeta') = f(\zeta') = \zeta'$, puis par récurrence que pour tout $p \in \mathbb{N}^*$, $f^p(\zeta') = \zeta'$. L'élément ζ' est alors également point fixe de f^n. Comme f^n admet un unique point fixe, on a nécessairement $\zeta' = \zeta$. On en conclut que f admet un unique point fixe.

Solution de l'exercice 9.11

L'espace vectoriel $E = \mathcal{C}^0([0,1], \mathbb{R})$ muni de la norme $\|\cdot\|_\infty$ est complet. Montrons que ϕ est une application contractante de $(E, \|\cdot\|_\infty)$ dans $(E, \|\cdot\|_\infty)$. Pour $g_1, g_2 \in \mathcal{C}^0([0,1], \mathbb{R})$, on a

$$\begin{aligned}
\|\phi(g_1) - \phi(g_2)\|_\infty &= \sup_{x\in[0,1]} |\phi(g_1)(x) - \phi(g_2)(x)| \\
&= \sup_{x\in[0,1]} \left| \int_0^1 k(x,y)\, g_1(y)\, \mathrm{d}y - \int_0^1 k(x,y)\, g_2(y)\, \mathrm{d}y \right| \\
&= \sup_{x\in[0,1]} \left| \int_0^1 k(x,y)\, (g_1(y) - g_2(y))\, \mathrm{d}y \right|.
\end{aligned}$$

Or,

$$\begin{aligned}
\left| \int_0^1 k(x,y)\, (g_1(y) - g_2(x))\, \mathrm{d}y \right| &\leqslant \int_0^1 |k(x,y)|\ |g_1(y) - g_2(y))|\ \mathrm{d}y \\
&\leqslant \int_0^1 \Big(\sup_{(x',y')\in[0,1]^2} |k(x',y')| \times \sup_{y'\in[0,1]} |g_1(y') - g_2(y'))| \Big)\, \mathrm{d}y \\
&= \sup_{(x',y')\in[0,1]^2} |k(x',y')| \times \sup_{y'\in[0,1]} |g_1(y') - g_2(y'))| \\
&= \frac{1}{2} \|g_1 - g_2\|_\infty.
\end{aligned}$$

On en déduit que ϕ est contractante de rapport $1/2$ et par conséquent que ϕ admet un unique point fixe g dans $(E, \|\cdot\|_\infty)$. L'application g est l'unique application continue sur $[0,1]$ telle que $\phi(g) = g$, autrement dit telle que pour tout $x \in [0,1]$ on ait $\phi(g)(x) = g(x)$. L'équation de Volterra admet donc une unique solution dans E.

L'espace $\mathcal{C}^0([0,1], \mathbb{R})$ muni de la norme $\|\cdot\|_1$ n'étant pas complet (voir p. 404), il est indispensable de choisir la norme $\|\cdot\|_\infty$ pour pouvoir appliquer le théorème du point fixe à ϕ.

Solution de l'exercice 9.12

1 - a) Soit $(\mathbf{y}_n)_n$ une suite d'éléments de E convergeant vers $\mathbf{0}_E$; pour tout $n \in \mathbb{N}$, on a $\mathbf{y}_n = \sqrt{\|\mathbf{y}_n\|_E}\,\mathbf{u}_n$ où la suite $(\mathbf{u}_n)_n$ est la suite d'éléments de E définie par

$$\forall n \in \mathbb{N} \quad \mathbf{u}_n = \begin{cases} \mathbf{y}_n / \sqrt{\|\mathbf{y}_n\|_E} & \text{si } \|\mathbf{y}_n\| \neq 0 \\ 0 & \text{si } \|\mathbf{y}_n\| = 0 \end{cases}.$$

Pour tout $n \in \mathbb{N}$ on a $\|\mathbf{u}_n\|_E = \sqrt{\|\mathbf{y}_n\|_E}$ (cette égalité a lieu dans les deux cas correspondant à la définition de la suite $(\mathbf{u}_n)_n$, car si $\|\mathbf{y}_n\|_E = 0$ alors $\|\mathbf{u}_n\|_E = 0 = \sqrt{\|\mathbf{y}_n\|_E}$). Comme la suite $(\mathbf{y}_n)_n$ converge vers $\mathbf{0}_E$, on déduit de la relation $\|\mathbf{u}_n\|_E = \sqrt{\|\mathbf{y}_n\|_E}$ que la suite $(\mathbf{u}_n)_n$ converge aussi vers 0. D'après les hypothèses, cela implique que la suite $(f(\mathbf{u}_n))_n$ est bornée[74], *i.e.* :

$$\exists K \in \mathbb{R}_+^* \quad \forall n \in \mathbb{N} \quad \|f(\mathbf{u}_n)\|_F \leqslant K.$$

L'application f étant linéaire, on a

$$0 \leqslant \|f(\mathbf{y}_n)\|_F = \|f(\|\mathbf{y}_n\|_E^{\frac{1}{2}}\,\mathbf{u}_n)\|_F = \|\mathbf{y}_n\|_E^{\frac{1}{2}}\,\|f(\mathbf{u}_n)\|_F \leqslant K\|\mathbf{y}_n\|_E^{\frac{1}{2}}.$$

Comme la suite $(\mathbf{y}_n)_n$ converge vers $\mathbf{0}_E$, on a $\lim_{n\to+\infty} \|\mathbf{y}_n\|_E^{\frac{1}{2}} = 0$ et le théorème d'encadrement permet de conclure que $\lim_{n\to+\infty} \|f(\mathbf{y}_n)\|_F = 0$.

b) On a ainsi établi que pour toute suite $(\mathbf{x}_n)_n$ d'éléments de E convergeant vers $\mathbf{0}_E$, la suite de terme général $f(\mathbf{x}_n)$ converge vers $\mathbf{0}_F$. Or comme f est une application linéaire, on a $\mathbf{0}_F = f(\mathbf{0}_E)$. D'après la proposition 9.8 p. 549, on en conclut que f est continue en $\mathbf{0}_E$ ce qui d'après le théorème 9.3 implique que f est une application continue sur E.

---

[74] Voir la définition 7.17 p. 392.

On a démontré le résultat suivant.

> Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux $\mathbb{K}$-espaces vectoriels normés et f une application linéaire de E dans F. Si pour toute suite $(\mathbf{x}_n)_n$ d'éléments de E convergeant pour $\|\cdot\|_E$ vers $\mathbf{0}_E$, la suite de terme général $f(\mathbf{x}_n)$ est bornée dans $(F, \|\cdot\|_F)$ alors f est une application linéaire continue.

2 - a) Considérons la suite $(h_n)_n$ d'éléments de $\mathcal{C}^0([0,1], \mathbb{R})$ définie par

$$\forall n \in \mathbb{N} \qquad h_n : x \in [0,1] \longmapsto \begin{cases} 0 & \text{si } 0 \leqslant x \leqslant 1 - \frac{1}{n} \\ n^{\frac{3}{2}} x + \sqrt{n}(1-n) & \text{si } 1 - \frac{1}{n} < x \leqslant 1 \end{cases}.$$

La suite $(h_n)_n$ converge pour la norme $\|\cdot\|_1$ vers la fonction nulle puisque[75]

$$\|h_n\|_1 = \int_0^1 |h_n(x)| \, \mathrm{d}x = \int_{1-\frac{1}{n}}^1 |h_n(x)| \, \mathrm{d}x = \frac{1}{2\sqrt{n}}$$

et par conséquent, $\lim_{n \to +\infty} \|h_n\|_1 = 0$.

Considérons la suite de terme général $f(h_n)$; pour tout $n \in \mathbb{N}$, on a $f(h_n) = h_n(1) = \sqrt{n}$. Cette suite n'étant pas bornée[76] dans $\mathbb{R}$, on déduit de la question précédente que f n'est pas une application continue de $\mathcal{C}^0([0,1], \mathbb{R})$ dans $\mathbb{R}$ pour la norme $\|\cdot\|_1$.

b) Considérons une suite $(h_n)_n$ d'éléments de $\mathcal{C}^0([0,1], \mathbb{R})$ convergeant vers la fonction nulle pour la norme $\|\cdot\|_\infty$, *i.e.* telle que $\lim_{n \to +\infty} \|h_n\|_\infty = 0$. Pour tout $n \in \mathbb{N}$, on a

$$0 \leqslant |f(h_n)| = |h_n(1)| \leqslant \sup_{x \in [0,1]} |h_n(x)| = \|h_n\|_\infty.$$

D'après le théorème d'encadrement, la suite de terme général $f(h_n)$ converge vers 0 dans $\mathbb{R}$ et par conséquent est bornée[77]. On déduit de la question 1 que f est une application continue de $\mathcal{C}^0([0,1], \mathbb{R})$ dans $\mathbb{R}$ pour la norme $\|\cdot\|_1$.

Solution de l'exercice 9.13

1 - Pour toute fonction y de classe $\mathcal{C}^1$ sur $[0,T]$, on a

$$\forall t \in [0,T] \qquad y(t) = y(0) + \int_0^t y'(s) \, \mathrm{d}s.$$

On en déduit que si y est solution du problème $(\mathcal{P})$ alors

$$\forall t \in [0,T] \qquad y(t) = \sqrt{1 + y(T)} + \int_0^t b(s) \, \mathrm{d}s.$$

---

[75] Remarquer que l'intégrale à calculer correspond à l'aire d'un triangle de base $1/n$ et de hauteur $\sqrt{n}$.

[76] Dans l'espace vectoriel $\mathbb{R}$ qui est de dimension finie, toutes les normes étant équivalentes, il est inutile de préciser la norme choisie.

[77] Voir la proposition 7.14 p. 394.

Réciproquement, l'application $y : t \in [0,T] \mapsto \sqrt{1+y(T)} + \int_0^t b(s)\,\mathrm{d}s$ prend la valeur $\sqrt{1+y(T)}$ et a pour dérivée b ; elle est donc de classe $\mathcal{C}^1$ sur $[0,T]$ (puisque b est continue) et est solution du problème $(\mathcal{P})$.

2 - Comme $\mathcal{C}^0([0,T],\mathbb{R})$ muni de la norme uniforme est complet (voir p. 403) pour montrer que A est complet, il suffit de montrer qu'il est fermé dans $\mathcal{C}^0([0,T],\mathbb{R})$ pour la norme uniforme (voir la proposition 7.19 p. 412).

Considérons une suite $(f_n)_n$ d'éléments de A qui converge dans $\mathcal{C}^0([0,T],\mathbb{R})$ pour la norme uniforme. D'après le théorème 7.6, montrer que A est fermé revient à montrer que la limite f de la suite $(f_n)_n$ appartient à A, autrement dit que f est continue et positive sur $[0,T]$. Puisque la suite $(f_n)_n$ converge vers f pour la norme uniforme, la suite $(f_n)_n$ converge uniformément vers f sur $[0,T]$. La continuité des termes de la suite sur $[0,T]$ implique la continuité de f sur $[0,T]$ (voir le théorème 2.3 p. 86). Pour tout $t \in [0,T]$, la suite numérique réelle de terme général $f_n(t)$ est une suite à termes positifs qui converge. Sa limite qui est le réel $f(t)$ est positive. La fonction f est donc bien à valeurs positives. On en conclut que A est fermé.

3 - Considérons l'application

$$\Phi : f \in A \longmapsto \Phi(f) \quad \text{où} \quad \Phi(f) : t \in [0,T] \longmapsto \sqrt{1+f(T)} + \int_0^t b(s)\,\mathrm{d}s.$$

Comme b est une fonction continue et positive, son intégrale indéfinie $t \in [0,T] \mapsto \int_0^t b(s)\,\mathrm{d}s$ est une fonction positive de classe $\mathcal{C}^1$. On a donc $\Phi(A) \subset A$. Pour appliquer le théorème du point fixe, il reste à établir que Φ est une application contractante sur A. Soient f_1 et f_2 deux fonctions appartenant à A ; on a

$$\begin{aligned}
\|\Phi(f_1) - \Phi(f_2)\|_\infty &= \sup_{t\in[0,T]} |\Phi(f_1)(t) - \Phi(f_2)(t)| \\
&= \sup_{t\in[0,T]} \left|\sqrt{1+f_1(T)} - \sqrt{1+f_2(T)}\right| \\
&= \left|\sqrt{1+f_1(T)} - \sqrt{1+f_2(T)}\right| \\
&= \frac{|f_1(T) - f_2(T)|}{\sqrt{1+f_1(T)} + \sqrt{1+f_2(T)}}
\end{aligned}$$

Comme f_1 et f_2 sont à valeurs positives, on a $\sqrt{1+f_1(T)} + \sqrt{1+f_2(T)} \geqslant 2$ et par conséquent

$$\begin{aligned}
\|\Phi(f_1) - \Phi(f_2)\|_\infty &\leqslant \frac{1}{2}\,|f_1(T) - f_2(T)| \leqslant \tfrac{1}{2} \sup_{t\in[0,T]} |f_1(t) - f_2(t)| \\
&= \tfrac{1}{2}\,\|f_1 - f_2\|_\infty.
\end{aligned}$$

Ainsi Φ est contractante de rapport $1/2$ et le théorème du point fixe assure qu'il existe une unique fonction $y \in A$ telle que $\Phi(y) = y$. On en conclut qu'il existe une unique fonction continue et positive sur $[0,T]$ qui soit solution du problème $(\mathcal{P})$.

Solution de l'exercice 9.14

1 - Notons $\mathcal{C}^0_+([0,1],\mathbb{R})$ l'ensemble des applications positives et continues sur $[0,1]$. Un exemple d'application linéaire de $\mathcal{C}^0_+([0,1],\mathbb{R})$ dans $\mathbb{R}$ qui vérifie la propriété $(\mathcal{P})$ est :

$$f : h \in \mathcal{C}^0_+([0,1],\mathbb{R}) \longmapsto \int_0^1 h(x)\,\mathrm{d}x$$

puisque l'intégrale d'une fonction Riemann-intégrable et positive est un réel positif[78] et que par ailleurs l'image de la fonction constante égale à 1 par f est bien égale à 1. Un autre exemple est l'application

$$f : h \in \mathcal{C}^0_+([0,1],\mathbb{R}) \longmapsto \sup_{x\in[0,1]} h(x).$$

En effet, une fonction continue sur un intervalle fermé et borné étant bornée et atteignant ses bornes[79], il existe $c \in [0,1]$ tel que $f(h) = \sup_{x\in[0,1]} h(x) = h(c) \geqslant 0$. De plus, l'image de la fonction constante égale à 1 par f vaut bien 1.

2 - a) Montrons que si f est une forme linéaire sur $\mathcal{C}^0([0,1],\mathbb{R})$ vérifiant la propriété $(\mathcal{P})$ alors f est une application croissante en établissant l'assertion suivante :

$$\forall (h_1,h_2) \in \mathcal{C}^0([0,1],\mathbb{R})^2 \quad \big(h_1 \leqslant h_2 \implies f(h_1) \leqslant f(h_2)\big).$$

Soient $h_1, h_2 \in \mathcal{C}^0([0,1],\mathbb{R})$ telles que $h_1 \leqslant h_2$; l'application $h = h_2 - h_1$ est une application continue et positive sur $[0,1]$. D'après la propriété $(\mathcal{P})$, on en déduit que $f(h) \geqslant 0$. Or, puisque f est une application linéaire, $f(h) = f(h_2) - f(h_1)$, autrement dit $f(h_1) \leqslant f(h_2)$. L'application f est bien croissante.

b) Montrer que pour tout $h \in \mathcal{C}^0([0,1],\mathbb{R})$ on a $|f(h)| \leqslant f(|h|)$ revient à montrer que $-f(|h|) \leqslant f(h) \leqslant f(|h|)$. Soit $h \in \mathcal{C}^0([0,1],\mathbb{R})$; on a $h \leqslant |h|$ car si $x \in [0,1]$ est tel que $h(x) \geqslant 0$ alors $h(x) = |h(x)|$ et si $x \in [0,1]$ est tel que $h(x) \geqslant 0$ alors $h(x) < 0$ et $|h(x)| > 0$. Comme f est une application croissante, voir la question précédente, on en déduit que $f(h) \leqslant f(|h|)$. De la même manière, on a $h \geqslant -|h|$, donc comme f est croissante on a $f(h) \geqslant f(-|h|)$ et f étant linéaire $f(-|h|) = -f(|h|)$. Le résultat est établi.

Comme $\|h\|_\infty = \sup_{x\in[0,1]} |h(x)|$, on a pour tout $x \in [0,1]$, $|h(x)| \leqslant \|h\|_\infty \, 1$. Puisque f est croissante, on en déduit que $f(|h|) \leqslant f(\|h\|_\infty \, 1)$. Or f étant linéaire, on a $f(\|h\|_\infty \, 1) = \|h\|_\infty \, f(1)$. Le résultat est établi.

Remarquons que la conjonction des deux résultats qui viennent d'être établis indique que pour tout $h \in \mathcal{C}^0([0,1],\mathbb{R})$ on a $|f(h)| \leqslant C\,\|h\|_\infty$ où $C = f(1)$.

3 - L'application f est une forme linéaire sur $\mathcal{C}^0([0,1],\mathbb{R})$ qui, d'après la question précédente, vérifie

$$\forall h \in \mathcal{C}^0([0,1],\mathbb{R}) \quad |f(h)| \leqslant C\,\|h\|_\infty$$

---

[78] Voir la proposition 18.4 p. 844 du *Cours de première année*.

[79] Voir le théorème 13.4 p. 596 du *Cours de première année*.

où $C = f(1)$. Cela permet d'affirmer d'après le théorème 9.3 que f une application continue de $(\mathcal{C}^0([0,1],\mathbb{R}), \|\cdot\|_\infty)$ dans $(\mathbb{R}, |\cdot|)$. De plus,

$$|||f||| = \sup_{\substack{h\in\mathcal{C}^0([0,1],\mathbb{R}) \\ h\neq 0}} \frac{|f(h)|}{\|h\|_\infty} \leqslant C = f(1).$$

Remarquons que si $h = 1$ alors $\|h\|_\infty = 1$ et par conséquent $\dfrac{|f(h)|}{\|h\|_\infty} = f(1)$. On en déduit que

$$|||f||| = \sup_{\substack{h\in\mathcal{C}^0([0,1],\mathbb{R}) \\ h\neq 0}} \frac{|f(h)|}{\|h\|_\infty}, \geqslant f(1)$$

ce qui implique finalement que $|||f||| = f(1)$.

CHAPITRE 10

Calcul différentiel

Nous avons dans le chapitre 9 introduit la notion de continuité pour une application d'un espace vectoriel normé $(E, \|\cdot\|_E)$ dans un espace vectoriel normé $(F, \|\cdot\|_F)$ et avons insisté sur le cas particulier, très important dans la pratique, des fonctions de $\mathbb{R}^n$ dans $\mathbb{R}^p$. Dans le prolongement du chapitre 9, nous allons introduire dans ce chapitre, pour une application d'un espace vectoriel normé $(E, \|\cdot\|_E)$ dans un espace vectoriel normé $(F, \|\cdot\|_F)$ la notion de différentielle, notion qui généralise celle de dérivée dans le cas des fonctions de $\mathbb{R}$ dans $\mathbb{R}$. Nous insisterons une fois encore sur le cas des fonctions de $\mathbb{R}^n$ dans $\mathbb{R}^p$, introduisant les opérateurs différentiels usuels de la physique que sont le gradient, la divergence et le rotationnel ainsi que les résultats essentiels pour les applications que sont le théorème d'inversion locale et la formule de Taylor. Dans ce chapitre, les espaces vectoriels considérés auront tous pour corps de base $\mathbb{R}$.

10.1 Différentiabilité dans un espace vectoriel

10.1.1 Applications différentiables

DÉFINITION 10.1 *Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux espaces vectoriels normés et U un ensemble ouvert dans $(E, \|\cdot\|_E)$. On dit qu'une application f de U dans F est* différentiable *en $\mathbf{a} \in U$ s'il existe une application linéaire continue $L \in \mathcal{L}_c(E, F)$ et une fonction ε de E dans F telles que, sur un voisinage de $\mathbf{a}$ dans U, on ait*

$$f(\mathbf{x}) = f(\mathbf{a}) + L(\mathbf{x} - \mathbf{a}) + \|\mathbf{x} - \mathbf{a}\|_E \, \varepsilon(\mathbf{x}) \tag{1}$$

avec $\lim_{\mathbf{x} \to \mathbf{a}} \varepsilon(\mathbf{x}) = \mathbf{0}_F$.

Remarques

1. Rappelons que la fonction ε de E dans F admet pour limite $\mathbf{0}_F$ en $\mathbf{a} \in E$ si

$$\forall \delta \in \mathbb{R}_+^* \quad \exists \eta \in \mathbb{R}_+^* \quad \forall \mathbf{x} \in E \quad \Big(\|\mathbf{x} - \mathbf{a}\|_E \leqslant \eta \implies \|\varepsilon(\mathbf{x})\|_F \leqslant \delta \Big).$$

2. Lorsque l'espace vectoriel E est de dimension finie, toutes les applications linéaires de E dans F sont continues[1].

3. Toutes les normes sur un espace vectoriel de dimension finie sont équivalentes[2] de sorte que lorsque E ou F sont de dimension finie, on peut utiliser n'importe quelle norme sur ces espaces afin de montrer qu'une application est différentiable. ◇

Exemples

1. Soit f une application de $\mathbb{R}^2$ dans $\mathbb{R}$ constante, c'est-à-dire telle que

$$\exists C \in \mathbb{R} \quad \forall (x_1, x_2) \in \mathbb{R}^2 \quad f(x_1, x_2) = C.$$

L'application f est différentiable en $\mathbf{a}$ pour tout $\mathbf{a} = (a_1, a_2) \in \mathbb{R}^2$. En effet, pour tout $(x_1, x_2) \in \mathbb{R}^2$ on peut écrire :

$$f(x_1, x_2) = f(a_1, a_2) + L(x_1 - a_1, x_2 - a_2) + \|(x_1 - a_1, x_2 - a_2)\|_{\mathbb{R}^2}\, \varepsilon(x_1, x_2),$$

où L est l'application linéaire nulle de $\mathbb{R}^2$ dans $\mathbb{R}$, $\|\cdot\|$ désigne une norme quelconque de $\mathbb{R}^2$ et $\varepsilon : (x_1, x_2) \in \mathbb{R}^2 \mapsto 0 \in \mathbb{R}$ tend vers 0 quand (x_1, x_2) tend vers $(0, 0)$.

2. L'application $f : (x_1, x_2) \in \mathbb{R}^2 \mapsto x_1^2 + 2(x_1 + x_2) + x_2^2 \in \mathbb{R}$ est différentiable en $(0, 0)$. En effet, puisque $f(0, 0) = 0$, on a pour tout $(x_1, x_2) \in \mathbb{R}^2$

$$f(x_1, x_2) = f(0, 0) + 2(x_1 + x_2) + x_1^2 + x_2^2.$$

En posant $L : (x_1, x_2) \in \mathbb{R}^2 \mapsto 2(x_1 + x_2)$ et $\varepsilon : (x_1, x_2) \in \mathbb{R}^2 \mapsto \sqrt{x_1^2 + x_2^2}$, on obtient

$$f(x_1, x_2) = f(0, 0) + L(x_1, x_2) + \|(x_1, x_2)\|_2\, \varepsilon(x_1, x_2),$$

où L est une application linéaire de $\mathbb{R}^2$ dans $\mathbb{R}$ et où pour tout $(x_1, x_2) \in \mathbb{R}^2$, $\varepsilon(x_1, x_2) = \|(x_1, x_2)\|_2$ tend vers 0 quand (x_1, x_2) tend vers $(0, 0)$.

Remarques

1. En utilisant les notations de Landau[3], la relation (1) s'écrit aussi

$$f(\mathbf{x}) = f(\mathbf{a}) + L(\mathbf{x} - \mathbf{a}) + o_{\mathbf{a}}(\|\mathbf{x} - \mathbf{a}\|_E).$$

---

[1] Voir le théorème 9.4 p. 582.

[2] Voir le théorème 7.2 p. 367.

[3] Voir le *Cours de première année*, chap. 15.

2. On peut aussi reformuler la définition 10.1 de la manière suivante : une application f est différentiable en $\mathbf{a} \in U$ s'il existe une application linéaire continue $L \in \mathcal{L}_c(E, F)$ telle que pour tout $\mathbf{h}$ au voisinage de $\mathbf{0}_E$, on ait :

$$f(\mathbf{a}+\mathbf{h}) = f(\mathbf{a}) + L(\mathbf{h}) + o_{\mathbf{0}}(\|\mathbf{h}\|_E).$$

◇

Exemple Montrons que l'application $f : (x_1, x_2) \in \mathbb{R}^2 \mapsto x_1^2 + 2(x_1 + x_2) + x_2^2$ est différentiable en un point $\mathbf{a} = (a_1, a_2)$ quelconque de $\mathbb{R}^2$ en exploitant la remarque précédente. Dans un premier temps, formons la différence entre les termes $f(a_1 + h_1, a_2 + h_2)$ et $f(a_1, a_2)$:

$$f(a_1+h_1, a_2+h_2) - f(a_1, a_2) = 2a_1h_1 + h_1^2 + 2h_1 + 2h_2 + 2a_2h_2 + h_2^2.$$

Dans cette expression, on identifie les termes qui sont « linéaires » en la variable $\mathbf{h} = (h_1, h_2)$ et on pose

$$L : (h_1, h_2) \in \mathbb{R}^2 \mapsto 2a_1h_1 + 2h_1 + 2h_2 + 2a_2h_2.$$

On en déduit que

$$f(a_1+h_1, a_2+h_2) - f(a_1, a_2) = L(h_1, h_2) + \|(h_1, h_2)\|_2 \, \varepsilon(h_1, h_2),$$

où $\varepsilon : (h_1, h_2) \in \mathbb{R}^2 \mapsto \sqrt{h_1^2 + h_2^2}$ tend vers 0 lorsque (h_1, h_2) tend vers $(0, 0)$.

PROPOSITION 10.1 *Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux espaces vectoriels normés, U un ensemble ouvert dans $(E, \|\cdot\|_E)$ et f une application de U dans F. Si f est différentiable en $\mathbf{a} \in U$ alors l'application linéaire L de la définition 10.1 est unique. Elle est appelée* différentielle *de f en $\mathbf{a}$ et est notée* $\mathrm{d}f_{\mathbf{a}}$.

Démonstration Supposons que l'on dispose de deux applications L_1 et L_2 satisfaisant la définition 10.1. Dans un voisinage $V_{\mathbf{a}}$ de $\mathbf{a}$, on a[(4)]

$$f(\mathbf{x}) = f(\mathbf{a}) + L_1(\mathbf{x}-\mathbf{a}) + \|\mathbf{x}-\mathbf{a}\|_E \, \varepsilon_1(\mathbf{x}),$$
$$f(\mathbf{x}) = f(\mathbf{a}) + L_2(\mathbf{x}-\mathbf{a}) + \|\mathbf{x}-\mathbf{a}\|_E \, \varepsilon_2(\mathbf{x}),$$

où ε_1 et ε_2 sont deux applications de E dans F de limite nulle en $\mathbf{a}$. L'application linéaire $L = L_1 - L_2$ vérifie pour tout $\mathbf{x} \in V_{\mathbf{a}}$:

$$L(\mathbf{x}-\mathbf{a}) = \|\mathbf{x}-\mathbf{a}\|_E \, \varepsilon(\mathbf{x}) \quad \text{où} \quad \varepsilon(\mathbf{x}) = \varepsilon_1(\mathbf{x}) - \varepsilon_2(\mathbf{x}).$$

(4) L'intersection de deux voisinage de $\mathbf{a}$ étant un voisinage de $\mathbf{a}$, on peut considérer ces deux relations sur un même voisinage de $\mathbf{a}$.

Soit $\mathbf{u} \in U$ fixé ; en considérant des éléments $\mathbf{x}$ de la forme $\mathbf{x} = \mathbf{a} + t\mathbf{u}$ avec $t \in \mathbb{R}$ tels que $\mathbf{x} \in V_{\mathbf{a}}$, la relation précédente s'écrit :

$$L(t\mathbf{u}) = \|t\mathbf{u}\|_E \, \varepsilon(t\mathbf{u} + \mathbf{a}).$$

D'une part, l'application L étant linéaire, on a $L(t\mathbf{u}) = tL(\mathbf{u})$ et d'autre part[5] $\|t\mathbf{u}\|_E = |t| \, \|\mathbf{u}\|_E$. Pour tout réel t non nul, on a donc

$$L(\mathbf{u}) = \frac{|t|}{t} \|\mathbf{u}\|_E \, \varepsilon(t\mathbf{u} + \mathbf{a}).$$

Or, $\lim\limits_{t \to 0} \mathbf{a} + t\mathbf{u} = \mathbf{a}$ et $\lim\limits_{\mathbf{x} \to \mathbf{a}} \varepsilon(\mathbf{x}) = \mathbf{0}_F$ donc $\lim\limits_{t \to 0} \varepsilon(\mathbf{a} + t\mathbf{u}) = \mathbf{0}_F$. On a par conséquent

$$\lim_{t \to 0^+} \frac{|t|}{t} \|\mathbf{u}\|_E \, \varepsilon(t\mathbf{u} + \mathbf{a}) = \mathbf{0}_F \quad \text{et} \quad \lim_{t \to 0^-} \frac{|t|}{t} \|\mathbf{u}\|_E \, \varepsilon(t\mathbf{u} + \mathbf{a}) = \mathbf{0}_F.$$

On en déduit que $L(\mathbf{u}) = \mathbf{0}_F$ pour tout $\mathbf{u} \in U$. L'application L est donc l'application linéaire nulle sur U, et donc[6] sur E. On en conclut que $L_1 = L_2$, autrement dit que l'application linéaire L de la définition 10.1 est unique. □

PROPOSITION 10.2 *Soient f une application définie sur un ouvert U de E à valeurs dans F et $\mathbf{a}$ un élément de U. Si f est différentiable en $\mathbf{a}$ alors elle est continue en $\mathbf{a}$.*

Démonstration Si f est différentiable en $\mathbf{a}$ alors, d'après la définition 10.1, il existe une application linéaire continue L de E dans F telle que pour $\mathbf{x}$ au voisinage de $\mathbf{a}$, on ait :

$$f(\mathbf{x}) = f(\mathbf{a}) + L(\mathbf{x} - \mathbf{a}) + \|\mathbf{x} - \mathbf{a}\|_E \, \varepsilon(\mathbf{x}) \quad \text{et} \quad \lim_{\mathbf{x} \to \mathbf{a}} \varepsilon(\mathbf{x}) = \mathbf{0}_F.$$

L'application linéaire L étant continue sur E, on a $\lim\limits_{\mathbf{x} \to \mathbf{a}} L(\mathbf{x} - \mathbf{a}) = L(\mathbf{0}_E) = \mathbf{0}_F$ et par conséquent $\lim\limits_{\mathbf{x} \to \mathbf{a}} f(\mathbf{x}) = f(\mathbf{a})$. Cela exprime la continuité de f en $\mathbf{a}$. □

DÉFINITION 10.2

✗ *Une application f définie sur un ouvert U de E à valeurs dans F est dite* différentiable sur U *si elle est différentiable en tout point de U.*

✗ *Soit f une application définie sur un ouvert U de E à valeurs dans F, différentiable sur U. On appelle* différentielle de f *et on note* $\mathrm{d}f$ *l'application*

$$\begin{aligned} \mathrm{d}f : U &\longrightarrow \mathcal{L}_c(E, F) \\ \mathbf{a} &\longmapsto \mathrm{d}f_{\mathbf{a}} \end{aligned}$$

[5] Voir la définition 7.3.

[6] Une application linéaire L d'un espace vectoriel E dans un espace vectoriel F qui est nulle sur un ouvert U non vide de E est nécessairement nulle sur E, voir l'exercice 9.4 p. 573.

On déduit de la proposition 10.2 que si une application f d'un ouvert U de E à valeurs dans F est différentiable sur U alors elle est continue sur U.

Exemple Considérons l'application[7] $\Phi : M \in \mathrm{GL}_n(\mathbb{R}) \mapsto M^{-1} \in \mathrm{GL}_n(\mathbb{R})$ qui à une matrice inversible de $\mathcal{M}_n(\mathbb{R})$ associe son inverse. Pour une matrice $H \in \mathcal{M}_n(\mathbb{R})$ telle que $M + H \in \mathrm{GL}_n(\mathbb{R})$, commençons par remarquer que

$$M\big((M+H)^{-1} - M^{-1}\big)(M+H) = M - (M+H) = -H$$

d'où

$$(M+H)^{-1} - M^{-1} = -M^{-1}H(M+H)^{-1}.$$

On a donc $(M+H)^{-1} = M^{-1} - M^{-1}H(M+H)^{-1}$ de sorte que

$$\begin{aligned}\Phi(M+H) - \Phi(M) &= (M+H)^{-1} - M^{-1} \\ &= -M^{-1}HM^{-1} + M^{-1}HM^{-1}H(M+H)^{-1}.\end{aligned}$$

Vérifions que cette relation est formée d'un « terme linéaire en H » et d'un « terme petit par rapport à H ». Posons[8] :

$$L : H \in \mathcal{M}_n(\mathbb{R}) \longmapsto -M^{-1}HM^{-1} \quad \text{et}$$

$$\varepsilon : H \in \mathcal{M}_n(\mathbb{R}) \longmapsto \begin{cases} \dfrac{1}{\|H\|} M^{-1}HM^{-1}H(M+H)^{-1} & \text{si } H \neq 0, \\ 0 & \text{si } H = 0. \end{cases}$$

D'une part, l'application L est linéaire puisque $\forall (\lambda, \mu) \in \mathbb{R}^2\ \forall (A, B) \in \mathcal{M}_n(\mathbb{R})^2$,

$$\begin{aligned}L(\lambda A + \mu B) &= -M^{-1}(\lambda A + \mu B)M^{-1} = -\lambda M^{-1}AM^{-1} - \mu M^{-1}BM^{-1} \\ &= \lambda L(A) + \mu L(B)\end{aligned}$$

et elle est continue car d'après le théorème 9.4 p. 582, en dimension finie toutes les applications linéaires sont continues. D'autre part, on a[8] :

$$0 \leqslant \|\varepsilon(H)\| \leqslant \|M^{-1}\|^2\ \|H\|\ \|(M+H)^{-1}\|.$$

Par le théorème d'encadrement[9], on en déduit que $\varepsilon(H)$ tend vers 0 lorsque H tend vers 0. Ceci prouve que Φ est différentiable sur $\mathrm{GL}_n(\mathbb{R})$ et que sa différentielle $d\Phi$ est

$$d\Phi : M \in \mathrm{GL}_n(\mathbb{R}) \longmapsto d\Phi_M,$$

où $d\Phi_M : H \in \mathcal{M}_n(\mathbb{R}) \longmapsto -M^{-1}HM^{-1}$.

---

[7] L'ensemble $\mathrm{GL}_n(\mathbb{R})$ des matrices inversibles est un ouvert de l'espace vectoriel $\mathcal{M}_n(\mathbb{R})$, voir p. 561.

[8] La norme $\|\cdot\|$ considérée sur l'espace vectoriel $\mathcal{M}_n(\mathbb{R})$ peut être n'importe quelle norme subordonnée à une norme vectorielle sur $\mathbb{R}^n$, c'est-à-dire n'importe quelle norme sur $\mathcal{M}_n(\mathbb{R})$ telle que pour toutes matrices A et B de $\mathcal{M}_n(\mathbb{R})$ on a $\|AB\| \leqslant \|A\|\ \|B\|$. Les normes usuelles sur l'ensemble $\mathcal{M}_n(\mathbb{R})$, voir la proposition 7.5 p. 365, vérifient une telle propriété.

[9] Voir le *Cours de première année* p. 577.

Liens entre la notion de différentiabilité et de dérivabilité pour les fonctions d'une variable réelle

PROPOSITION 10.3 *Soit f une application définie sur un intervalle ouvert I de $\mathbb{R}$ à valeurs dans $\mathbb{R}$. L'application f est différentiable sur I si et seulement si f est dérivable sur I. Dans ce cas, la différentielle de f en $a \in I$ est*

$$\mathrm{d}f_a : h \in \mathbb{R} \longmapsto hf'(a) \in \mathbb{R}.$$

Démonstration $\trianglerighteq$ Supposons que l'application f est différentiable en $a \in I$. Il existe une application linéaire $\mathrm{d}f_a$ de $\mathbb{R}$ dans $\mathbb{R}$ telle que pour h au voisinage de 0 avec $a + h \in I$, on ait

$$f(a+h) = f(a) + \mathrm{d}f_a(h) + |h|\,\varepsilon(h) \quad \text{avec} \quad \lim_{h\to 0} \varepsilon(h) = 0.$$

L'application $\mathrm{d}f_a$ étant linéaire, pour $h \in \mathbb{R}$ on a $\mathrm{d}f_a(h) = h\mathrm{d}f_a(1)$ et le taux d'accroissement de f en a vaut :

$$\frac{f(a+h)-f(a)}{h} = \frac{\mathrm{d}f_a(h)}{h} + \frac{|h|}{h}\varepsilon(h) = \mathrm{d}f_a(1) + \frac{|h|}{h}\varepsilon(h).$$

On en déduit que le taux d'accroissement de f en a admet $\mathrm{d}f_a(1)$ pour limite lorsque h tend vers 0 et par conséquent que l'application f est dérivable en a et de dérivée $f'(a) = \mathrm{d}f_a(1)$.

$\trianglelefteq$ Réciproquement, supposons que f est dérivable en a. On a alors

$$f'(a) = \lim_{h\to 0} \frac{f(a+h)-f(a)}{h},$$

et par conséquent :

$$\lim_{h\to 0} \frac{f(a+h)-f(a)-hf'(a)}{h} = 0.$$

Cela signifie que $f(a+h) - f(a) - hf'(a) = o_0(h)$, autrement dit que l'on a $f(a+h) = f(a) + L(h) + o_0(h)$ où $L : h \in \mathbb{R} \mapsto hf'(a) \in \mathbb{R}$ est une application linéaire. On en déduit que f est différentiable en a de différentielle en a l'application

$$\mathrm{d}f_a : h \in \mathbb{R} \longmapsto hf'(a) \in \mathbb{R}.$$

□

En appliquant la proposition 10.3 à chacune des fonctions coordonnées f_1, $f_2, \ldots, f_p$ d'une fonction f de $\mathbb{R}$ dans $\mathbb{R}^p$, on obtient le résultat suivant.

PROPOSITION 10.4 *Soit f une application définie sur un intervalle ouvert I de $\mathbb{R}$ à valeurs dans $\mathbb{R}^p$. L'application f est différentiable sur I si et seulement si les p fonctions coordonnées f_1, $f_2, \ldots, f_p$ de f sont dérivables sur I. Dans ce cas, la différentielle de f en $a \in I$ est*

$$\mathrm{d}f_a : h \in \mathbb{R} \longmapsto \left(hf_1'(a), hf_2'(a), \ldots, hf_p'(a)\right) \in \mathbb{R}^p.$$

Par la suite, lorsque l'on étudiera une application définie sur $I \subset \mathbb{R}$ et à valeurs dans $\mathbb{R}^p$, on utilisera assez fréquemment la notation f' pour désigner la différentielle $\mathrm{d}f$ de f, *i.e.* on notera

$$f'(a) = (f_1'(a), f_2'(a), \ldots, f_p'(a)) \in \mathbb{R}^p.$$

Différentiabilité des applications linéaires

D'après le théorème 9.4 p. 582, toute application linéaire d'un espace vectoriel normé de dimension finie E à valeurs dans un espace vectoriel normé F est continue. On montre ici un résultat concernant la différentiabilité d'une telle application.

PROPOSITION 10.5 *Soit f une application linéaire d'un espace vectoriel E de dimension finie à valeurs dans un espace vectoriel F. L'application f est différentiable sur E et on a*

$$\forall \mathbf{a} \in E \quad \mathrm{d}f_{\mathbf{a}} = f.$$

Démonstration Puisque l'application f est linéaire, pour tous $\mathbf{a} \in E$ et $\mathbf{h} \in E$ on a $f(\mathbf{a}+\mathbf{h}) = f(\mathbf{a}) + f(\mathbf{h})$. D'après la définition 10.1, cela implique que f est différentiable en $\mathbf{a}$ et que sa différentielle[(10)] est $\mathrm{d}f_{\mathbf{a}} : \mathbf{h} \in E \mapsto f(\mathbf{h}) \in F$. □

Exemples

1. Pour $n \in \mathbb{N}^*$, considérons l'application $f : P \in \mathbb{R}_n[X] \mapsto P' \in \mathbb{R}_{n-1}[X]$ qui à un polynôme associe son polynôme dérivé. Cette application est linéaire et l'espace vectoriel $\mathbb{R}_n[X]$ est de dimension finie. On en déduit que l'application f est différentiable sur $\mathbb{R}_n[X]$ et que, pour tout $P \in \mathbb{R}_n[X]$ on a $\mathrm{d}f_P = f$.

2. Considérons l'application linéaire $f : P \in \mathbb{R}[X] \mapsto P' \in \mathbb{R}[X]$. Contrairement à l'exemple précédent, l'espace vectoriel sur lequel est définie l'application f n'est pas de dimension finie. On va montrer que la différentiabilité de f dépend du choix de la norme $\|\cdot\|$ sur l'espace $\mathbb{R}[X]$. Il résulte de la preuve de la proposition 10.5, que si f est différentiable sur $\mathbb{R}[X]$, sa différentielle en $P \in \mathbb{R}[X]$ est nécessairement $\mathrm{d}f_P = f$. L'application f sera différentiable sur $\mathbb{R}[X]$ si pour tout $P \in \mathbb{R}[X]$, l'application linéaire $\mathrm{d}f_P$ est continue sur $\mathbb{R}[X]$ pour la norme $\|\cdot\|$.

(10) Le fait que E soit de dimension finie nous permet d'affirmer que l'application $\mathrm{d}f_{\mathbf{a}}$ est bien continue.

Munissons l'espace vectoriel $\mathbb{R}[X]$ de la norme $\|\cdot\|_\infty$ définie pour $P = \sum_{k=1}^{n} a_k X^k$ par $\|P\|_\infty = \max_{k\in\{1,\dots,n\}} |a_k|$. Considérons, pour $n \in \mathbb{N}^*$, le polynôme $P_n = X^n$. On a $\|P_n\|_\infty = 1$ et $\|f(P_n)\|_\infty = \|nX^{n-1}\|_\infty = n$. Cela montre qu'il n'existe pas de constante $C \in \mathbb{R}$ telle que pour tout polynôme $P \in \mathbb{R}[X]$ on ait $\|f(P)\|_\infty \leqslant C\|P\|_\infty$. Ainsi, l'application linéaire f n'est pas continue lorsque $\mathbb{R}[X]$ est muni de la norme $\|\cdot\|_\infty$, voir la définition 9.5 p. 546. On en déduit que $\mathrm{d}f_P$ n'est pas continue sur $\mathbb{R}[X]$ muni de la norme $\|\cdot\|_\infty$ et donc que f n'est pas différentiable en $P \in \mathbb{R}[X]$. On peut également conclure en utilisant la proposition 10.2 : f n'étant pas continue sur $\mathbb{R}[X]$ muni de la norme $\|\cdot\|_\infty$, elle ne peut pas être différentiable.

On peut aussi munir l'espace vectoriel $\mathbb{R}[X]$ de la norme $\|\cdot\|_0$ définie par

$$P \in \mathbb{R}[X] \longmapsto |P(0)| + \|P'\|_\infty.$$

Notons que l'application linéaire f de $(\mathbb{R}[X], \|\cdot\|_0)$ dans $(\mathbb{R}[X], \|\cdot\|_\infty)$ est continue puisque pour tout polynôme $P \in \mathbb{R}[X]$ on a[11]

$$\|f(P)\|_\infty = \|P'\|_\infty \leqslant |P(0)| + \|P'\|_\infty = \|P\|_0.$$

Le même raisonnement que celui effectué lors de la preuve de la proposition 10.5 permet d'affirmer que, bien que l'espace vectoriel $\mathbb{R}[X]$ ne soit pas de dimension finie, l'application f de $(\mathbb{R}[X], \|\cdot\|_0)$ dans $(\mathbb{R}[X], \|\cdot\|_\infty)$ est différentiable sur $\mathbb{R}[X]$ et pour tout $P \in \mathbb{R}[X]$ on a $\mathrm{d}f_P = f$.

Exercice 10.1 Soit E un espace vectoriel normé de dimension finie égale à n et $\mathcal{B} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ une base de E.
1 - Déterminer la différentielle d'une forme bilinéaire[12] sur $E \times E$.
2 - Déterminer la différentielle d'une forme quadratique[13] sur E.

Différentiabilité de la norme

On a montré p. 572 que dans un espace vectoriel normé $(E, \|\cdot\|)$, l'application $N : \mathbf{x} \in E \mapsto \|\mathbf{x}\|$ était continue sur $(E, \|\cdot\|)$. On va vérifier que cette application n'est pas différentiable en $\mathbf{0}_E$.

Raisonnons par l'absurde. Si N est différentiable en $\mathbf{0}_E$ alors d'après la définition 10.1, il existe une application linéaire continue $\mathrm{d}N_{\mathbf{0}_E}$ de E dans $\mathbb{R}$, et une application ε de E dans $\mathbb{R}$ de limite nulle en $\mathbf{0}_E$ telles que pour tout $\mathbf{x}$ dans un voisinage de $\mathbf{0}_E$ on ait

$$\|\mathbf{x}\| = \|\mathbf{0}_E\| + \mathrm{d}N_{\mathbf{0}_E}(\mathbf{x}) + \|\mathbf{x}\|\,\varepsilon(\mathbf{x}).$$

En écrivant cette égalité pour $-\mathbf{x}$, on obtient

$$\|-\mathbf{x}\| = \|\mathbf{0}_E\| + \mathrm{d}N_{\mathbf{0}_E}(-\mathbf{x}) + \|-\mathbf{x}\|\,\varepsilon(-\mathbf{x}).$$

[11] On a recours ici à la quatrième condition du théorème 9.3 p. 574.
[12] Voir la définition 8.1 p. 438.
[13] Voir la définition 8.6 p. 448.

Puisque $\|\mathbf{0}_E\| = 0$, que $\| - \mathbf{x}\| = \|\mathbf{x}\|$ et que $\mathrm{d}N_{\mathbf{0}_E}$ est une application linéaire, on obtient en sommant les deux égalités ci-dessus :

$$2\|\mathbf{x}\| = \|\mathbf{x}\| + \| - \mathbf{x}\| = \mathrm{d}N_{\mathbf{0}_E}(\mathbf{x}) - \mathrm{d}N_{\mathbf{0}_E}(\mathbf{x}) + \|\mathbf{x}\|\,\varepsilon_1(\mathbf{x}) = \|\mathbf{x}\|\,\varepsilon_1(\mathbf{x})$$

où $\varepsilon_1 : \mathbf{x} \mapsto \varepsilon(\mathbf{x}) + \varepsilon(-\mathbf{x})$. On en déduit que pour tout $\mathbf{x} \neq \mathbf{0}_E$ on a $\varepsilon_1(\mathbf{x}) = 2$. Cela est incompatible avec la propriété :

$$\lim_{\mathbf{x}\to\mathbf{0}_E} \varepsilon_1(\mathbf{x}) = \lim_{\mathbf{x}\to\mathbf{0}_E} \varepsilon(\mathbf{x}) + \lim_{\mathbf{x}\to\mathbf{0}_E} \varepsilon(-\mathbf{x}) = 0.$$

Remarque En un point $\mathbf{a} \neq \mathbf{0}_E$, on ne pas conclure sur la différentiabilité de l'application norme. En effet, selon le choix de la norme N sur E et selon le choix du point $\mathbf{a} \in E$, l'application N peut être différentiable en $\mathbf{a}$ ou non différentiable en $\mathbf{a}$. Par exemple, si la norme est subordonnée à un produit scalaire, c'est-à-dire s'il existe un produit scalaire $\langle \cdot, \cdot \rangle$ sur E tel que pour tout $\mathbf{x} \in E$ on ait $N(\mathbf{x}) = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$ alors en vertu des propriétés de bilinéarité du produit scalaire et de la dérivabilité de l'application $x \in \mathbb{R}_+^* \mapsto \sqrt{x}$, on montre que N est différentiable en tout $\mathbf{a} \in E \setminus \{\mathbf{0}\}$. Par contre, si on considère la norme sur $\mathbb{R}^2$ définie par $N(x_1, x_2) = |x_1| + |x_2|$ alors on peut montrer (comme on l'a fait ci-dessus) que l'application N n'est pas différentiable en $\mathbf{a} = (1, 0)$. $\diamond$

Applications de classe $\mathcal{C}^1$

Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux espaces vectoriels normés et U un ensemble ouvert dans $(E, \|\cdot\|_E)$. Si une application f de U dans F est différentiable en tout point $\mathbf{a}$ de U alors on dit que l'application f est différentiable sur U. On appelle différentielle de f l'application

$$\mathrm{d}f : \mathbf{a} \in U \longmapsto \mathrm{d}f_{\mathbf{a}} \in \mathcal{L}_c(E, F).$$

Une norme sur l'espace des applications linéaires de E dans F est définie par[14]

$$\forall L \in \mathcal{L}_c(E, F) \quad |||\, L \,|||_{\mathcal{L}_c(E,F)} = \sup_{\mathbf{h}\in E} \frac{\|L(\mathbf{h})\|_F}{\|\mathbf{h}\|_E}.$$

Ainsi, l'application $\mathrm{d}f$ sera continue sur U si pour tout $\mathbf{a} \in U$, on a

$$\lim_{\mathbf{x}\to\mathbf{a}} |||\mathrm{d}f_{\mathbf{x}} - \mathrm{d}f_{\mathbf{a}}|||_{\mathcal{L}_c(E,F)} = 0,$$

où $|||\mathrm{d}f_{\mathbf{x}} - \mathrm{d}f_{\mathbf{a}}|||_{\mathcal{L}_c(E,F)} = \sup_{\mathbf{h}\in E} \dfrac{\|\mathrm{d}f_{\mathbf{x}}(\mathbf{h}) - \mathrm{d}f_{\mathbf{a}}(\mathbf{h})\|_F}{\|\mathbf{h}\|_E}$.

Les considérations sur la continuité de l'application différentielle nous amènent à introduire la définition suivante.

[14] Voir la proposition 9.23 p. 577.

DÉFINITION 10.3 *Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux espaces vectoriels normés et U un ensemble ouvert dans $(E, \|\cdot\|_E)$. Une application f de U dans F est dite de* classe $\mathcal{C}^1$ *sur U si elle est différentiable sur U et si* $\mathrm{d}f$ *est continue sur U.*

Exemple Nous avons vu p. 611 que l'application

$$f : (x_1, x_2) \in \mathbb{R}^2 \longmapsto x_1^2 + 2(x_1 + x_2) + x_2^2$$

est différentiable sur $\mathbb{R}^2$ et que sa différentielle en $\mathbf{a} = (a_1, a_2) \in \mathbb{R}^2$ est

$$\mathrm{d}f_{\mathbf{a}} : (h_1, h_2) \in \mathbb{R}^2 \mapsto 2a_1h_1 + 2h_1 + 2h_2 + 2a_2h_2.$$

Considérons pour $\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2$ et pour $\mathbf{h} = (h_1, h_2) \in \mathbb{R}^2$ la quantité

$$\Delta = \frac{|\mathrm{d}f_{\mathbf{x}}(h_1, h_2) - \mathrm{d}f_{\mathbf{a}}(h_1, h_2)|}{\|\mathbf{h}\|_\infty}.$$

On a

$$\begin{aligned} \Delta &= \frac{|2(x_1 - a_1)h_1 + 2(x_2 - a_2)h_2|}{\|\mathbf{h}\|_\infty} \leqslant 2|x_1 - a_1|\frac{|h_1|}{\|\mathbf{h}\|_\infty} + 2|x_2 - a_2|\frac{|h_2|}{\|\mathbf{h}\|_\infty} \\ &\leqslant 2|x_1 - a_1| + 2|x_2 - a_2| = 2\|\mathbf{x} - \mathbf{a}\|_1. \end{aligned}$$

Lorsque $\mathbf{x}$ tend vers $\mathbf{a}$, la quantité $|x_1 - a_1| + |x_2 - a_2|$ tend vers zéro. On en déduit que pour tout $\mathbf{a} \in \mathbb{R}^2$, la différentielle $\mathrm{d}f$ est continue en $\mathbf{a}$. L'application $\mathrm{d}f$ est donc continue sur $\mathbb{R}^2$.

Une application différentiable sur un ouvert U n'est pas nécessairement une application de classe $\mathcal{C}^1$ sur l'ouvert U. Nous verrons ultérieurement que l'application f définie sur $\mathbb{R}^2$ par

$$f(x_1, x_2) = \begin{cases} x_2^2 \sin\left(\dfrac{x_1}{x_2}\right) & \text{si } x_2 \neq 0 \\ 0 & \text{si } x_2 = 0 \end{cases}$$

est différentiable sur $\mathbb{R}^2$ mais n'est pas de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$.

10.1.2 Opérations sur les applications différentiables

Somme d'applications différentiables

Tout comme la dérivée dans le cas des fonctions de $\mathbb{R}$ dans $\mathbb{R}$, la notion de différentielle se comporte bien vis-à-vis de l'addition et de la multiplication par un scalaire : si f et g sont deux applications de U dans E, différentiables sur U, et si λ et μ sont deux réels alors $\lambda f + \mu g$ est aussi différentiable sur U. De plus, la différentielle est donnée par

$$\mathrm{d}(\lambda f + \mu g) = \lambda\,\mathrm{d}f + \mu\,\mathrm{d}g.$$

Exemple L'application $f : (x_1, x_2) \in \mathbb{R}^2 \mapsto x_1^2 + 2(x_1 + x_2) + x_2^2$ étant la somme de l'application linéaire $L : (x_1, x_2) \in \mathbb{R}^2 \mapsto 2(x_1 + x_2)$ et de la forme quadratique $q : (x_1, x_2) \in \mathbb{R}^2 \mapsto x_1^2 + x_2^2$, on a $\mathrm{d}f = \mathrm{d}L + \mathrm{d}q$. Or, pour tout $(a_1, a_2) \in \mathbb{R}^2$ et tout $(h_1, h_2) \in \mathbb{R}^2$ on a

- d'une part $\mathrm{d}L_{(a_1,a_2)}(h_1, h_2) = L(h_1, h_2)$, voir la proposition 10.5 ;
- et d'autre part $\mathrm{d}q_{(a_1,a_2)}(h_1, h_2) = 2B((a_1, a_2), (h_1, h_2))$ où B est la forme polaire associée à la forme quadratique q, voir l'exercice 10.1 p. 616.

On en déduit que[15]

$$\mathrm{d}f_{(a_1,a_2)} : (h_1, h_2) \in \mathbb{R}^2 \mapsto 2h_1 + 2h_2 + 2a_1h_1 + 2a_2h_2.$$

Composée d'applications différentiables

On a également le résultat suivant concernant la différentiabilité de la composée de deux applications.

PROPOSITION 10.6 *Soient E, F et G trois espaces vectoriels normés, U un ouvert de E et V un ouvert de F. Soit f une application de U dans V, différentiable en $\mathbf{a} \in U$ et g une application de V dans G, différentiable en $\mathbf{b} = f(\mathbf{a})$. L'application composée $g \circ f$ de U dans G est différentiable en $\mathbf{a}$ et on a*

$$d(g \circ f)_{\mathbf{a}} = \mathrm{d}g_{\mathbf{b}} \circ \mathrm{d}f_{\mathbf{a}}.$$

Démonstration L'application g étant différentiable en $f(\mathbf{a})$, pour $\mathbf{y}$ dans un voisinage de $\mathbf{b} = f(\mathbf{a})$, on a

$$g(\mathbf{y}) = g(\mathbf{b}) + \mathrm{d}g_{\mathbf{b}}(\mathbf{y} - \mathbf{b}) + \|\mathbf{y} - \mathbf{b}\|_F \, \varepsilon_1(\mathbf{y}) \quad \text{avec} \quad \lim_{\mathbf{y} \to \mathbf{b}} \varepsilon_1(\mathbf{y}) = \mathbf{0}_G.$$

En prenant pour $\mathbf{y}$ des valeurs de la forme $\mathbf{y} = f(\mathbf{x})$ avec $\mathbf{x}$ dans un voisinage de $\mathbf{a}$, on obtient

$$g(f(\mathbf{x})) = g(f(\mathbf{a})) + \mathrm{d}g_{\mathbf{b}}(f(\mathbf{x}) - f(\mathbf{a})) + \|f(\mathbf{x}) - f(\mathbf{a})\|_F \, \varepsilon_1(f(\mathbf{x})) \qquad (2)$$

avec $\lim_{\mathbf{x} \to \mathbf{a}} \varepsilon_1(f(\mathbf{x})) = \mathbf{0}_G$. L'application f étant différentiable en $\mathbf{a}$, on a aussi

$$f(\mathbf{x}) - f(\mathbf{a}) = \mathrm{d}f_{\mathbf{a}}(\mathbf{x} - \mathbf{a}) + \|\mathbf{x} - \mathbf{a}\|_E \, \varepsilon_2(\mathbf{x}) \quad \text{avec} \quad \lim_{\mathbf{x} \to \mathbf{a}} \varepsilon_2(\mathbf{x}) = \mathbf{0}_F. \qquad (3)$$

En combinant les relations (2) et (3), on obtient

$$g(f(\mathbf{x})) = g(f(\mathbf{a})) + \mathrm{d}g_{f(\mathbf{a})}(\mathrm{d}f_{\mathbf{a}}(\mathbf{x} - \mathbf{a})) + \|\mathbf{x} - \mathbf{a}\|_E \, \varepsilon(\mathbf{x})$$

où on a posé

$$\varepsilon(\mathbf{x}) = \mathrm{d}g_{f(\mathbf{a})}(\varepsilon_2(\mathbf{x})) + \frac{\|\mathrm{d}f_{\mathbf{a}}(\mathbf{x} - \mathbf{a}) + \|\mathbf{x} - \mathbf{a}\|_E \, \varepsilon_2(\mathbf{x})\|_F}{\|\mathbf{x} - \mathbf{a}\|_E} \varepsilon_1(f(\mathbf{x})).$$

[15] Signalons que le calcul de la différentielle de f en ayant recours à la définition a été effectué p. 611.

Comme $\mathrm{d}f$ et $\mathrm{d}g$ sont des applications linéaires et continues, la fonction ε admet zéro pour limite lorsque $\mathbf{x}$ tend vers $\mathbf{a}$. On en conclut que $g \circ f$ est différentiable en $\mathbf{a}$, de différentielle en $\mathbf{a}$ l'application $d(g\circ f)_{\mathbf{a}} : \mathbf{h} \in E \longmapsto \mathrm{d}g_{f(\mathbf{a})}(\mathrm{d}f_{\mathbf{a}}(\mathbf{h}))$. □

Exemple Considérons l'application $\Psi : \mathrm{M} \in \mathrm{GL}_n(\mathbb{R}) \mapsto \mathrm{Tr}(\mathrm{M}^{-1})$ qui est la composée de l'application trace avec l'application « inversion de matrice ». Schématiquement on a :

$$\begin{aligned} \mathrm{GL}_n(\mathbb{R}) &\longrightarrow \mathrm{GL}_n(\mathbb{R}) \longrightarrow \mathbb{R} \\ \mathrm{M} &\longmapsto \mathrm{M}^{-1} \longmapsto \mathrm{Tr}(\mathrm{M}^{-1}). \end{aligned}$$

L'application trace étant linéaire, d'après la proposition 10.5, sa différentielle en tout point est égale à elle-même, *i.e.*

$$\forall \mathrm{M} \in \mathrm{GL}_n(\mathbb{R}) \quad \mathrm{dTr}_{\mathrm{M}} : \mathrm{H} \in \mathcal{M}_n(\mathbb{R}) \longmapsto \mathrm{Tr}(H).$$

La différentielle de l'application $\Phi : \mathrm{M} \in \mathrm{GL}_n(\mathbb{R}) \mapsto \mathrm{M}^{-1}$ a été étudiée p. 613 ; on a

$$\forall \mathrm{M} \in \mathrm{GL}_n(\mathbb{R}) \quad \mathrm{d}\Phi_{\mathrm{M}} : \mathrm{H} \in \mathcal{M}_n(\mathbb{R}) \longmapsto -\mathrm{M}^{-1}\mathrm{H}\mathrm{M}^{-1}.$$

La proposition 10.6 indique que $\mathrm{d}\Psi_{\mathrm{M}} = \mathrm{dTr}_{\Phi(\mathrm{M})} \circ \mathrm{d}\Phi_{\mathrm{M}}$ pour tout $\mathrm{M} \in \mathrm{GL}_n(\mathbb{R})$. Autrement dit, on a

$$\forall \mathrm{M} \in \mathrm{GL}_n(\mathbb{R}) \quad \mathrm{d}\Psi_{\mathrm{M}} : \mathrm{H} \in \mathcal{M}_n(\mathbb{R}) \longmapsto -\mathrm{Tr}(\mathrm{M}^{-1}\mathrm{H}\mathrm{M}^{-1}).$$

Exercice 10.2 Soient U l'ensemble des matrices de $\mathcal{M}_n(\mathbb{R})$ de trace non nulle et $\xi : \mathrm{M} \mapsto 1/\mathrm{Tr}(\mathrm{M})$.
1 - Justifier que U est un sous-ensemble ouvert de $\mathcal{M}_n(\mathbb{R})$.
2 - En utilisant la proposition 10.6, montrer que la différentielle de ξ est $\mathrm{d}\xi : M \in U \mapsto \mathrm{d}\xi_M \in \mathcal{L}_c(\mathcal{M}_n(\mathbb{R}), \mathbb{R})$ où $\mathrm{d}\xi_M : H \in \mathcal{M}_n(\mathbb{R}) \mapsto -\mathrm{Tr}(H)/\mathrm{Tr}(\mathrm{M})^2$.

Au sujet du produit et du quotient d'applications différentiables

Parmi les exemples de calcul d'applications différentielles remarquons que l'on a su calculer[16] la différentielle d'une forme quadratique, c'est-à-dire[17] la différentielle d'un produit de formes linéaires et que l'on a aussi déterminé p. 620 la différentielle de l'application $\mathrm{M} \in \mathrm{GL}_n(\mathbb{R}) \mapsto \mathrm{M}^{-1}$, c'est-à-dire la différentielle de l'inverse[18] de l'application $id : \mathrm{M} \in \mathrm{GL}_n(\mathbb{R}) \mapsto \mathrm{M}$.

Plus généralement, on peut se poser la question suivante : si deux applications f et g définies sur un ouvert U de E à valeurs dans F sont différentiables, est-ce que le produit $f \times g$ et le quotient f/g sont différentiables ?

---

[16] Voir l'exercice 10.1 p. 616.

[17] Rappelons, voir le théorème de Gauss p. 469, que toute forme quadratique se décompose en une somme de carrés de formes linéaires.

[18] À ne pas confondre avec la réciproque.

Une première remarque est de constater que ces deux opérations, produit et quotient entre applications, n'ont pas toujours de sens. Par exemple si $F = \mathbb{R}^2$ alors on ne sait pas faire le produit de deux éléments de $\mathbb{R}^2$. Pour pouvoir définir le produit ou le quotient, il faut supposer que l'ensemble d'arrivée F possède non seulement une structure d'espace vectoriel normé mais puisse aussi être muni d'une loi de composition interne « produit ».

Avec la structure d'*algèbre normé*, voir la définition 9.3 p. 544, on dispose de la structure nécessaire pour définir la différentielle d'un produit.

PROPOSITION 10.7 *Soient $(E, \|\cdot\|_E)$ un espace vectoriel normé, $(F, \|\cdot\|_F)$ une algèbre normée, U un ensemble ouvert dans $(E, \|\cdot\|_E)$ et f et g deux applications de U dans F. Si les applications f et g sont différentiables en $\mathbf{x} \in U$ alors le produit $f \times g$ est différentiable en $\mathbf{x}$ et la différentielle de $f \times g$ en $\mathbf{x}$ est l'application $\mathrm{d}(f \times g)_{\mathbf{x}}$ définie par*

$$\forall \mathbf{h} \in E \quad \mathrm{d}(f \times g)_{\mathbf{x}}(\mathbf{h}) = \mathrm{d}f_{\mathbf{x}}(\mathbf{h}) \times g(\mathbf{x}) \; + \; f(\mathbf{x}) \times \mathrm{d}g_{\mathbf{x}}(\mathbf{h}).$$

Démonstration Pour montrer que $f \times g$ est différentiable en $\mathbf{x} \in U$, de différentielle l'application $\mathrm{d}(f \times g)_{\mathbf{x}} : \mathbf{h} \in E \mapsto \mathrm{d}f_{\mathbf{x}}(\mathbf{h}) \times g(\mathbf{x}) + f(\mathbf{x}) \times \mathrm{d}g_{\mathbf{x}}(\mathbf{h})$, nous allons montrer que pour tout $\mathbf{h} \in E$, on a

$$(f \times g)(\mathbf{x} + \mathbf{h}) = (f \times g)(\mathbf{x}) + \mathrm{d}(f \times g)_{\mathbf{x}}(\mathbf{h}) + o_{\mathbf{0}}(\|\mathbf{h}\|_E).$$

On peut remarquer que comme les applications $\mathrm{d}f_{\mathbf{x}}$ et $\mathrm{d}g_{\mathbf{x}}$ sont des applications linéaires continues, l'application $\mathrm{d}(f \times g)_{\mathbf{x}}$ est une application linéaire continue.

Puisque les applications f et g sont différentiables en $\mathbf{x}$, pour tout $\mathbf{h}$ au voisinage de $\mathbf{0}$ dans E on a

$$\begin{aligned} f(\mathbf{x} + \mathbf{h}) &= f(\mathbf{x}) + \mathrm{d}f_{\mathbf{x}}(\mathbf{h}) + o_{\mathbf{0}}(\|\mathbf{h}\|_E), \\ g(\mathbf{x} + \mathbf{h}) &= g(\mathbf{x}) + \mathrm{d}g_{\mathbf{x}}(\mathbf{h}) + o_{\mathbf{0}}(\|\mathbf{h}\|_E) = g(\mathbf{x}) + o_{\mathbf{0}}(1). \end{aligned} \tag{4}$$

Or[19] :

$$(f \times g)(\mathbf{x}+\mathbf{h}) - (f \times g)(\mathbf{x}) = \big(f(\mathbf{x}+\mathbf{h}) - f(\mathbf{x})\big) \times g(\mathbf{x}+\mathbf{h}) + f(\mathbf{x}) \times \big(g(\mathbf{x}+\mathbf{h}) - g(\mathbf{x})\big).$$

En utilisant les relations (4), on en déduit que

$$\begin{aligned} (f \times g)(\mathbf{x} + \mathbf{h}) - (f \times g)(\mathbf{x}) &= \big(\mathrm{d}f_{\mathbf{x}}(h) + o_{\mathbf{0}}(\|\mathbf{h}\|_E)\big) \times \big(g(\mathbf{x}) + o_{\mathbf{0}}(1)\big) \\ &\qquad + f(\mathbf{x}) \times \big(\mathrm{d}g_{\mathbf{x}}(\mathbf{h}) + o_{\mathbf{0}}(\|\mathbf{h}\|_E)\big) \\ &= \mathrm{d}f_{\mathbf{x}}(\mathbf{h}) \times g(\mathbf{x}) + f(\mathbf{x}) \times \mathrm{d}g_{\mathbf{x}}(\mathbf{h}) + o_{\mathbf{0}}(\|\mathbf{h}\|_E) \\ &= \mathrm{d}(f \times g)_{\mathbf{x}}(\mathbf{h}) + o_{\mathbf{0}}(\|\mathbf{h}\|_E). \end{aligned}$$

Le résultat est démontré. □

[19] On sera attentif à la forme de cette expression : la loi « produit » de l'algèbre F n'est pas nécessairement commutative. C'est le cas par exemple de l'ensemble des matrices carrées de taille n ($n \geqslant 2$) pour lequel le produit n'est pas commutatif.

Exemples

1. Considérons deux applications f et g définies sur un ouvert I de $\mathbb{R}$ à valeurs dans $\mathbb{R}$ et considérons l'application produit $f \times g$ par

$$f \times g : x \in I \longmapsto f(x)\, g(x) \in \mathbb{R}.$$

La proposition 10.7 indique que si f et g sont différentiables sur I, autrement dit si f et g sont dérivables[20] sur I alors le produit $f \times g$ est aussi différentiable (donc dérivable) sur I. De plus, on a

$$\begin{aligned}\forall x \in I \quad \forall h \in \mathbb{R} \quad \mathrm{d}(f \times g)_x(h) &= f'(x)h\, g(x) + f(x)\, g'(x)h \\ &= (f'(x)\, g(x) + f(x)\, g'(x))\, h = (f \times g)'(x)\, h.\end{aligned}$$

On retrouve la formule de la dérivée d'un produit de deux applications réelles d'une variable réelle[21].

2. L'application $\mathrm{id} : \mathrm{M} \in \mathcal{M}_n(\mathbb{R}) \mapsto M$ étant une application linéaire, elle est différentiable sur $\mathcal{M}_n(\mathbb{R})$ et

$$\forall \mathrm{M} \in \mathcal{M}_n(\mathbb{R}) \quad \mathrm{d}(\mathrm{id})_{\mathrm{M}} : \mathrm{H} \in \mathcal{M}_n(\mathbb{R}) \longmapsto \mathrm{H}.$$

L'espace d'arrivée $\mathcal{M}_n(\mathbb{R})$ pouvant être muni d'une structure d'algèbre (voir p. 544), on peut s'intéresser à l'application $f = \mathrm{id} \times \mathrm{id} : \mathrm{M} \in \mathcal{M}_n(\mathbb{R}) \mapsto M^2$. D'après la proposition 10.7, cette application est différentiable sur $\mathcal{M}_n(\mathbb{R})$ et on a

$$\forall \mathrm{M} \in \mathcal{M}_n(\mathbb{R}) \quad \mathrm{d}f_{\mathrm{M}} : \mathrm{H} \in \mathcal{M}_n(\mathbb{R}) \longmapsto \mathrm{HM} + \mathrm{MH}.$$

Rappelons[22] qu'un élément $\mathbf{a}$ d'un anneau $(A, +, \times)$ est inversible s'il existe $\mathbf{b} \in A$ tel que $\mathbf{a} \times \mathbf{b} = \mathbf{1}_A$ et $\mathbf{b} \times \mathbf{a} = \mathbf{1}_A$ où l'élément $\mathbf{1}_A$ est l'unité de l'anneau A (c'est-à-dire l'élément neutre pour la loi $\times$). L'inverse d'un élément inversible $\mathbf{a} \in A$ est unique[23] et est noté $\mathbf{a}^{-1}$.

DÉFINITION 10.4 *L'ensemble de tous les éléments* inversibles *d'une algèbre* $(A, +, \cdot, \times)$ *est l'ensemble, noté* A^*, *défini par*

$$A^* = \{\mathbf{a} \in A \ ; \ \exists \mathbf{b} \in A \quad \mathbf{a} \times \mathbf{b} = \mathbf{1}_A \quad \textit{et} \quad \mathbf{b} \times \mathbf{a} = \mathbf{1}_A\}$$

où l'élément $\mathbf{1}_A$ *est l'unité de l'anneau* $(A, +, \times)$.

---

[20] Voir p. 614 le lien entre dérivabilité et différentiabilité dans le cas de fonctions de $\mathbb{R}$ dans $\mathbb{R}$.

[21] Voir la Proposition 16.3 p. 718 du *Cours de première année*.

[22] Voir le *Cours de première année* p. 68.

[23] Supposons que $\mathbf{b}$ et $\mathbf{c}$ soient deux inverses de l'élément $\mathbf{a}$. On a donc $\mathbf{a} \times \mathbf{b} = \mathbf{1}_A$ et $\mathbf{a} \times \mathbf{c} = \mathbf{1}_A$. Par différence, on en déduit que $\mathbf{a} \times (\mathbf{b} - \mathbf{c}) = \mathbf{0}_A$. En multipliant cette relation à gauche par $\mathbf{b}$, on a $\mathbf{b} - \mathbf{c} = \mathbf{0}_A$ et donc $\mathbf{b} = \mathbf{c}$.

Exemples

1. L'ensemble des éléments inversibles de l'algèbre des nombres réels $(\mathbb{R}, +, \cdot, \times)$ est l'ensemble $\mathbb{R}^* = \{x \in \mathbb{R} \ ; \ x \neq 0\}$.

2. Considérons l'algèbre $(\mathcal{M}_n(\mathbb{R}), +, \cdot, \times)$. L'élément unité est la matrice identité I_n et l'ensemble des éléments inversibles de l'algèbre $\mathcal{M}_n(\mathbb{R})$ est l'ensemble $\mathrm{GL}_n(\mathbb{R})$ des matrices inversibles.

Si une application f définie sur un ouvert U d'un espace vectoriel normé E est à valeurs dans l'ensemble des éléments inversibles F^* d'une algèbre normée F alors on définit l'application inverse de f, notée $\frac{1}{f}$ par

$$\frac{1}{f} : \mathbf{x} \in U \longmapsto f(\mathbf{x})^{-1} \in F^*.$$

On prendra garde à ne pas confondre l'inverse $\frac{1}{f}$ d'une application f avec d'une part $f(\mathbf{x})^{-1}$ l'inverse de l'élément $f(\mathbf{x})$ pour la multiplication dans le groupe $(F^*, \times)$ et avec d'autre part f^{-1} la bijection réciproque de f (dans la mesure où elle existe).

Proposition 10.8 *Soient $(E, \|\cdot\|_E)$ un espace vectoriel normé, $(F, \|\cdot\|_F)$ une algèbre normée, U un ensemble ouvert dans $(E, \|\cdot\|_E)$ et f une application de U dans l'ensemble F^* des éléments inversibles de l'algèbre normée F. Si l'application f est différentiable en $\mathbf{x} \in U$ alors son inverse $\frac{1}{f}$ est différentiable en $\mathbf{x}$ et*

$$\forall \mathbf{h} \in E \quad \mathrm{d}\left(\frac{1}{f}\right)_{\mathbf{x}}(\mathbf{h}) = -f(\mathbf{x})^{-1} \ \mathrm{d}f_{\mathbf{x}}(\mathbf{h}) \ f(\mathbf{x})^{-1}.$$

Démonstration La démonstration de ce résultat qui reprend, en la généralisant, l'approche utilisée p. 620 pour déterminer la différentielle de l'application $\mathrm{M} \in \mathrm{GL}_n(\mathbb{R}) \mapsto \mathrm{M}^{-1}$ est laissée en exercice. □

Corollaire 10.1 *Soient $(E, \|\cdot\|_E)$ un espace vectoriel normé, U un ensemble ouvert dans $(E, \|\cdot\|_E)$ et f une application de U dans $\mathbb{R}^*$. Si l'application f est différentiable en $\mathbf{x} \in U$ alors son inverse $\frac{1}{f}$ est différentiable en $\mathbf{x}$ et*

$$\forall \mathbf{h} \in E \quad \mathrm{d}\left(\frac{1}{f}\right)_{\mathbf{x}}(\mathbf{h}) = -\frac{\mathrm{d}f_{\mathbf{x}}(\mathbf{h})}{f(\mathbf{x})^2}.$$

Si $E = F = \mathbb{R}$ alors les notions de dérivabilité et de différentiabilité coïncident[(24)] : si f est une application d'un ouvert U de $\mathbb{R}$ à valeurs dans $\mathbb{R}$ qui est différentiable en $x \in U$ alors elle est dérivable en x et on a

$$f'(x) = \mathrm{d}f_x(1).$$

[(24)] Voir la remarque de la p. 614.

La proposition 10.1 s'énonce alors en disant que l'inverse de toute application dérivable f de $U \subset \mathbb{R}$ dans $\mathbb{R}^*$ est dérivable et que

$$\forall x \in U \quad \left(\frac{1}{f}\right)'(x) = -\frac{f'(x)}{f(x)^2}.$$

Signalons qu'une approche analogue peut être mise en œuvre pour étudier les applications de classe $\mathcal{C}^1$: la somme, la composée, le produit, etc. d'applications de classe $\mathcal{C}^1$ sont aussi de classe $\mathcal{C}^1$.

10.1.3 Utilisation des différentielles en physique

Pour conclure cette partie, donnons deux exemples d'utilisation des différentielles pour l'étude de problèmes physiques.

Conservation de l'énergie le long d'une trajectoire

Le principe fondamental de la dynamique exprime le fait que dans un repère galiléen, la somme des forces $\mathbf{F}$ qui s'exercent sur un point matériel de masse m est proportionnelle à son accélération. Si on note $\mathbf{x}(t) \in \mathbb{R}^3$ la position et $\mathbf{v}(t) = \mathbf{x}'(t) \in \mathbb{R}^3$ la vitesse de ce point au temps $t \in \mathbb{R}^+$ alors son accélération est donnée par[25] $\mathbf{v}'(t)$ et le principe fondamental de la dynamique s'écrit :

$$m\,\mathbf{v}'(t) = \mathbf{F}(\mathbf{x}(t)).$$

L'inconnue dans ce problème est le vecteur position $\mathbf{x}(t)$ à chaque instant t. L'application $\mathbf{x} : t \in \mathbb{R}^+ \mapsto \mathbf{x}(t) \in \mathbb{R}^3$ est un arc paramétré dans l'espace[26]. La trajectoire du point, qui est la courbe de l'espace décrite par cet arc paramétré, est l'ensemble

$$\left\{(x_1, x_2, x_3) \in \mathbb{R}^3 \ ; \ \exists t \in \mathbb{R}^+ \quad (x_1, x_2, x_3) = \mathbf{x}(t)\right\}.$$

Par exemple, si un point de coordonnées $(x_1, x_2, x_3) \in \mathbb{R}^3$ est soumis à la seule gravité alors la force s'exerçant sur lui est

$$\mathbf{F}(x_1, x_2, x_3) = \begin{pmatrix} 0 \\ 0 \\ -mg \end{pmatrix}.$$

Les équations $\mathbf{v}(t) = \mathbf{x}'(t)$ et $m\mathbf{v}'(t) = \mathbf{F}(\mathbf{x}(t))$ forment un système différentiel[27] d'ordre 1 qui se réécrit, en posant $\mathbf{X} = (\mathbf{x}, \mathbf{v})$, sous la forme $\mathbf{X}'(t) = f(\mathbf{X}(t))$ où f est l'application $f : (\mathbf{x}, \mathbf{v}) \in \mathbb{R}^3 \times \mathbb{R}^3 \mapsto (\mathbf{v}, \mathbf{F}(\mathbf{x}))$.

---

[25] Au sujet de la notation $\mathbf{v}'(t)$ pour la différentielle d'une application d'un ouvert de $\mathbb{R}$ à valeurs dans $\mathbb{R}^3$, on pourra consulter la proposition 10.4 p. 614.

[26] L'étude des arcs paramétrés est effectuée dans le chap. 11.

[27] Voir le *Cours de première année* p. 979.

Dans de nombreuses situations physiques, la force $\mathbf{F}$ « dérive » d'un potentiel, c'est-à-dire qu'il existe une application V de $\mathbb{R}^3$ dans $\mathbb{R}$, différentiable, telle que

$$\forall (\mathbf{x},\mathbf{y}) \in \mathbb{R}^3 \times \mathbb{R}^3 \quad \mathbf{F}(\mathbf{x}) \cdot \mathbf{y} = -m \, \mathrm{d}V_{\mathbf{x}}(\mathbf{y}).$$

L'énergie totale du « système » est alors introduite de la façon suivante :

$$E(\mathbf{x},\mathbf{v}) = \frac{m}{2}\|\mathbf{v}\|^2 + V(\mathbf{x}).$$

Par exemple, la force de gravité dérive du potentiel $V : (x_1,x_2,x_3) \in \mathbb{R}^3 \mapsto g x_3$ et l'énergie associée est $E((x_1,x_2,x_3),\mathbf{v}) = \frac{1}{2}m\|\mathbf{v}\|^2 + g\, x_3$.

L'introduction des applications différentiables, et en particulier de la proposition 10.6, permet d'établir le principe suivant.

L'énergie totale d'un système soumis à une force dérivant d'un potentiel est conservée le long des trajectoires.

Démonstration Considérons un système soumis à une force $\mathbf{F}$ qui dérive d'un potentiel V. Prouver que l'énergie totale est conservée le long des trajectoires revient à montrer que l'application composée $E \circ \mathbf{X} : t \in \mathbb{R} \mapsto E(\mathbf{X}(t)) \in \mathbb{R}$ est une application constante. Puisque l'application $E \circ \mathbf{X}$ est une fonction réelle de la variable réelle qui est continue, il suffit de prouver que sa dérivée[28] est nulle.

La formule de différentiation des applications composées (voir la proposition 10.6) indique que pour tout $t \in \mathbb{R}$ on a

$$\begin{aligned}(E \circ \mathbf{X})'(t) &= \mathrm{d}E_{\mathbf{X}(t)}\big(\mathbf{X}'(t)\big) = \mathrm{d}E_{\mathbf{X}(t)}\big(f(\mathbf{X}(t))\big) \\ &= \mathrm{d}E_{(\mathbf{x}(t),\mathbf{v}(t))}\big(\mathbf{v}(t),\mathbf{F}(\mathbf{x}(t))\big).\end{aligned} \tag{5}$$

L'application $E : (\mathbf{x},\mathbf{v}) \in \mathbb{R}^3 \times \mathbb{R}^3 \mapsto \frac{m}{2}\|\mathbf{v}\|^2 + V(\mathbf{x}) \in \mathbb{R}$ étant la somme d'une forme quadratique et de l'application $(\mathbf{x},\mathbf{v}) \mapsto V(\mathbf{x})$, sa différentielle en $(\mathbf{x},\mathbf{v}) \in \mathbb{R}^3 \times \mathbb{R}^3$ vaut

$$\mathrm{d}E_{(\mathbf{x},\mathbf{v})} : (\mathbf{y},\mathbf{w}) \in \mathbb{R}^3 \times \mathbb{R}^3 \longmapsto m\mathbf{v} \cdot \mathbf{w} + m \, \mathrm{d}V_{\mathbf{x}}(\mathbf{y}).$$

Étant donné que la force $\mathbf{F}$ dérive du potentiel V, on a $\mathrm{d}V_{\mathbf{x}}(\mathbf{y}) = -m\mathbf{F}(\mathbf{x}) \cdot \mathbf{y}$ et on déduit de la relation (5) que :

$$(E \circ \mathbf{X})'(t) = m\mathbf{v}(t) \cdot \mathbf{F}(\mathbf{x}(t)) - m\mathbf{F}(\mathbf{x}(t)) \cdot \mathbf{v}(t) = 0.$$

L'application $E \circ \mathbf{X}$ est donc à dérivée nulle, ce qui permet de conclure à la conservation de l'énergie le long des trajectoires. □

[28] On peut montrer que si l'application $\mathbf{F}$ est de classe $\mathcal{C}^1$ alors la solution de l'équation $\mathbf{X}'(t) = f(\mathbf{X}(t))$ sera aussi de classe $\mathcal{C}^1$, ainsi l'application $E \circ \mathbf{X}$ est dérivable.

Application à la mécanique des milieux continus

En mécanique des milieux continus, on appelle loi de comportement, une relation qui exprime la contrainte exercée sur une particule en fonction du mouvement passé et des changements d'état que la particule a subis. Les lois de comportement les plus simples se traduisent par une expression mathématique de la contrainte représentée par une matrice symétrique $S \in \mathcal{S}_3(\mathbb{R})$ en fonction du taux de déformation $D \in \mathcal{S}_3(\mathbb{R})$. Elle est suggérée par des essais expérimentaux sur le milieu qu'on étudie.
Pour un fluide polymère comme du blanc d'œuf ou du dentifrice, une loi de comportement parfois utilisée pour relier la contrainte S à la déformation D est

$$S = \mu \operatorname{Tr}(D^2)\, D \quad \text{où } \mu \in \mathbb{R}. \tag{6}$$

Pour calculer la différentielle de l'application $f : D \in \mathcal{S}_3(\mathbb{R}) \mapsto S \in \mathcal{S}_3(\mathbb{R})$ ainsi définie, on forme le taux d'accroissement, pour tout $H \in \mathcal{S}_3(\mathbb{R})$:

$$\begin{aligned} f(D+H) &= \mu \operatorname{Tr}((D+H)^2)\,(D+H) \\ &= f(D) + \underbrace{\mu \operatorname{Tr}(D^2)\, H + 2\mu \operatorname{Tr}(DH)\, D}_{\text{linéaire en H}} + \underbrace{\mu \operatorname{Tr}(H^2)\,(D+H) + 2\mu \operatorname{Tr}(DH)\, H}_{=\mathcal{O}_0(\|H\|)}. \end{aligned}$$

On obtient :

$$df_D : H \in \mathcal{S}_3(\mathbb{R}) \longmapsto 2\mu \operatorname{Tr}(DH)\, D + \mu \operatorname{Tr}(D^2)\, H.$$

Si on soumet l'écoulement de fluide à un cisaillement (entre deux plaques parallèles distantes d'une longueur L et dont la vitesse relative est V), le taux de déformation D_0 est donné par

$$D_0 = \begin{pmatrix} 0 & a & 0 \\ a & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{où } a = \frac{V}{L}.$$

La relation (6) indique que la contrainte associée à un tel écoulement est

$$f(D_0) = 2\mu a^2 D_0.$$

Supposons qu'on souhaite déterminer la contrainte S lorsque le fluide est soumis à un écoulement proche d'un cisaillement, c'est-à-dire à un écoulement dont le taux de déformation est de la forme $D = D_0 + H$ avec $H = (h_{i,j})_{(i,j)\in\{1,2,3\}^2}$ modélisant une petite déformation. D'après la définition 10.1, on a la relation

$$f(D_0 + H) = 2\mu a^2 D_0 + 2\mu \operatorname{Tr}(D_0 H) D_0 + \mu \operatorname{Tr}(D_0^2) H + \mathcal{O}_0(\|H\|).$$

Comme $\operatorname{Tr}(D_0 H) = 2a h_{2,1}$ et $\operatorname{Tr}(D_0^2) = 2a^2$, la perturbation (en première approximation) qu'entraîne la déformation H sur la contrainte est égale à

$$4\mu a h_{2,1} D_0 + 2\mu a^2 H.$$

10.2 Différentiabilité en dimension finie

Dans le chapitre 9, nous avons vu que pour des applications entre espaces vectoriels de dimension finie un certain nombre de notions prenaient une forme simplifiée par rapport au cadre général. C'est également le cas en ce qui concerne les résultats liés à la différentiabilité. Moyennant le choix d'une base, un espace vectoriel réel de dimension finie $n \in \mathbb{N}^*$ peut être identifié à l'espace vectoriel $\mathbb{R}^n$.

10.2.1 Dérivation selon un vecteur

Définition 10.5 (Dérivée directionnelle)
Soient f une application définie sur un ouvert U de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^p$ et $\mathbf{v} \in \mathbb{R}^n$ un vecteur non nul. On dit que l'application f est dérivable en $\mathbf{a} \in U$ selon le vecteur $\mathbf{v}$ *(ou dans la direction du vecteur $\mathbf{v}$) si*

$$\lim_{t \to 0} \frac{f(\mathbf{a}+t\mathbf{v}) - f(\mathbf{a})}{t} \quad \textit{existe.}$$

Cette limite est notée $\mathrm{d}_{\mathbf{v}} f_{\mathbf{a}}$.

Remarques

1. L'existence de la dérivée directionnelle ne dépend pas du vecteur mais uniquement de sa direction. En effet, si f est dérivable en $\mathbf{a}$ selon un vecteur $\mathbf{v} \in \mathbb{R}^n \setminus \{0\}$ alors f est aussi dérivable en $\mathbf{a}$ selon tous les vecteurs de la forme $\lambda\mathbf{v}$, $\lambda \in \mathbb{R}^*$ puisque

$$\lim_{t \to 0} \frac{f(\mathbf{a}+t\lambda\mathbf{v}) - f(\mathbf{a})}{t} = \lambda \lim_{s \to 0} \frac{f(\mathbf{a}+s\mathbf{v}) - f(\mathbf{a})}{s}.$$

Plus exactement cette égalité indique que si l'application f est dérivable en $\mathbf{a} \in U$ selon le vecteur $\mathbf{v}$ alors

$$\forall \lambda \in \mathbb{R}^* \quad \mathrm{d}_{\lambda\mathbf{v}} f_{\mathbf{a}} = \lambda\, \mathrm{d}_{\mathbf{v}} f_{\mathbf{a}}.$$

2. Lorsque $n = 1$ on a seulement deux directions correspondant aux valeurs $v = 1$ et $v = -1$. Une application f définie sur un intervalle ouvert I de $\mathbb{R}$, à valeurs dans $\mathbb{R}^p$, est dérivable en $a \in I$ selon $v = 1$ si

$$\lim_{t \to 0} \frac{f(a+t) - f(a)}{t} \quad \text{existe,}$$

et elle est dérivable en $a \in I$ selon $v = -1$ si

$$\lim_{t \to 0} \frac{f(a-t) - f(a)}{t} = -\lim_{s \to 0} \frac{f(a+s) - f(a)}{s} \quad \text{existe.}$$

Autrement dit, dans le cas où $n = 1$, il y a équivalence entre « dérivable » et « dérivable selon un vecteur ».

3. D'après le théorème 7.2, en dimension finie toutes les normes sont équivalentes et en conséquence l'existence d'une limite est indépendante de la norme choisie. Ainsi, en dimension finie, l'existence d'une dérivée directionnelle est elle aussi indépendante du choix de la norme. ◇

Exemple Considérons l'application $f : (x_1, x_2) \in \mathbb{R}^2 \mapsto x_1^2 - 2x_1x_2 + |\sin(x_2)|$ et le vecteur $\mathbf{v}_1 = (2, 1)$. Pour déterminer si cette application est dérivable selon le vecteur $\mathbf{v}_1$ en $(0, 0)$, on forme pour $t \in \mathbb{R}^*$ le taux d'accroissement suivant :

$$\frac{f(0+2t, 0+t) - f(0,0)}{t} = \frac{|\sin(t)|}{t}.$$

Puisque cette quantité n'a pas de limite lorsque t tend vers 0, l'application f n'est pas dérivable en $(0, 0)$ selon le vecteur $\mathbf{v}_1$. Par contre, f est dérivable selon le vecteur $\mathbf{v}_2 = (1, 0)$ en $(0, 0)$ puisque le taux d'accroissement en $(0, 0)$ qui vaut

$$\frac{f(0+t, 0) - f(0,0)}{t} = t$$

admet pour limite 0 lorsque t tend vers 0. L'application f est dérivable en $(0, 0)$ dans la direction $\mathbf{v}_2 = (1, 0)$ et on a $\mathrm{d}_{\mathbf{v}_2} f_{(0,0)} = 0$.

Exercice 10.3
1 - Montrer que l'application $f : (x_1, x_2) \in \mathbb{R}^2 \mapsto \sin(x_1 + x_2) \in \mathbb{R}$ est dérivable selon les deux vecteurs $\mathbf{v}_1 = (-1, 1)$ et $\mathbf{v}_2 = (1, 1)$ en tout $\mathbf{a} \in \mathbb{R}^2$ et calculer $\mathrm{d}_{\mathbf{v}_1} f_{\mathbf{a}}$ et $\mathrm{d}_{\mathbf{v}_2} f_{\mathbf{a}}$.
2 - Soient g une application de $\mathbb{R}$ dans $\mathbb{R}$ dérivable sur $\mathbb{R}$ et f l'application $f : (x_1, x_2) \in \mathbb{R}^2 \mapsto g(x_1 + x_2)$. Déterminer les dérivées de f dans toutes les directions $\mathbf{v} \in \mathbb{R}^2_*$ et en tout $\mathbf{a} \in \mathbb{R}^2$.

THÉORÈME 10.1 *Soit f une application définie sur un ouvert U de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^p$. Si l'application f est différentiable en $\mathbf{a} \in U$ alors elle admet en $\mathbf{a}$ une dérivée selon tout vecteur $\mathbf{v} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ et on a $\mathrm{d}_{\mathbf{v}} f_{\mathbf{a}} = \mathrm{d}f_{\mathbf{a}}(\mathbf{v})$.*

Démonstration Puisque f est différentiable en $\mathbf{a}$, pour $\mathbf{x}$ dans un voisinage de $\mathbf{a}$, on a

$$f(\mathbf{x}) = f(\mathbf{a}) + \mathrm{d}f_{\mathbf{a}}(\mathbf{x} - \mathbf{a}) + \|\mathbf{x} - \mathbf{a}\| \, \varepsilon(\mathbf{x}) \quad \text{avec} \quad \lim_{\mathbf{x} \to \mathbf{a}} \varepsilon(\mathbf{x}) = \mathbf{0},$$

où $\mathrm{d}f_{\mathbf{a}}$ désigne la différentielle de f en $\mathbf{a}$. Considérons des éléments $\mathbf{x}$ de U de la forme $\mathbf{x} = \mathbf{a} + t\mathbf{v}$ où t est un réel appartenant à un voisinage de 0. En notant $\widetilde{\varepsilon}$ la fonction $t \in \mathbb{R} \mapsto \varepsilon(\mathbf{a} + t\mathbf{v})$ la relation précédente devient

$$f(\mathbf{a} + t\mathbf{v}) = f(\mathbf{a}) + t \, \mathrm{d}f_{\mathbf{a}}(\mathbf{v}) + |t| \, \|\mathbf{v}\| \, \widetilde{\varepsilon}(t) \quad \text{avec} \quad \lim_{t \to 0} \widetilde{\varepsilon}(t) = 0.$$

Lorsque $t \neq 0$, elle se réécrit sous la forme

$$\frac{f(\mathbf{a}+t\mathbf{v})-f(\mathbf{a})}{t} = \mathrm{d}f_{\mathbf{a}}(\mathbf{v}) + \frac{|t|}{t}\,\|\mathbf{v}\|\,\widetilde{\varepsilon}(t).$$

Comme $|t|/t$ est borné par 1 et que $\widetilde{\varepsilon}(t)$ tend vers 0 lorsque t tend vers 0, on en déduit que le taux d'accroissement $(f(\mathbf{a}+t\mathbf{v})-f(\mathbf{a}))/t$ admet $\mathrm{d}f_{\mathbf{a}}(\mathbf{v})$ comme limite lorsque t tend vers 0. Ceci montre que l'application f est dérivable selon le vecteur $\mathbf{v}$ au point $\mathbf{a} \in U$. □

Remarque On a vu dans les exemples donnés p. 611 et p. 618 que l'application $f : (x_1, x_2) \in \mathbb{R}^2 \mapsto x_1^2 + 2(x_1 + x_2) + x_2^2$ est différentiable et que sa différentielle en $(a_1, a_2) \in \mathbb{R}^2$ est

$$\mathrm{d}f_{(a_1,a_2)} : (h_1, h_2) \in \mathbb{R}^2 \mapsto 2a_1h_1 + 2h_1 + 2h_2 + 2a_2h_2.$$

Le théorème 10.1 permet de retrouver l'expression de la différentielle en calculant les dérivées directionnelles de f en $(a_1, a_2) \in \mathbb{R}^2$ dans la direction $(h_1, h_2) \in \mathbb{R}^2$; on a :

$$\begin{aligned}\mathrm{d}_{(h_1,h_2)}f_{(a_1,a_2)} &= \lim_{t\to 0} \frac{f(a_1+th_1, a_2+th_2)-f(a_1,a_2)}{t}\\ &= 2a_1h_1 + 2h_1 + 2h_2 + 2a_2h_2.\end{aligned}$$

Attention toutefois au fait que le théorème 10.1 ne permet pas d'établir que l'application f est différentiable en $(a_1, a_2) \in \mathbb{R}^2$. Cela est une hypothèse du théorème. ◇

Le théorème 10.1 n'admet pas de réciproque : une application peut admettre des dérivées selon tous les vecteurs sans pour autant être différentiable (voir l'exemple suivant).

Exemple Considérons l'application f de $\mathbb{R}^2$ dans $\mathbb{R}$ définie par

$$f(x_1, x_2) = \begin{cases} \dfrac{x_1x_2}{x_1^2 + x_2} & \text{lorsque } x_2 \neq -x_1^2,\\ 0 & \text{sinon.}\end{cases}$$

Cette application n'étant pas continue en $\mathbf{0}$ (on vérifie par exemple que l'on a $\lim\limits_{x\to 0} f(x, x^3 - x^2) = -1 \neq 0$), d'après la proposition 10.2, elle n'est pas différentiable en $\mathbf{0}$. Par contre, elle admet des dérivées directionnelles en $\mathbf{0}$ selon tous les vecteurs puisque si $\mathbf{v} = (\alpha, \beta)$ désigne un vecteur non nul de $\mathbb{R}^2$ alors on a

$$\frac{f(\mathbf{0}+t\mathbf{v})-f(\mathbf{0})}{t} = \frac{f(t\alpha, t\beta)}{t} = \frac{\alpha\beta}{t\alpha^2 + \beta}.$$

Cette quantité admet toujours une limite quand t tend vers 0 qui vaut 0 si $\beta = 0$ et α sinon.

10.2.2 Dérivées partielles

Soit f une application d'un ouvert U de $\mathbb{R}^n$ dans $\mathbb{R}^p$ différentiable en $\mathbf{a} \in U$; la différentielle de f en $\mathbf{a}$ est une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^p$ ce qui implique que pour tous vecteurs $\mathbf{u}$, $\mathbf{v}$ de $\mathbb{R}^n$ non nuls et pour tous réels λ et μ tels que $\lambda\mathbf{u} + \mu\mathbf{v} \neq \mathbf{0}$ on a

$$\mathrm{d}_{\lambda\mathbf{u}+\mu\mathbf{v}} f_{\mathbf{a}} = \mathrm{d}f_{\mathbf{a}}(\lambda\mathbf{u} + \mu\mathbf{v}) = \lambda\,\mathrm{d}f_{\mathbf{a}}(\mathbf{u}) + \mu\,\mathrm{d}f_{\mathbf{a}}(\mathbf{v}) = \lambda\,\mathrm{d}_{\mathbf{u}} f_{\mathbf{a}} + \mu\,\mathrm{d}_{\mathbf{v}} f_{\mathbf{a}}.$$

On en déduit que si on connaît les dérivées directionnelles selon k vecteurs non nuls alors on dispose de la dérivée directionnelle selon n'importe quel vecteur non nul engendré par ces k vecteurs. En particulier, si on connaît les dérivées directionnelles selon les n vecteurs d'une base $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ de $\mathbb{R}^n$ alors on connaît la dérivée selon tous les vecteurs $\mathbf{v}$ non nuls de $\mathbb{R}^n$: si on désigne par $v_1, \ldots, v_n$ les composantes de $\mathbf{v}$ dans la base $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ alors on a, pour tout $\mathbf{a} \in U$

$$\mathrm{d}_{\mathbf{v}} f_{\mathbf{a}} = \sum_{i=1}^{n} v_i\,\mathrm{d}_{\mathbf{e}_i} f_{\mathbf{a}}. \tag{7}$$

Cette constatation nous amène à introduire la définition suivante.

Définition 10.6 (Dérivées partielles)
Soient $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ une base de $\mathbb{R}^n$, U un ouvert de $\mathbb{R}^n$ et f une application de U dans $\mathbb{R}^p$ différentiable en $\mathbf{a} \in U$. On appelle i–ième dérivée partielle *de f en $\mathbf{a}$ et on note $\frac{\partial f}{\partial x_i}(\mathbf{a})$ ou $\partial_i f(\mathbf{a})$, la dérivée de f en $\mathbf{a}$ selon le vecteur de base $\mathbf{e}_i$, i.e.*

$$\frac{\partial f}{\partial x_i}(\mathbf{a}) = \mathrm{d}_{\mathbf{e}_i} f_{\mathbf{a}}.$$

Remarques

1. Le plus souvent $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ est la base canonique de $\mathbb{R}^n$, mais on aura également recours ultérieurement à la base des coordonnées sphérique et cylindrique de $\mathbb{R}^3$.

2. Si l'application f considérée dans la définition 10.6 est différentiable sur U, on peut définir l'application $\partial_i f : \mathbf{a} \in U \mapsto \partial_i f(\mathbf{a}) \in \mathbb{R}^p$ qui est appelée *i–ième dérivée partielle de f*. Cette application $\partial_i f$ est aussi notée $\frac{\partial f}{\partial x_i}$. ◇

Exemple Soit f l'application $(x_1, x_2) \in \mathbb{R}^2 \mapsto (x_1\cos(x_2), x_1\sin(x_2)) \in \mathbb{R}^2$. Pour tout $(x_1, x_2) \in \mathbb{R}^2$ et pour tout $t \in \mathbb{R}^*$ on a

$$\frac{f(x_1 + t, x_2) - f(x_1, x_2)}{t} = (\cos(x_2), \sin(x_2)).$$

Ainsi, l'application f admet une dérivée partielle selon le premier vecteur de la base canonique de $\mathbb{R}^2$ donnée par :

$$\forall (x_1, x_2) \in \mathbb{R}^2 \quad \frac{\partial f}{\partial x_1}(x_1, x_2) = (\cos(x_2), \sin(x_2)).$$

Pour tout $(x_1, x_2) \in \mathbb{R}^2$ et pour tout $t \in \mathbb{R}^*$ on a

$$\frac{f(x_1, x_2 + t) - f(x_1, x_2)}{t} = \left(x_1 \frac{\cos(x_2 + t) - \cos(x_2)}{t}, x_1 \frac{\sin(x_2 + t) - \sin(x_2)}{t} \right).$$

Or, on a

$$\lim_{t \to 0} \frac{\cos(x_2 + t) - \cos(x_2)}{t} = -\sin(x_2)$$

$$\text{et} \quad \lim_{t \to 0} \frac{\sin(x_2 + t) - \sin(x_2)}{t} = \cos(x_2)$$

ce qui permet d'affirmer que f admet une dérivée partielle selon le second vecteur de la base canonique de $\mathbb{R}^2$ donnée par :

$$\forall (x_1, x_2) \in \mathbb{R}^2 \quad \frac{\partial f}{\partial x_2}(x_1, x_2) = (-x_1 \sin(x_2), x_1 \cos(x_2)).$$

Exercice 10.4 Montrer que l'application f définie sur $\mathbb{R}^3$ par

$$f : (x_1, x_2, x_3) \in \mathbb{R}^3 \mapsto (x_1 \cos(x_2) \sin(x_3), x_1 \sin(x_2) \sin(x_3), x_1 \cos(x_3)) \in \mathbb{R}^3$$

admet des dérivées partielles selon les trois vecteurs de la base canonique de $\mathbb{R}^3$ et calculer les dérivées partielles $\partial_1 f$, $\partial_2 f$ et $\partial_3 f$.

Lien entre dérivées partielles et fonctions partielles

La i–ième dérivée partielle d'une application f définie sur un ouvert U de $\mathbb{R}^n$ et à valeurs dans $\mathbb{R}^p$, qui est différentiable sur U, n'est rien d'autre que la dérivée de la i–ième application partielle[29] de f. D'après la définition 10.6, on a en effet pour tout $\mathbf{a} = (a_1, \ldots, a_n) \in U$,

$$\begin{aligned} \frac{\partial f}{\partial x_i}(\mathbf{a}) = \mathrm{d}_{\mathbf{e}_i} f_{\mathbf{a}} &= \lim_{h \to 0} \frac{f(\mathbf{a} + h\,\mathbf{e}_i) - f(\mathbf{a})}{h} \\ &= \lim_{h \to 0} \frac{\varphi_{\mathbf{a},i}(a_i + h) - \varphi_{\mathbf{a},i}(a_i)}{h} = \varphi'(a_i) \end{aligned}$$

où $\varphi_{\mathbf{a},i} : x_i \in \mathbb{R} \mapsto f(a_1, \ldots, a_{i-1}, x_i, a_{i+1}, \ldots, a_n)$ désigne la i–ième application partielle de f en $\mathbf{a}$.

On en déduit que pour calculer la i–ième dérivée partielle de l'application f en $\mathbf{a} = (a_1, \ldots, a_n) \in U$, on peut procéder de la manière suivante : on fige toutes les variables $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$ et on calcule la dérivée de la quantité $f(x_1, \ldots, x_n)$ par rapport à la seule variable x_i ; on évalue ensuite l'expression obtenue en $(a_1, \ldots, a_n)$.

[29] Voir la définition 9.6 p. 554.

Exemple Calculons les dérivées partielles en $(0,0,0)$ de l'application

$$f : (x_1, x_2, x_3) \in \mathbb{R}^3 \mapsto \big(x_1 + \mathrm{sh}(x_1x_2x_3),\ x_2\,\mathrm{ch}(x_1 + x_2 + x_3)\big) \in \mathbb{R}^2.$$

En dérivant l'expression de $f(x_1, x_2, x_3)$ par rapport à x_1 en considérant x_2 et x_3 fixés, on obtient

$$\frac{\partial f}{\partial x_1}(x_1, x_2, x_3) = \big(1 + x_2x_3\,\mathrm{ch}(x_1x_2x_3),\ x_2\,\mathrm{sh}(x_1 + x_2 + x_3)\big)$$

et en $(0,0,0)$, on a $\partial_1 f(0,0,0) = (1,0)$. En dérivant l'expression de $f(x_1, x_2, x_3)$ par rapport à x_2 en considérant x_1 et x_3 fixés, on obtient

$$\frac{\partial f}{\partial x_2}(x_1, x_2, x_3) = \big(x_1x_3\,\mathrm{ch}(x_1x_2x_3),\ \mathrm{ch}(x_1 + x_2 + x_3) + x_2\,\mathrm{sh}(x_1 + x_2 + x_3)\big)$$

et en $(0,0,0)$, on a $\partial_2 f(0,0,0) = (0,1)$. Enfin, en dérivant l'expression de $f(x_1, x_2, x_3)$ par rapport à x_3 en considérant x_1 et x_2 fixés, on obtient

$$\frac{\partial f}{\partial x_3}(x_1, x_2, x_3) = \big(x_1x_2\,\mathrm{ch}(x_1x_2x_3),\ x_2\,\mathrm{sh}(x_1 + x_2 + x_3)\big)$$

et en $(0,0,0)$, on a $\partial_3 f(0,0,0) = (0,0)$.

Exercice 10.5 Calculer les deux dérivées partielles des applications :
1 - $f : (x_1, x_2) \in \mathbb{R}^2_* \longmapsto x_2 \dfrac{x_1^2 - x_2^2}{x_1^2 + x_2^2}$. 2 - $f : (x_1, x_2) \in \mathbb{R}^2_* \longmapsto \dfrac{4x_1^2x_2^3}{(x_1^2 + x_2^2)^2}$.

Expression de la différentielle à l'aide des dérivées partielles

Le théorème 10.1 indique qu'une application différentiable admet des dérivées partielles selon tous les vecteurs de base et fournit l'expression des dérivées partielles à l'aide de la différentielle. La proposition suivante indique qu'à l'inverse, la différentielle en $\mathbf{a}$ d'une application différentiable en $\mathbf{a}$ peut être exprimée à l'aide des dérivées partielles en $\mathbf{a}$.

PROPOSITION 10.9 *Soit f une application définie sur un ouvert U de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^p$, différentiable en $\mathbf{a} \in U$. La différentielle de f en $\mathbf{a}$ est donnée par*

$$\mathrm{d}f_{\mathbf{a}} : \mathbf{h} \in \mathbb{R}^n \longmapsto \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(\mathbf{a}),$$

où les réels h_i correspondent aux composantes de $\mathbf{h}$ dans la base où sont calculées les dérivées partielles.

Nous avons introduit la notion de dérivées partielles dans le cas d'applications de $\mathbb{R}^n$ dans $\mathbb{R}^p$. Moyennant le choix d'une base de l'espace de départ, la définition que nous avons donnée d'une dérivée partielle s'étend à une application d'un espace vectoriel de dimension finie E sur un espace vectoriel de dimension finie F. Nous allons illustrer cette remarque à l'aide d'un exemple.

Exemple Considérons l'application f qui à toute matrice de $\mathcal{M}_2(\mathbb{R})$ associe son polynôme caractéristique :

$$\begin{aligned} f : \mathcal{M}_2(\mathbb{R}) &\longrightarrow \mathbb{R}_2[X] \\ \mathrm{M} &\longmapsto \det(\mathrm{M} - X\,\mathrm{I}_2). \end{aligned}$$

Pour $\mathrm{M} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_2(\mathbb{R})$ on a $f(\mathrm{M}) = X^2 - (a+d)X + (ad - bc)$. L'espace vectoriel $\mathcal{M}_2(\mathbb{R})$ est un espace vectoriel de dimension 4. Choisissons pour base de cet espace la famille composée des matrices suivantes :

$$\mathrm{E}_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \mathrm{E}_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \mathrm{E}_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \quad \text{et} \quad \mathrm{E}_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Pour calculer les dérivées partielles de l'application f en I_2 selon cette base, il suffit de déterminer les limites des taux d'accroissement

$$\frac{f(\mathrm{I}_2 + t\,\mathrm{E}_{ij}) - f(\mathrm{I}_2)}{t}$$

lorsque t tend vers zéro. Ainsi, la dérivée partielle selon le premier vecteur de base est

$$\begin{aligned} \partial_1 f(\mathrm{I}_2) &= \lim_{t \to 0} \frac{f(\mathrm{I}_2 + t\,\mathrm{E}_{11}) - f(\mathrm{I}_2)}{t} \\ &= \lim_{t \to 0} \frac{(X^2 - (2+t)X + (1+t)) - (X^2 - 2X + 1)}{t} = 1 - X. \end{aligned}$$

De la même manière, on obtient pour les autres dérivées partielles (correspondant aux autres vecteurs de la base E_{12}, E_{21} puis E_{22}) : $\partial_2 f(\mathrm{I}_2) = \partial_3 f(\mathrm{I}_2) = 0$ et $\partial_4 f(\mathrm{I}_2) = 1 - X$.

À l'aide de ces dérivées partielles, on peut[30] en utilisant la proposition 10.9 obtenir l'expression de la différentielle de l'application f en I_2 : pour toute matrice $\mathrm{H} \in \mathcal{M}_2(\mathbb{R})$ de composantes $h_{ij} \in \mathbb{R}$ on a

$$\begin{aligned} \mathrm{d}f_{\mathrm{I}_2}(\mathrm{H}) &= \partial_1 f(\mathrm{I}_2)\ h_{11} + \partial_2 f(\mathrm{I}_2)\ h_{12} + \partial_3 f(\mathrm{I}_2)\ h_{21} + \partial_4 f(\mathrm{I}_2)\ h_{22} \\ &= (1 - X)(h_{11} + h_{22}) = (1 - X)\mathrm{Tr}(\mathrm{H}). \end{aligned}$$

[30] Pour appliquer la proposition 10.9, il faut en toute rigueur commencer par établir que l'application f est différentiable. Nous verrons avec le théorème 10.2 que la propriété de différentiabilité d'une application peut être obtenue à l'aide des dérivées partielles.

Exercice 10.6 L'espace vectoriel $\mathbb{R}_2[X]$ des polynômes de degré inférieur ou égal à 2 est un espace vectoriel de dimension 3 que l'on peut munir de la base $\mathcal{B} = \{1, 1+X, 1+X^2\}$. Calculer selon cette base les trois dérivées partielles en $Q = 1 + X^2$ de l'application

$$\begin{aligned} f \ : \quad \mathbb{R}_2[X] \quad &\longrightarrow \quad \mathbb{R}_2[X] \\ a + bX + cX^2 &\longmapsto \sin(ab)X - \cos(c)X^2. \end{aligned}$$

L'intérêt des dérivées partielles est, comme on vient de le voir, de ramener le calcul de différentielles au calcul de dérivées de fonctions d'une seule variable réelle. Les dérivées partielles permettent aussi de caractériser les applications de classe $\mathcal{C}^1$ comme l'indique le théorème suivant.

THÉORÈME 10.2 *Une application d'un ouvert U de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^p$ est de classe $\mathcal{C}^1$ sur U si et seulement si elle admet des dérivées partielles par rapport à tous les vecteurs d'une base de $\mathbb{R}^n$ qui sont continues sur U.*

Démonstration $\trianglerighteq$ Supposons que l'application f admette des dérivées partielles selon tous les vecteurs de la base canonique de $\mathbb{R}^n$ et que ces dérivées partielles soient continues sur U. D'après la proposition 10.9, pour tout $\mathbf{a} \in U$, l'application

$$\mathbf{h} \in \mathbb{R}^n \longmapsto \sum_{i=1}^{n} h_i \, \partial_i f(\mathbf{a})$$

est la différentielle $\mathrm{d}f_{\mathbf{a}}$ de l'application f en $\mathbf{a}$. Afin de prouver que f est de classe $\mathcal{C}^1$, nous allons montrer que l'application $\mathrm{d}f : \mathbf{a} \in U \mapsto \mathrm{d}f_{\mathbf{a}}$ ainsi définie est une application continue sur U, autrement dit, nous allons montrer[31] que pour tout $\mathbf{a} \in U$ on a

$$\lim_{\mathbf{x} \to \mathbf{a}} |||\mathrm{d}f_{\mathbf{a}} - \mathrm{d}f_{\mathbf{x}}|||_{\mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)} = 0,$$

où

$$|||\mathrm{d}f_{\mathbf{a}} - \mathrm{d}f_{\mathbf{x}}|||_{\mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)} = \sup_{\substack{\mathbf{h} \in \mathbb{R}^n \\ \mathbf{h} \neq \mathbf{0}}} \frac{\|(\mathrm{d}f_{\mathbf{a}} - \mathrm{d}f_{\mathbf{x}})(\mathbf{h})\|_{\infty, \mathbb{R}^p}}{\|h\|_{\infty, \mathbb{R}^n}}.$$

Pour tous $(\mathbf{a}, \mathbf{x}) \in U^2$ et $\mathbf{h} \in \mathbb{R}^n$ avec $\mathbf{h} \neq \mathbf{0}$, on a

$$\begin{aligned} \|(\mathrm{d}f_{\mathbf{a}} - \mathrm{d}f_{\mathbf{x}})(\mathbf{h})\|_{\infty, \mathbb{R}^p} &= \left\| \sum_{i=1}^{n} h_i \left(\partial_i f(\mathbf{a}) - \partial_i f(\mathbf{x}) \right) \right\|_{\infty, \mathbb{R}^p} \\ &\leqslant \sum_{i=1}^{n} |h_i| \, \left\| \partial_i f(\mathbf{a}) - \partial_i f(\mathbf{x}) \right\|_{\infty, \mathbb{R}^p}. \end{aligned}$$

[31] Voir la définition 9.5 p. 546.

Pour tout $i \in \{1, \dots, n\}$ on a $|h_i| \leqslant \|h\|_{\infty,\mathbb{R}^n}$ de sorte que

$$\frac{\|(\mathrm{d}f_{\mathbf{a}} - \mathrm{d}f_{\mathbf{x}})(\mathbf{h})\|_{\infty,\mathbb{R}^p}}{\|h\|_{\infty,\mathbb{R}^n}} \leqslant \sum_{i=1}^{n} \left\|\partial_i f(\mathbf{a}) - \partial_i f(\mathbf{x})\right\|_{\infty,\mathbb{R}^p}.$$

Cette inégalité étant vraie pour tout vecteur $\mathbf{h} \in \mathbb{R}^n_*$, on en déduit que

$$0 \leqslant \|\mathrm{d}f_{\mathbf{a}} - \mathrm{d}f_{\mathbf{x}}\|_{\mathcal{L}(\mathbb{R}^n,\mathbb{R}^p)} \leqslant \sum_{i=1}^{n} \left\|\partial_i f(\mathbf{a}) - \partial_i f(\mathbf{x})\right\|_{\infty,\mathbb{R}^p}. \tag{8}$$

Par hypothèse, les dérivées partielles étant continues sur U, pour tout $\mathbf{a} \in U$ on a

$$\lim_{\mathbf{x}\to\mathbf{a}} \partial_i f(\mathbf{x}) = \partial_i f(\mathbf{a})$$

ce qui implique que tous les termes de la somme dans le membre de droite de l'inégalité (8) tendent vers 0 lorsque $\mathbf{x}$ tend vers $\mathbf{a}$. Le théorème d'encadrement permet d'affirmer que $\mathrm{d}f_{\mathbf{x}}$ tend vers $\mathrm{d}f_{\mathbf{a}}$ lorsque $\mathbf{x}$ tend vers $\mathbf{a}$ ce qui prouve que l'application $\mathrm{d}f$ est continue en $\mathbf{a}$ pour tout $\mathbf{a} \in U$.

$\trianglerighteq$ Réciproquement, supposons que l'application f soit de classe $\mathcal{C}^1$ sur U. L'application f étant différentiable sur U, les dérivées partielles selon tous les vecteurs de la base canonique $(\mathbf{e}_1, \dots, \mathbf{e}_n)$ de $\mathbb{R}^n$ existent en tout $\mathbf{a} \in U$. Il s'agit de montrer que ces dérivées partielles sont continues sur U, c'est-à-dire que pour tout $\mathbf{a} \in U$ et pour tout $i \in \{1, \dots, n\}$ on a

$$\lim_{\mathbf{x}\to\mathbf{a}} \partial_i f(\mathbf{x}) = \partial_i f(\mathbf{a}).$$

Soient $i \in \{1, \dots, n\}$ et $(\mathbf{a}, \mathbf{x}) \in U^2$; par définition de la dérivée partielle selon le vecteur $\mathbf{e}_i$, on a

$$\partial_i f(\mathbf{a}) - \partial_i f(\mathbf{x}) = \mathrm{d}f_{\mathbf{a}}(\mathbf{e}_i) - \mathrm{d}f_{\mathbf{x}}(\mathbf{e}_i) = (\mathrm{d}f_{\mathbf{a}} - \mathrm{d}f_{\mathbf{x}})(\mathbf{e}_i).$$

On en déduit que

$$0 \leqslant \left\|\partial_i f(\mathbf{a}) - \partial_i f(\mathbf{x})\right\|_{\infty,\mathbb{R}^p} \leqslant \left\|(\mathrm{d}f_{\mathbf{a}} - \mathrm{d}f_{\mathbf{x}})(\mathbf{e}_i)\right\|_{\infty,\mathbb{R}^p} \leqslant \|\mathrm{d}f_{\mathbf{a}} - \mathrm{d}f_{\mathbf{x}}\|_{\mathcal{L}(\mathbb{R}^n,\mathbb{R}^p)}.$$

Comme l'application $\mathrm{d}f$ est continue, la limite de $\|\mathrm{d}f_{\mathbf{a}} - \mathrm{d}f_{\mathbf{x}}\|_{\mathcal{L}(\mathbb{R}^n,\mathbb{R}^p)}$ lorsque $\mathbf{x}$ tend vers $\mathbf{a}$ est nulle. Le théorème d'encadrement permet d'affirmer que $\partial_i f(\mathbf{x})$ tend vers $\partial_i f(\mathbf{a})$ lorsque $\mathbf{x}$ tend vers $\mathbf{a}$ ce qui prouve que la i–ième dérivée partielle de f est une application continue en $\mathbf{a}$. □

10.2.3 Matrice jacobienne

Il est commode, lorsque l'on est amené à manipuler les dérivées partielles d'une fonction f de $\mathbb{R}^n$ dans $\mathbb{R}^p$, d'introduire une matrice, appelée *matrice jacobienne* de f, dont les éléments sont les dérivées partielles de f en un point.

Définition 10.7 *Soit f une application d'un ouvert U de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^p$, différentiable en $\mathbf{a} \in U$. La* matrice jacobienne *de f en $\mathbf{a}$ est la matrice de $\mathcal{M}_{n,p}(\mathbb{R})$ définie par*

$$\mathrm{J}_f(\mathbf{a}) = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1}(\mathbf{a}) & \dfrac{\partial f_1}{\partial x_2}(\mathbf{a}) & \dfrac{\partial f_1}{\partial x_n}(\mathbf{a}) \\ \vdots & \vdots & \vdots \\ \dfrac{\partial f_p}{\partial x_1}(\mathbf{a}) & \dfrac{\partial f_p}{\partial x_2}(\mathbf{a}) & \dfrac{\partial f_p}{\partial x_n}(\mathbf{a}) \end{pmatrix}.$$

Lorsque $n = p$, on appelle jacobien *de f en $\mathbf{a}$ le déterminant de la matrice jacobienne de f en $\mathbf{a}$: $\mathcal{J}_f(\mathbf{a}) = \det(\mathrm{J}_f(\mathbf{a}))$.*

Remarque On appelle encore matrice jacobienne d'une application f d'un ouvert U de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^p$ différentiable, la « matrice » dont les éléments sont des applications de U dans $\mathbb{R}$:

$$\mathrm{J}_f = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1} & \dfrac{\partial f_1}{\partial x_2} & \dfrac{\partial f_1}{\partial x_n} \\ \vdots & \vdots & \vdots \\ \dfrac{\partial f_p}{\partial x_1} & \dfrac{\partial f_p}{\partial x_2} & \dfrac{\partial f_p}{\partial x_n} \end{pmatrix}.$$

On sera vigilant au fait que ce tableau n'est pas une matrice au sens défini en première année[32] car les éléments de ce tableau sont des applications de U dans $\mathbb{R}$ et l'ensemble des applications de U dans $\mathbb{R}$ n'est pas un corps.

De même, lorsque $n = p$ par abus de langage on appelle jacobien d'une application f, le « déterminant » de la « matrice » jacobienne de f. Le jacobien de f est ainsi une application de U dans $\mathbb{R}$. ◇

Exemple L'application[33] $f : (x_1, x_2) \in \mathbb{R}^2 \mapsto (x_1 \cos(x_2), x_1 \sin(x_2)) \in \mathbb{R}^2$ est une application qui admet des dérivées partielles selon tous les vecteurs de la base canonique de $\mathbb{R}^2$ (voir p. 630) et pour $(x_1, x_2) \in \mathbb{R}^2$ on a

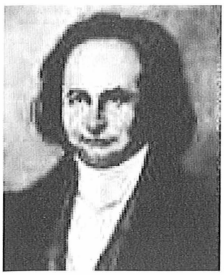
$$\frac{\partial f}{\partial x_1}(x_1, x_2) = (\cos(x_2), \sin(x_2)) \quad \text{et} \quad \frac{\partial f}{\partial x_2}(x_1, x_2) = (-x_1 \sin(x_2), x_1 \cos(x_2)).$$

La matrice jacobienne de f en $(x_1, x_2) \in \mathbb{R}^2$ vaut

$$\mathrm{J}_f(x_1, x_2) = \begin{pmatrix} \cos(x_2) & -x_1 \sin(x_2) \\ \sin(x_2) & x_1 \cos(x_2) \end{pmatrix}.$$

et le jacobien de f en $(x_1, x_2) \in \mathbb{R}^2$ qui est le déterminant de cette matrice vaut $\mathcal{J}_f(x_1, x_2) = x_1$.

JACOBI, Carl-Gustav-Jakob (1804, Postdam - 1851, Berlin).



On doit à Jacobi des découvertes majeures en théorie des nombres (il a calculé notamment le nombre de façon de décomposer un entier en somme de quatre carrés). En même temps qu'Abel, mais indépendamment, il a établi des résultats fondamentaux concernant les fonctions elliptiques. En mécanique, il a prolongé des travaux d'Euler concernant le problème des trois corps[34]. Il est surtout resté célèbre pour son étude systématique des déterminants, dans un mémoire intitulé *De determinantibus functionalibus* paru en 1841.

Exercice 10.7 Déterminer la matrice jacobienne ainsi que le jacobien de l'application[35]

$$f : (x_1, x_2, x_3) \in \mathbb{R}^3 \mapsto (x_1 \cos(x_2)\sin(x_3), x_1 \sin(x_2)\sin(x_3), x_1 \cos(x_3)) \in \mathbb{R}^3.$$

Considérons une fonction f de $\mathbb{R}^n$ dans $\mathbb{R}^p$, différentiable en $\mathbf{a} \in \mathbb{R}^n$ et une fonction g de $\mathbb{R}^p$ dans $\mathbb{R}^q$, différentiable en $\mathbf{b} = f(\mathbf{a})$. Les applications coordonnées $(g \circ f)_i$, $i \in \{1, \ldots, q\}$, de la fonction $g \circ f$ sont des fonctions à valeurs dans $\mathbb{R}$ dont les dérivées partielles se calculent en utilisant les règles de calcul de la dérivée[36] de fonctions de $\mathbb{R}$ dans $\mathbb{R}$. Pour tout couple $(i, j) \in \{1, \ldots, q\} \times \{1, \ldots, n\}$, on a

$$\partial_j (g \circ f)_i(\mathbf{a}) = \sum_{k=1}^{p} \partial_k g_i(f(\mathbf{a})) \times \partial_j f_k(\mathbf{a}) = \sum_{k=1}^{p} \partial_k g_i(\mathbf{b}) \times \partial_j f_k(\mathbf{a}), \qquad (9)$$

où les fonctions f_i, $i \in \{1, \ldots, p\}$ et g_j, $j \in \{1, \ldots, q\}$ correspondent aux applications coordonnées de f et g respectivement. La relation (9) peut s'interpréter matriciellement en utilisant les matrices jacobiennes des applications f, g et $f \circ g$.

---

[32] Dans le *Cours de première année*, on a défini des matrices à coefficients dans un corps.

[33] Cette application intervient dans le changement en coordonnées polaires, voir p. 666.

[34] Le problème des 3 corps peut s'énoncer sous la forme suivante : on considère trois corps massiques placés dans l'espace $\mathbb{R}^3$ (on peut penser à des planètes). Ces trois corps sont en interaction gravitationnelle les uns avec les autres : chacun de ces corps attire les deux autres selon une loi énoncée par Newton - la loi de la gravitation universelle. Comment déterminer les trajectoires suivies par ces trois corps ?

[35] Cette application intervient dans le changement de coordonnées sphériques, voir p. 672.

[36] Si f et g sont deux applications de $\mathbb{R}$ dans $\mathbb{R}$ dérivables alors on a $(g \circ f)' = (g' \circ f) \times f'$.

Proposition 10.10 *Soient f une application d'un ouvert U de $\mathbb{R}^n$ à valeurs dans un ouvert V de $\mathbb{R}^p$, et g une application de V dans $\mathbb{R}^q$. Si l'application f est différentiable en $\mathbf{a} \in U$ et si l'application g est différentiable en $\mathbf{b} = f(\mathbf{a})$ alors l'application $g \circ f$ est différentiable en $\mathbf{a}$ et on a*

$$\mathrm{J}_{g\circ f}(\mathbf{a}) = \mathrm{J}_g(\mathbf{b}) \times \mathrm{J}_f(\mathbf{a}). \tag{10}$$

Démonstration Notons $\varphi = g \circ f$. Comme f est différentiable en $\mathbf{a} \in U$ et que g est différentiable en $\mathbf{b} = f(\mathbf{a})$, d'après la proposition 10.6 l'application φ est différentiable en $\mathbf{a}$.

Le produit C de la matrice $\mathrm{J}_g(\mathbf{b}) \in \mathcal{M}_{p,q}(\mathbb{R})$ par la matrice $\mathrm{J}_f(\mathbf{a}) \in \mathcal{M}_{n,p}(\mathbb{R})$ est une matrice appartenant à $\mathcal{M}_{n,q}(\mathbb{R})$. Les coefficients de C sont

$$c_{ij} = \sum_{k=1}^{p} (\mathrm{J}_g(\mathbf{b}))_{ik} \ (\mathrm{J}_f(\mathbf{a}))_{kj} = \sum_{k=1}^{p} \partial_k g_i(\mathbf{b}) \ \partial_j f_k(\mathbf{a}).$$

D'après la relation (9), ce terme est égal à $\partial_j(g \circ f)_i(\mathbf{a})$. On a donc pour tous $i \in \{1, \ldots, p\}$ et $j \in \{1, \ldots, q\}$ la relation $(\mathrm{J}_{g\circ f}(\mathbf{a}))_{ij} = \partial_i(g \circ f)_i(\mathbf{a})$, ce qui permet de conclure que $\mathrm{J}_{g\circ f}(\mathbf{a}) = \mathrm{J}_g(\mathbf{b}) \times \mathrm{J}_f(\mathbf{a})$. □

Exemple Soient g une application de $\mathbb{R}^2$ dans $\mathbb{R}$, différentiable sur $\mathbb{R}^2$ et $f : (x_1, x_2) \in \mathbb{R}^2 \mapsto (x_1 \cos(x_2), x_1 \sin(x_2)) \in \mathbb{R}^2$. Intéressons-nous au calcul des dérivées partielles de l'application composée $\varphi = g \circ f$. Puisque les applications g et f sont toutes les deux différentiables[37], d'après la proposition 10.6, la composée $\varphi = g \circ f$ est différentiable sur $\mathbb{R}^2$. Nous allons donner deux méthodes permettant de calculer les dérivées partielles de l'application φ, la première en utilisant la formule (9) et la seconde utilisant la proposition 10.10. Dans les deux cas, on notera f_1 et f_2 les deux fonctions coordonnées de l'application f de sorte que

$$\varphi : (x_1, x_2) \in \mathbb{R}^2 \longmapsto g(f_1(x_1, x_2), f_2(x_1, x_2)) \in \mathbb{R}.$$

En utilisant la formule (9), on a

$$\begin{aligned}\frac{\partial \varphi}{\partial x_1}(x_1, x_2) = &\frac{\partial f_1}{\partial x_1}(x_1, x_2) \frac{\partial g}{\partial x_1}(f_1(x_1, x_2), f_2(x_1, x_2)) \\ &+ \frac{\partial f_2}{\partial x_1}(x_1, x_2) \frac{\partial g}{\partial x_2}(f_1(x_1, x_2), f_2(x_1, x_2)),\end{aligned}$$

et

$$\begin{aligned}\frac{\partial \varphi}{\partial x_2}(x_1, x_2) = &\frac{\partial f_1}{\partial x_2}(x_1, x_2) \frac{\partial g}{\partial x_1}(f_1(x_1, x_2), f_2(x_1, x_2)) \\ &+ \frac{\partial f_2}{\partial x_2}(x_1, x_2) \frac{\partial g}{\partial x_2}(f_1(x_1, x_2), f_2(x_1, x_2)),\end{aligned}$$

[37] L'application f a déjà été étudiée p. 630 et p. 636.

Ayant calculé p. 630 les dérivées partielles des applications f_1 et f_2, on obtient

$$\begin{aligned}\frac{\partial\varphi}{\partial x_1}(x_1,x_2) = \cos(x_2)\,\frac{\partial g}{\partial x_1}(f_1(x_1,x_2),f_2(x_1,x_2)) \\ + \sin(x_2)\,\frac{\partial g}{\partial x_2}(f_1(x_1,x_2),f_2(x_1,x_2)),\end{aligned} \tag{11}$$

et

$$\begin{aligned}\frac{\partial\varphi}{\partial x_2}(x_1,x_2) = -x_1\sin(x_2)\,\frac{\partial g}{\partial x_1}(f_1(x_1,x_2),f_2(x_1,x_2)) \\ + x_1\cos(x_2)\,\frac{\partial g}{\partial x_2}(f_1(x_1,x_2),f_2(x_1,x_2)).\end{aligned} \tag{12}$$

D'après la proposition 10.10, on peut également obtenir les dérivées partielles de $g\circ f$ en effectuant le produit matriciel entre la matrice jacobienne de g en $f(x_1,x_2)$ et la matrice jacobienne de f en (x_1,x_2) (qui a été calculée à la p. 636). On a

$$\begin{aligned}\mathrm{J}_\varphi(x_1,x_2) &= \mathrm{J}_g(f(x_1,x_2))\times\mathrm{J}_f(x_1,x_2)\\ &= \begin{pmatrix}\dfrac{\partial g}{\partial x_1}(f_1(x_1,x_2),f_2(x_1,x_2)) & \dfrac{\partial g}{\partial x_2}(f_1(x_1,x_2),f_2(x_1,x_2))\end{pmatrix}\\ &\qquad\times\begin{pmatrix}\cos(x_2) & -x_1\sin(x_2)\\ \sin(x_2) & x_1\cos(x_2)\end{pmatrix}\\ &= \\ &\begin{pmatrix}\cos(x_2)\,\dfrac{\partial g}{\partial x_1}(f_1(x_1,x_2),f_2(x_1,x_2)) + \sin(x_2)\,\dfrac{\partial g}{\partial x_2}(f_1(x_1,x_2),f_2(x_1,x_2))\\ -x_1\sin(x_2)\,\dfrac{\partial g}{\partial x_1}(f_1(x_1,x_2),f_2(x_1,x_2)) + x_1\cos(x_2)\,\dfrac{\partial g}{\partial x_2}(f_1(x_1,x_2),f_2(x_1,x_2))\end{pmatrix}\end{aligned}$$

et on retrouve les deux relations (11) et (12).

On se rend compte sur cet exemple de l'ambiguïté de certaines notations lors de la manipulation des dérivées partielles. Ainsi $\frac{\partial g}{\partial x_1}(f_1(x_1,x_2),f_2(x_1,x_2))$ doit être compris comme la dérivée partielle de l'application g selon le premier vecteur de la base canonique de $\mathbb{R}^2$ calculée au point $(f_1(x_1,x_2),f_2(x_1,x_2))$ et non comme la dérivée par rapport à x_1 de la quantité $g(f_1(x_1,x_2),f_2(x_1,x_2))$. On peut pour cette raison préférer pour un calcul comme celui-ci noter $\partial_i g$ les dérivées partielles de g plutôt que $\frac{\partial g}{\partial x_i}$.

Exercice 10.8 Soient g une application de $\mathbb{R}^3$ dans $\mathbb{R}$, différentiable sur $\mathbb{R}^3$ et f l'application

$$f : (x_1,x_2,x_3)\in\mathbb{R}^3 \mapsto (x_1\cos(x_2)\sin(x_3), x_1\sin(x_2)\sin(x_3), x_1\cos(x_3))\in\mathbb{R}^3.$$

Calculer les dérivées partielles de l'application composée $\varphi = g\circ f$ en fonction des dérivées partielles de l'application g.

Remarque Dans le cas d'une application $\varphi : t \in I \mapsto g(f_1(t), \ldots, f_p(t))$ où I est un intervalle fermé de $\mathbb{R}$, $f_1, \ldots, f_p$ des applications de I dans $\mathbb{R}$ dérivables sur I et g une application de $\mathbb{R}^p$ dans $\mathbb{R}$ différentiable sur $\mathbb{R}^p$, la formule (9) s'écrit

$$\forall t \in I \quad \varphi'(t) = f_1'(t)\, \partial_1 g(f_1(t), \ldots, f_p(t)) + \cdots + f_p'(t)\, \partial_p g(f_1(t), \ldots, f_p(t)).$$

◇

Exemple Deux bateaux naviguent en ligne droite sur un plan d'eau (assimilé à un domaine du plan $\mathbb{R}^2$). La position $(a_1(t), a_2(t))$ du premier bateau à l'instant $t \in [0, 2]$ est donnée par

$$a_1(t) = t + 1 \quad \text{et} \quad a_2(t) = 1 - 2t,$$

tandis que la position $(b_1(t), b_2(t))$ du second bateau est donnée à l'instant $t \in [0, 2]$ par

$$b_1(t) = 3t + 1 \quad \text{et} \quad b_2(t) = t - 3.$$

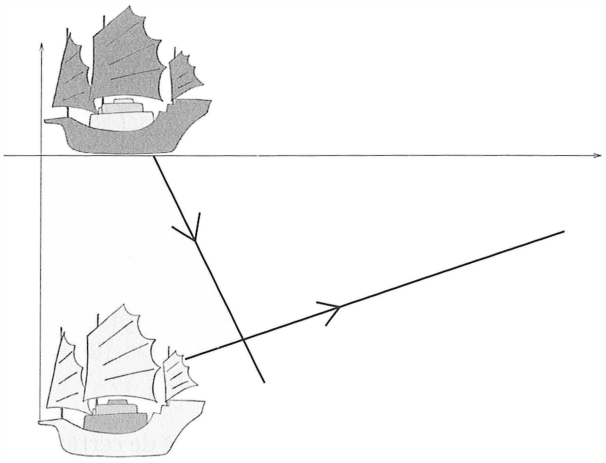


Fig. 1 Trajectoires de deux bateaux sur un plan d'eau.

À chaque instant t, le carré de la distance entre les deux bateaux est donné par $\varphi(t) = g(a_1(t), a_2(t), b_1(t), b_2(t))$ où $g(x_1, x_2, y_1, y_2) = (y_1 - x_1)^2 + (y_2 - x_2)^2$ est le carré de la distance entre les points (x_1, x_2) et (y_1, y_2). Afin de déterminer l'instant où les deux bateaux seront les plus proches, calculons le minimum de l'application φ sur $[0, 2]$. Les applications a_1, a_2, b_1, b_2 et g étant toutes différentiables sur leur ensemble de définition, il en est de même pour l'application φ. Sa dérivée est donnée par

$$\begin{aligned}\varphi'(t) &= a_1'(t)\, \frac{\partial g}{\partial x_1}(a_1(t), a_2(t), b_1(t), b_2(t)) + a_2'(t)\, \frac{\partial g}{\partial x_2}(a_1(t), a_2(t), b_1(t), b_2(t)) \\ &\quad + b_1'(t)\, \frac{\partial g}{\partial x_3}(a_1(t), a_2(t), b_1(t), b_2(t)) + b_2'(t)\, \frac{\partial g}{\partial x_4}(a_1(t), a_2(t), b_1(t), b_2(t)) \\ &= 11t - 12.\end{aligned}$$

L'application φ' s'annule pour $t_0 = 12/11$. Au temps $t_0 = 12/11$, l'application φ atteint un minimum car[38] on a aussi $\varphi''(12/11) = 11 > 0$. Les deux bateaux sont donc les plus proches l'un de l'autre au temps $t_0 = 12/11$, la distance entre les deux bateaux valant alors $\sqrt{\varphi(12/11)} = 8\sqrt{10}/11$. Si les deux bateaux sont assimilés à des disques de rayon 1, vont-ils entrer en collision ?[39]

10.2.4 Cas des applications à valeurs dans $\mathbb{R}$

Lors du calcul de la différentielle d'une application de $\mathbb{R}^n$ dans $\mathbb{R}^p$, on se rend compte que l'espace de départ $\mathbb{R}^n$ et l'espace d'arrivée $\mathbb{R}^p$ ont des rôles différents : pour calculer la différentielle d'une application f d'un ouvert U de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^p$, on est amené à reproduire le même processus calculatoire pour chaque application coordonnée $f_1, \ldots, f_p$ qui sont des applications à valeurs dans $\mathbb{R}$.

Considérons une application f définie sur un ouvert U de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}$, différentiable en $\mathbf{a} \in U$. Pour tout $\mathbf{h} = (h_1, \ldots, h_n) \in \mathbb{R}^n$ on a

$$\mathrm{d}f_{\mathbf{a}}(\mathbf{h}) = \sum_{j=1}^{n} h_j \frac{\partial f}{\partial x_j}(\mathbf{a}).$$

La matrice jacobienne en $\mathbf{a}$ est réduite à une matrice à 1 ligne et n colonnes ; la relation précédente peut aussi s'écrire

$$\boxed{\mathrm{d}f_{\mathbf{a}}(\mathbf{h}) = \mathrm{J}_f(\mathbf{a})\ {}^t\mathbf{h}.} \tag{13}$$

Exemple Considérons l'application

$$f : (x_1, x_2) \in \mathbb{R}^2 \longmapsto \begin{cases} x_1 x_2 \dfrac{x_1^2 - x_2^2}{x_1^2 + x_2^2} & \text{si } (x_1, x_2) \neq (0,0) \\ 0 & \text{si } (x_1, x_2) = (0,0) \end{cases}$$

et déterminons sa dérivée partielle première. Au voisinage de tout point (x_1, x_2) autre que l'origine la dérivée partielle de f par rapport à x_1 vaut :

$$\frac{\partial f}{\partial x_1}(x_1, x_2) = x_2 \frac{x_1^2 - x_2^2}{x_1^2 + x_2^2} + \frac{4x_1^2 x_2^3}{(x_1^2 + x_2^2)^2}.$$

Pour calculer cette dérivée partielle première en $(0,0)$, utilisons la définition 10.5 et évaluons la limite du taux d'accroissement de f en $(0,0)$ dans la direction $\mathbf{e}_1$; on a

$$\frac{f(t,0) - f(0,0)}{t} = \frac{0-0}{t} = 0.$$

---

[38] Une application φ d'un ouvert U de $\mathbb{R}$ à valeurs dans $\mathbb{R}$ qui est de classe C^2 admet un minimum en $t_0 \in U$ dès que $\varphi'(t_0) = 0$ et $\varphi''(t_0) > 0$, voir le *Cours de première année*, Proposition 16.10, p. 736.

[39] Non, car $8\sqrt{10}/11 > 2$.

Ainsi, l'application f admet une dérivée partielle selon le premier vecteur de la base canonique de $\mathbb{R}^2$ donnée par

$$\frac{\partial f}{\partial x_1} : (x_1, x_2) \in \mathbb{R}^2 \longmapsto \begin{cases} x_2 \dfrac{x_1^2 - x_2^2}{x_1^2 + x_2^2} + \dfrac{4x_1^2 x_2^3}{(x_1^2 + x_2^2)^2} & \text{si } (x_1, x_2) \neq (0,0) \\ 0 & \text{si } (x_1, x_2) = (0,0) \end{cases}.$$

Un calcul analogue permet de déterminer la seconde dérivée partielle de f. On a

$$\frac{\partial f}{\partial x_2} : (x_1, x_2) \in \mathbb{R}^2 \longmapsto \begin{cases} x_1 \dfrac{x_1^2 - x_2^2}{x_1^2 + x_2^2} - \dfrac{4x_1^3 x_2^2}{(x_1^2 + x_2^2)^2} & \text{si } (x_1, x_2) \neq (0,0) \\ 0 & \text{si } (x_1, x_2) = (0,0) \end{cases}.$$

On en déduit que les matrices jacobiennes de f en $(1,1)$ et en $(0,0)$ sont respectivement

$$\mathrm{J}_f(1,1) = \begin{pmatrix} 1 & -1 \end{pmatrix} \quad \text{et} \quad \mathrm{J}_f(0,0) = \begin{pmatrix} 0 & 0 \end{pmatrix}.$$

En utilisant la formule (13), on obtient l'expression de la différentielle de f aux deux points $(1,1)$ et $(0,0)$; pour $\mathbf{h} = (h_1, h_2) \in \mathbb{R}^2$ on a

$$\mathrm{d}f_{(1,1)}(h_1, h_2) = \begin{pmatrix} 1 & -1 \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} = h_1 - h_2$$

et $$\mathrm{d}f_{(0,0)}(h_1, h_2) = \begin{pmatrix} 0 & 0 \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} = 0.$$

On notera l'approche utilisée dans l'exemple précédent pour calculer les dérivées partielles de f qui est une application définie par un prolongement par continuité. Pour les points où f est définie par une expression littérale, ici $\mathbb{R}^2_*$, on a recours à la méthode présentée p. 632. Pour le point correspondant au prolongement par continuité, ici $(0,0)$, on a recours à la définition 10.5.

Comment établir la différentiabilité d'une application de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}$?

Nous avons pu constater qu'établir qu'une application était différentiable en ayant recours à la définition 10.1 n'était pas toujours aisé. Tout comme pour étudier la dérivabilité des fonctions de $\mathbb{R}$ dans $\mathbb{R}$, on peut « décomposer » la fonction de $\mathbb{R}^n$ dans $\mathbb{R}$ en sommes, produits, quotients ou composées d'applications dont il est plus aisé d'établir la différentiabilité et utiliser les propriétés algébriques présentées à la section 10.1.2. Pour établir la différentiabilité de cette manière, on a souvent recours au fait que les projections canoniques $(x_1, \ldots, x_n) \in \mathbb{R}^n \mapsto x_i \in \mathbb{R}$ sont des applications différentiables sur $\mathbb{R}^n$. Une approche analogue peut être mise en œuvre pour montrer qu'une application est de classe $\mathcal{C}^1$. Nous allons illustrer cette façon de procéder sur un exemple.

Exemple Montrons que l'application f de $\mathbb{R}^2$ dans $\mathbb{R}$ définie par

$$f(x_1, x_2) = \begin{cases} x_2^2 \sin\left(\dfrac{x_1}{x_2}\right) & \text{si } x_2 \neq 0 \\ 0 & \text{si } x_2 = 0 \end{cases}$$

est différentiable sur $\mathbb{R}^2$ mais n'est pas de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$.
L'application f est de classe $\mathcal{C}^1$ sur l'ouvert $U = \{(x_1, x_2) \in \mathbb{R}^2 \ ; \ x_2 \neq 0\}$ car en restriction à U, f est le produit de l'application $(x_1, x_2) \in U \mapsto x_2^2$ qui est de classe $\mathcal{C}^1$ sur U et de l'application $(x_1, x_2) \in U \mapsto \sin(x_1/x_2)$ qui est de classe $\mathcal{C}^1$ sur U en tant que composée de l'application $(x_1, x_2) \in U \mapsto x_1/x_2$ de classe $\mathcal{C}^1$ sur U (le dénominateur ne s'annulant pas sur U) et de la fonction sinus qui est de classe $\mathcal{C}^1$ sur $\mathbb{R}$. Nous allons prouver successivement que :

1. L'application f est continue sur $\mathbb{R}^2$,
2. L'application f est différentiable sur $\mathbb{R}^2$,
3. L'application f n'est pas de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$.

Comme f est de classe $\mathcal{C}^1$ sur U, il s'agit d'étudier la continuité, la différentiabilité et la continuité des dérivées partielles sur $V = \{(x_1, 0) \ ; \ x_1 \in \mathbb{R}\}$.

1. Commençons par montrer la continuité de f sur V, c'est-à-dire, étudions la continuité de f en tous les points de la forme $(a_1, 0)$. Pour tout $t \in \mathbb{R}$ on a $|\sin t| \leqslant |t|$ et par conséquent pour $x_1 \in \mathbb{R}$ et $x_2 \in \mathbb{R}^*$ on a $|f(x_1, x_2)| \leqslant |x_1 x_2|$. On en déduit que $\lim_{(x_1,x_2)\to(a_1,0)} f(x_1, x_2) = 0$. Comme $f(a_1, 0) = 0$, l'application f est continue en tout point de la forme $(a_1, 0)$, $a_1 \in \mathbb{R}$.

2. Montrons que l'application f admet des dérivées partielles en $(a_1, 0)$. Pour calculer la première dérivée partielle, on forme le taux d'accroissement de f en $(a_1, 0)$ dans la direction $\mathbf{e}_1$:

$$\frac{f(a_1 + t, 0) - f(a_1, 0)}{t} = 0 \quad \text{pour } t \in \mathbb{R}^*.$$

Celui-ci étant constant, sa limite lorsque t tend vers 0 existe (et vaut 0). De même, pour calculer la seconde dérivée partielle, on forme le taux d'accroissement de f en $(a_1, 0)$ dans la direction $\mathbf{e}_2$:

$$\frac{f(a_1, t) - f(a_1, 0)}{t} = t \sin\left(\frac{a_1}{t}\right) \quad \text{pour } t \in \mathbb{R}^*.$$

Comme la fonction sinus est bornée, ce taux d'accroissement admet 0 pour limite lorsque t tend vers 0. On en déduit que les deux dérivées partielles de f existent en tout point de la forme $(a_1, 0)$ et qu'elles valent 0.

Montrons maintenant que l'application f est différentiable en $(a_1, 0)$. D'après la proposition 10.9, si f est différentiable en $(a_1, 0)$ alors sa différentielle en $(a_1, 0)$ est donnée par

$$\mathrm{d}f_{(a_1,0)} : \mathbf{h} \in \mathbb{R}^2 \longmapsto \sum_{i=1}^{2} \frac{\partial f}{\partial x_i}(a_1, 0) h_i$$

et coïncide avec l'application nulle. Comme pour tout $(h_1, h_2) \in \mathbb{R}^2$ on a

$$\begin{aligned} f(a_1 + h_1, h_2) - f(a_1, 0) &= h_2^2 \sin\left(\frac{a_1 + h_1}{h_2}\right) \\ &= \mathrm{d}f_{(a_1,0)}(h_1, h_2) + \sqrt{h_1^2 + h_2^2}\, \varepsilon(h_1, h_2) \end{aligned}$$

avec

$$\varepsilon : (h_1, h_2) \in \mathbb{R}^2 \longmapsto \frac{h_2^2}{\sqrt{h_1^2 + h_2^2}} \sin\left(\frac{a_1 + h_1}{h_2}\right),$$

et que la fonction ε vérifie pour tout $(h_1, h_2) \in \mathbb{R}^2$

$$|\varepsilon(h_1, h_2)| \leqslant \frac{h_2^2}{\sqrt{h_1^2 + h_2^2}} \leqslant \frac{h_1^2 + h_2^2}{\sqrt{h_1^2 + h_2^2}} \leqslant \sqrt{h_1^2 + h_2^2} = \|(h_1, h_2)\|_2,$$

on en déduit que $\lim_{(h_1,h_2)\to(0,0)} \varepsilon(h_1, h_2) = 0$ et on peut conclure que f est différentiable en $(a_1, 0)$, de différentielle nulle.

On en conclut que f est différentiable sur $\mathbb{R}^2$.

3. Intéressons-nous à la continuité de la seconde dérivée partielle de f qui est donnée par :

$$\frac{\partial f}{\partial x_2} : (x_1, x_2) \in \mathbb{R}^2 \longmapsto \begin{cases} 2x_2 \sin\left(\dfrac{x_1}{x_2}\right) - x_1 \cos\left(\dfrac{x_1}{x_2}\right) & \text{si } x_2 \neq 0 \\ 0 & \text{si } x_2 = 0 \end{cases}$$

Cette application n'est pas continue en un point de la forme $(a_1, 0)$ avec $a_1 \neq 0$ puisque la limite de $\frac{\partial f}{\partial x_2}(a_1, x_2)$ lorsque x_2 tend vers 0 n'existe pas (le terme $2x_2 \sin(a_1/x_2)$ tend vers 0 lorsque x_2 tend vers 0 mais le terme $a_1 \cos(a_1/x_2)$ n'a pas de limite). Puisque[40] la dérivée seconde de f n'est pas continue sur $\mathbb{R}^2$, l'application f n'est pas de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$.

Exercice 10.9 Soit f l'application de $\mathbb{R}^3$ dans $\mathbb{R}$ définie par

$$f(x_1, x_2, x_3) = \begin{cases} \dfrac{x_1 x_2}{x_3^2} & \text{si } x_3 \neq 0, \\ 0 & \text{si } x_3 = 0. \end{cases}$$

1 - Calculer les trois dérivées partielles de f au point $(1, 1, 1)$.
2 - Montrer que f est différentiable en ce point et calculer sa différentielle.
3 - L'application f est-elle différentiable en $(0, 0, 0)$?

[40] C'est le théorème 10.2, et plus exactement sa contraposée qui nous permet d'affirmer que si l'une des dérivées partielles n'est pas continue alors l'application n'est pas de classe $\mathcal{C}^1$.

10.3 Applications de classe $\mathcal{C}^q$

10.3.1 Notion de dérivées d'ordre supérieur

Soit f une application définie[41] sur un ouvert U de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^p$. Si f est une application différentiable sur U alors sa différentielle $\mathrm{d}f$ est une application définie sur U et à valeurs dans l'espace vectoriel $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)$ des applications linéaires de $\mathbb{R}^n$ dans $\mathbb{R}^p$. On peut étudier la différentielle de $\mathrm{d}f$ en un point $\mathbf{a} \in U$ qui si elle existe est notée $\mathrm{d}^2 f_{\mathbf{a}}$. Par définition $\mathrm{d}^2 f_{\mathbf{a}}$ est une application linéaire continue de $\mathbb{R}^n$ dans $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)$, c'est-à-dire que l'on a $\mathrm{d}^2 f_{\mathbf{a}} \in \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^p))$.

Quelques cas particuliers

$\trianglerighteq$ Soit f une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^p$. On sait que[42] toute application linéaire est différentiable et que pour tout $\mathbf{a} \in \mathbb{R}^n$ on a $\mathrm{d}f_{\mathbf{a}} = f$. Ainsi, l'application $\mathrm{d}f$ est une application constante de $\mathbb{R}^n$ dans $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)$. Or[43] la différentielle d'une application constante est l'application nulle, par conséquent pour tout $\mathbf{a} \in \mathbb{R}^n$ on a $\mathrm{d}^2 f_{\mathbf{a}} = 0$.

$\trianglerighteq$ Soit f une application définie sur un intervalle ouvert I de $\mathbb{R}$ à valeurs dans $\mathbb{R}$ deux fois dérivable sur I (au sens usuel des fonctions de $\mathbb{R}$ dans $\mathbb{R}$). Cette application f est différentiable[44] en tout point $a \in I$ et on a

$$\mathrm{d}f_a : h \in \mathbb{R} \longmapsto f'(a)\, h \in \mathbb{R}.$$

La différentielle de f est l'application

$$\mathrm{d}f : a \in I \longmapsto \big(h \mapsto f'(a)\, h\big) \in \mathcal{L}(\mathbb{R}, \mathbb{R}).$$

Calculons la différentielle en $a \in I$ de l'application $\mathrm{d}f$. Pour $(h, k) \in \mathbb{R}^2$, on a $\mathrm{d}f_{a+k}(h) = f'(a+k)h$. Puisque l'application f est une application de I dans $\mathbb{R}$ deux fois dérivable sur I, la formule de Taylor-Young[45] s'écrit

$$f'(a+k) = f'(a) + f''(a)k + o_0(k^2)$$

de sorte que

$$\mathrm{d}f_{a+k}(h) = f'(a)\ h + f''(a)\ kh + o_0(k^2)\ h = \mathrm{d}f_a(h) + f''(a)\ kh + o_0(k^2)\ h.$$

Ainsi, pour tout $k \in \mathbb{R}$ on a $\mathrm{d}f_{a+k} = \mathrm{d}f_a + L(k) + \mathcal{O}_0(k^2)$ où l'application L définie par $L : k \in \mathbb{R} \mapsto (h \mapsto f''(a)kh) \in \mathcal{L}(\mathbb{R}, \mathbb{R})$ est une application linéaire. Ceci montre que l'application $\mathrm{d}f$ est différentiable en $a \in I$ et que

$$\mathrm{d}^2 f_a : k \in \mathbb{R} \mapsto \big(h \mapsto f''(a)\ kh\big) \in \mathcal{L}(\mathbb{R}, \mathbb{R}).$$

---

[41] Nous nous plaçons ici dans le cas d'une application définie sur un ouvert de $\mathbb{R}^n$ mais les résultats énoncés dans cette partie peuvent se généraliser aux applications définies sur un ouvert d'un espace vectoriel normé E et à valeurs dans un espace vectoriel normé F.

[42] Voir la proposition 10.5 p. 615.

[43] Voir l'exemple p. 610.

[44] Voir la proposition 10.3 p. 614.

[45] Voir le théorème 17.1 p. 778 du *Cours de première année*.

Cadre général

L'ensemble $\mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^p))$ auquel appartiennent les applications $\mathrm{d}^2 f_{\mathbf{a}}$ peut s'identifier à l'ensemble des applications bilinéaires[46] de $\mathbb{R}^n \times \mathbb{R}^n$ dans $\mathbb{R}^p$, noté $\mathcal{L}_2(\mathbb{R}^n, \mathbb{R}^p)$. Considérons une application bilinéaire $\varphi \in \mathcal{L}_2(\mathbb{R}^n, \mathbb{R}^p)$; pour tout $\mathbf{x} \in \mathbb{R}^n$ fixé, l'application $\mathbf{y} \in \mathbb{R}^n \mapsto \varphi(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^p$ est une application linéaire appartenant à l'espace vectoriel $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)$ et par conséquent l'application

$$\mathbf{x} \in \mathbb{R}^n \longmapsto \big(\mathbf{y} \mapsto \varphi(\mathbf{x}, \mathbf{y})\big) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)$$

est une application linéaire appartenant à l'espace vectoriel $\mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^p))$. On peut montrer que l'application

$$\begin{aligned} \Phi : \mathcal{L}_2(\mathbb{R}^n, \mathbb{R}^p) &\longrightarrow \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)) \\ \varphi &\longmapsto \big(\mathbf{x} \mapsto \varphi(\mathbf{x}, .)\big) \end{aligned}$$

est un isomorphisme d'espace vectoriel. Il permet d'identifier les deux espaces vectoriels. Ainsi, en général l'application $\mathrm{d}^2 f_a$ sera identifié à une application bilinéaire de $\mathbb{R}^n \times \mathbb{R}^n$ dans $\mathbb{R}^p$.

Considérons à nouveau le cas d'une application f définie sur un intervalle ouvert I de $\mathbb{R}$ et à valeurs dans $\mathbb{R}$, deux fois dérivable sur I (au sens usuel des fonctions de $\mathbb{R}$ dans $\mathbb{R}$). Pour $a \in I$ fixé, notons φ l'application bilinéaire de $\mathbb{R} \times \mathbb{R}$ dans $\mathbb{R}$ définie par

$$\varphi : (h, k) \in \mathbb{R} \times \mathbb{R} \longmapsto f''(a)hk.$$

Cette application bilinéaire peut être identifiée, en utilisant l'isomorphisme Φ, à l'application

$$\Phi(\varphi) : k \in \mathbb{R} \mapsto \big(h \mapsto f''(a)kh\big) \in \mathcal{L}(\mathbb{R}, \mathbb{R}).$$

Autrement dit, on a $\Phi(\varphi) = \mathrm{d}^2 f_a$ et la différentielle seconde de f en $a \in I$ pourra être identifiée à la forme bilinéaire $\mathrm{d}^2 f_a : (h, k) \in \mathbb{R}^2 \mapsto f''(a)kh \in \mathbb{R}$.

On peut itérer le procédé et définir la différentielle d'ordre $q \in \mathbb{N}^*$ d'une application f en $\mathbf{a} \in U$ que l'on note $\mathrm{d}^q f_{\mathbf{a}}$. Cette application est isomorphe à une application de $\mathcal{L}_q(\mathbb{R}^n, \mathbb{R}^p)$, l'espace vectoriel des applications q-linéaires de $(\mathbb{R}^n)^q$ dans $\mathbb{R}^p$, c'est-à-dire linéaire en chacune des q variables. Lorsque la différentielle d'ordre q d'une application f existe, on dit que f est q fois différentiable.

On introduit par ailleurs, la définition suivante.

---

[46] Rappelons, voir la définition 8.1 p. 438, qu'une application de $E \times E$ à valeurs dans F est dite *bilinéaire* lorsqu'elle est linéaire par rapport à chacune de ses variables :

$$\forall y \in E \quad x \mapsto f(x, y) \quad \text{est linéaire} \quad \text{et} \quad \forall x \in E \quad y \mapsto f(x, y) \quad \text{est linéaire.}$$

DÉFINITION 10.8 *Soient f une application d'un ouvert U de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^p$ et q un entier naturel non nul.*

✗ *On dit que f est de* classe C^q *si f est q fois différentiable et si $\mathrm{d}^q f$ est continue.*

✗ *On dit que f est de* classe C^∞ *si elle est de classe C^r pour tout $r \in \mathbb{N}^*$.*

10.3.2 Dérivations partielles d'ordre supérieur

Soit f une application d'un ouvert U de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^p$, différentiable sur U ; elle admet des dérivées partielles sur U et dans le cas où ces dérivées partielles sont elles-mêmes des applications différentiables, on peut calculer leurs dérivées partielles. On définit ainsi ce que l'on appelle les *dérivées partielles secondes.*

Exemple Considérons l'application de $\mathbb{R}^2$ dans $\mathbb{R}$ définie par

$$f : (x_1, x_2) \in \mathbb{R}^2 \longmapsto \begin{cases} x_1 x_2 \dfrac{x_1^2 - x_2^2}{x_1^2 + x_2^2} & \text{si } (x_1, x_2) \neq (0,0) \\ 0 & \text{si } (x_1, x_2) = (0,0) \end{cases}.$$

Nous avons établi p. 641 que cette application est différentiable sur $\mathbb{R}^2$ et que sa première dérivée partielle d'ordre 1 est

$$\frac{\partial f}{\partial x_1} : (x_1, x_2) \in \mathbb{R}^2 \longmapsto \begin{cases} x_2 \dfrac{x_1^2 - x_2^2}{x_1^2 + x_2^2} + \dfrac{4x_1^2 x_2^3}{(x_1^2 + x_2^2)^2} & \text{si } (x_1, x_2) \neq (0,0) \\ 0 & \text{si } (x_1, x_2) = (0,0) \end{cases}.$$

Pour calculer la seconde dérivée partielle d'ordre 1 de l'application $\frac{\partial f}{\partial x_1}$ en $(0,0)$, on évalue la limite du taux d'accroissement de $\frac{\partial f}{\partial x_1}$ en $(0,0)$ dans la direction du vecteur $\mathbf{e}_2$; on a

$$\lim_{t \to 0} \frac{\frac{\partial f}{\partial x_1}(0,t) - \frac{\partial f}{\partial x_1}(0,0)}{t} = \lim_{t \to 0} \frac{-t - 0}{t} = -1.$$

On en déduit que la seconde dérivée partielle de $\frac{\partial f}{\partial x_1}$ en $(0,0)$ existe et vaut

$$\frac{\partial \frac{\partial f}{\partial x_1}}{\partial x_2}(0,0) = -1.$$

Cette dérivée partielle est notée $\dfrac{\partial^2 f}{\partial x_2 \partial x_1}(0,0)$ ou $\partial^2_{21} f(0,0)$.

Pour calculer la seconde dérivée partielle d'ordre 1 de l'application $\frac{\partial f}{\partial x_1}$ en $(x_1, x_2) \neq (0,0)$, on peut avoir recours à la méthode présentée p. 632. On dérive par rapport à la variable x_2, la variable x_1 étant fixée, la quantité

$$x_2 \frac{x_1^2 - x_2^2}{x_1^2 + x_2^2} + \frac{4x_1^2 x_2^3}{(x_1^2 + x_2^2)^2}.$$

On obtient[47]

$$\frac{\partial^2 f}{\partial x_2 \partial x_1}(x_1, x_2) = \frac{\partial \frac{\partial f}{\partial x_1}}{\partial x_2}(x_1, x_2) = \frac{x_1^6 + 9\,x_1^4 x_2^2 - 9\,x_1^2 x_2^4 - x_2^6}{\left(x_1^2 + x_2^2\right)^3}.$$

Plus généralement, pour $(i, j) \in \{1, \ldots, n\}^2$, la i–ième dérivée partielle de l'application $\frac{\partial f}{\partial x_j}$ au point $\mathbf{a} \in U$ est notée $\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{a})$ ou encore $\partial^2_{ij} f(\mathbf{a})$. Lorsque $i = j$, cette dérivée partielle est notée $\frac{\partial^2 f}{\partial x_i^2}(\mathbf{a})$ ou $\partial^2_{ii} f(\mathbf{a})$.

Lien entre dérivées partielles et différentielle d'ordre 2

Rappelons que pour une application f d'un ouvert U de $\mathbb{R}^n$ dans $\mathbb{R}^p$, différentiable en $\mathbf{a} \in U$, la différentielle de f en $\mathbf{a}$ est une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^p$ qui s'exprime à l'aide des dérivées partielles de f sous la forme[48] :

$$\mathrm{d}f_{\mathbf{a}} : \mathbf{h} \in \mathbb{R}^n \longmapsto \sum_{i=1}^{n} h_i\, \partial_i f(\mathbf{a}) \in \mathbb{R}^p.$$

Notons g l'application $\mathbf{a} \in U \mapsto \mathrm{d}f_{\mathbf{a}} \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)$ et supposons que cette application soit différentiable en $\mathbf{a} \in U$; on a alors

$$\mathrm{d}g_{\mathbf{a}} : \mathbf{k} \in \mathbb{R}^n \longmapsto \sum_{i=1}^{n} k_i\, \partial_i g(\mathbf{a}) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^p). \tag{14}$$

Pour tout $i \in \{1, \ldots, n\}$, la quantité $\partial_i g(\mathbf{a})$ est une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^p$ qui par définition satisfait

$$\partial_i g(\mathbf{a}) = \lim_{t \to 0} \frac{g(\mathbf{a} + t\,\mathbf{e}_i) - g(\mathbf{a})}{t},$$

où $\mathbf{e}_i$ désigne le i–ième vecteur de la base canonique de $\mathbb{R}^n$. Puisque $g(\mathbf{a}) = \mathrm{d}f_{\mathbf{a}}$, pour tout $\mathbf{h} \in \mathbb{R}^n$ on a $g(\mathbf{a})(\mathbf{h}) = \mathrm{d}f_{\mathbf{a}}(\mathbf{h}) = \sum_{j=1}^{n} h_j \partial_j f(\mathbf{a})$ de sorte que

$$\frac{g(\mathbf{a} + t\mathbf{e}_i) - g(\mathbf{a})}{t}(\mathbf{h}) = \sum_{j=1}^{n} h_j\, \frac{\partial_j f(\mathbf{a} + t\mathbf{e}_i) - \partial_j f(\mathbf{a})}{t}.$$

Lorsque le réel t tend vers 0, chaque fraction de la somme dans le membre de droite tend vers la dérivée seconde $\partial^2_{ij} f(\mathbf{a})$; on a ainsi pour tout $i \in \{1, \ldots, n\}$

$$\partial_i g(\mathbf{a})(\mathbf{h}) = \sum_{j=1}^{n} h_j\, \partial^2_{ij} f(\mathbf{a}). \tag{15}$$

[47] Les calculs ont été effectués au cours de l'exercice 10.5 p. 632.

[48] Voir la proposition 10.9 p. 632.

En utilisant l'expression (14) et la relation (15), la différentielle de g en $\mathbf{a}$, c'est-à-dire la différentielle seconde de f en $\mathbf{a}$ s'exprime[49] de la façon suivante :

$$\boxed{\mathrm{d}^2 f_{\mathbf{a}} : (\mathbf{h},\mathbf{k}) \in \mathbb{R}^n \times \mathbb{R}^n \longmapsto \sum_{i=1}^{n}\sum_{j=1}^{n} h_i k_j \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{a}).} \tag{16}$$

Ordre des dérivations : le théorème de Schwarz

Étant donné une application f définie sur un ouvert U de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^p$, on peut se demander si pour $\mathbf{a} \in U$ le vecteur $\partial^2_{ij} f(\mathbf{a})$ est égal au vecteur $\partial^2_{ji} f(\mathbf{a})$, autrement dit si l'ordre de la dérivation peut être permuté. La réponse à cette question est que cette égalité n'a pas lieu sans hypothèse supplémentaire. Reprenons l'exemple de la p. 647 et considérons l'application de $\mathbb{R}^2$ dans $\mathbb{R}$ définie par

$$f : (x_1,x_2) \in \mathbb{R}^2 \longmapsto \begin{cases} x_1 x_2 \dfrac{x_1^2 - x_2^2}{x_1^2 + x_2^2} & \text{si } (x_1,x_2) \neq (0,0) \\ 0 & \text{si } (x_1,x_2) = (0,0) \end{cases}.$$

Après avoir établi p. 647 que $\partial^2_{21} f(0,0) = -1$, calculons la dérivée partielle $\partial^2_{12} f$ en $(0,0)$. Autrement dit, évaluons la dérivée partielle selon le premier vecteur de la base canonique de $\mathbb{R}^2$ en $(0,0)$ de l'application[50]

$$\frac{\partial f}{\partial x_2} : (x_1,x_2) \in \mathbb{R}^2 \longmapsto \begin{cases} x_1 \dfrac{x_1^2 - x_2^2}{x_1^2 + x_2^2} - \dfrac{4x_1^3 x_2^2}{(x_1^2 + x_2^2)^2} & \text{si } (x_1,x_2) \neq (0,0) \\ 0 & \text{si } (x_1,x_2) = (0,0) \end{cases}.$$

On a

$$\lim_{t\to 0} \frac{\partial_2 f(t,0) - \partial_2 f(0,0)}{t} = \lim_{t\to 0} \frac{t-0}{t} = 1,$$

et par conséquent $\partial^2_{12} f(0,0) = 1$; on en déduit que $\partial^2_{21} f(0,0) \neq \partial^2_{12} f(0,0)$.

Le théorème suivant donne une condition suffisante pour qu'il y ait égalité entre $\partial^2_{ij} f(\mathbf{a})$ et $\partial^2_{ji} f(\mathbf{a})$. On remarquera que l'application f de l'exemple précédent ne satisfait pas aux hypothèses du théorème.

---

[49] *Via* l'identification entre les applications bilinéaires et les applications linéaires à valeurs dans l'ensemble des applications linéaires donnée par l'isomorphisme Φ, voir p. 646.

[50] Voir l'exemple à la p. 641 pour le calcul de $\frac{\partial f}{\partial x_2}$.

THÉORÈME 10.3 (Théorème de Schwarz)
Soient f une application d'un ouvert U de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^p$ et $\mathbf{a} \in U$.

- *Si pour $(i,j) \in \{1,\dots,n\}^2$, les deux dérivées partielles secondes $\dfrac{\partial^2 f}{\partial x_i \partial x_j}$ et $\dfrac{\partial^2 f}{\partial x_j \partial x_i}$ sont des applications continues en $\mathbf{a} \in U$ alors*

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{a}) = \frac{\partial^2 f}{\partial x_j \partial x_i}(\mathbf{a}).$$

- *Si l'application f est de classe C^2 sur U alors*

$$\forall \mathbf{a} \in U \quad \forall (i,j) \in \{1,\dots,n\}^2 \quad \frac{\partial f}{\partial x_i \partial x_j}(\mathbf{a}) = \frac{\partial f}{\partial x_j \partial x_i}(\mathbf{a}).$$

Remarques

1. Puisque d'après l'exemple précédent, l'application

$$f : (x_1,x_2) \in \mathbb{R}^2 \longmapsto \begin{cases} x_1 x_2 \dfrac{x_1^2 - x_2^2}{x_1^2 + x_2^2} & \text{si } (x_1,x_2) \neq (0,0) \\ 0 & \text{si } (x_1,x_2) = (0,0) \end{cases}$$

ne satisfait pas la conclusion du théorème de Schwarz en $(0,0)$, il s'ensuit qu'elle ne vérifie pas les hypothèses de ce théorème. Ainsi, on peut en conclure que au moins une des deux applications $\partial^2_{12} f$ ou $\partial^2_{21} f$ n'est pas continue en $(0,0)$.

2. Nous verrons qu'il est possible d'établir que les dérivées partielles secondes sont continues sans les calculer, ce qui justifie l'intérêt du théorème 10.3. ◇

Exercice 10.10 Considérons l'application f définie sur $\mathbb{R}^2$ par

$$f(x_1,x_2) = \begin{cases} x_1^2 \sin(x_2/x_1) & \text{lorsque } x_1 \neq 0, \\ 0 & \text{lorsque } x_1 = 0. \end{cases}$$

1 - Déterminer les applications $\partial_1 f$ et $\partial_2 f$ sur $\mathbb{R}^2$.
2 - Calculer $\partial^2_{21} f(0,0)$ et $\partial^2_{12} f(0,0)$.

Dans le cas d'une application de $\mathbb{R}^n$ dans $\mathbb{R}^p$, on peut établir qu'elle est de classe C^q en étudiant la continuité de ces dérivées partielles. Plus précisément, on a la proposition suivante.

PROPOSITION 10.11 *Soit f une application d'un ouvert U de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^p$ et $q \in \mathbb{N}^*$. Si toutes les dérivées partielles d'ordre q de f sont continues sur U alors l'application f est une application de classe C^q sur U.*

Démonstration Cette proposition se démontre par récurrence sur l'entier q en utilisant le théorème 10.2 p. 634. La rédaction du raisonnement par récurrence peut être effectuée à titre d'exercice. □

Remarques

1. Considérons une application f d'un ouvert U de $\mathbb{R}^n$ dans $\mathbb{R}^p$ de classe $\mathcal{C}^q$ sur U et un élément $\mathbf{a}$ de U. On sait[51] exprimer la différentielle de f en $\mathbf{a}$ à l'aide des dérivées partielles de f en $\mathbf{a}$:

$$\boxed{\mathrm{d}f_{\mathbf{a}} : \mathbf{h} \in \mathbb{R}^n \longmapsto \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(\mathbf{a}).}$$

On sait[52] aussi exprimer la différentielle seconde de f en $\mathbf{a}$ à l'aide des dérivées partielles seconde de f en $\mathbf{a}$:

$$\boxed{\mathrm{d}^2 f_{\mathbf{a}} : (\mathbf{h}, \mathbf{k}) \in \mathbb{R}^n \times \mathbb{R}^n \longmapsto \sum_{i=1}^{n} \sum_{j=1}^{n} h_i k_j \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{a}).}$$

De même, il est possible d'exprimer la différentielle d'ordre q de f en $\mathbf{a}$ à l'aide des dérivées partielles d'ordre q de f en $\mathbf{a}$:

$$\boxed{\mathrm{d}^q f_{\mathbf{a}} : (\mathbf{h}_1, \ldots, \mathbf{h}_q) \in (\mathbb{R}^n)^q \longmapsto \sum_{i_1=1}^{n} \cdots \sum_{i_q=1}^{n} h_{i_1} \cdots h_{i_q} \frac{\partial^q f}{\partial x_{i_1} \cdots \partial x_{i_q}}(\mathbf{a}).}$$

2. Le théorème de Schwarz se généralise aux applications de classe $\mathcal{C}^q$: si f est une application de classe $\mathcal{C}^q$ alors pour calculer une dérivée partielle d'ordre inférieure ou égale à q, on peut effectuer les dérivées partielles dans l'ordre que l'on souhaite. ◇

10.3.3 Propriétés des applications de classe $\mathcal{C}^q$

Pour $q \in \mathbb{N}^* \cup \{\infty\}$, on note $\mathcal{C}^q(U, \mathbb{R}^p)$ l'ensemble des applications définies sur l'ouvert U de $\mathbb{R}^n$ et à valeurs dans $\mathbb{R}^p$ qui sont de classe $\mathcal{C}^q$. On note aussi $\mathcal{C}^0(U, \mathbb{R}^p)$ ou plus simplement $\mathcal{C}(U, \mathbb{R}^p)$ l'ensemble des applications continues de U dans $\mathbb{R}^p$. Notons que si une application f définie sur un ouvert U de $\mathbb{R}^n$ et à valeurs dans $\mathbb{R}^p$ est une application de classe $\mathcal{C}^q$ avec $q \geqslant 2$, alors elle est aussi de classe $\mathcal{C}^{q-1}$; en d'autres termes on a les inclusions suivantes :

$$\mathcal{C}^\infty(U, \mathbb{R}^p) \subset \cdots \subset \mathcal{C}^q(U, \mathbb{R}^p) \subset \mathcal{C}^{q-1}(U, \mathbb{R}^p) \subset \cdots \subset \mathcal{C}^1(U, \mathbb{R}^p) \subset \mathcal{C}(U, \mathbb{R}^p).$$

[51] Voir la proposition 10.9.

[52] Voir la relation (16) p. 649.

Étant donné un ouvert U de $\mathbb{R}^n$, l'ensemble $\mathcal{C}(U, \mathbb{R}^p)$ des applications continues de U dans $\mathbb{R}^p$ est un espace vectoriel. Pour $q \in \mathbb{N}^* \cup \{\infty\}$, l'ensemble $\mathcal{C}^q(U, \mathbb{R}^p)$ est un sous-espace vectoriel de $\mathcal{C}(U, \mathbb{R}^p)$. Cela se traduit pour les applications de classe C^q par la proposition suivante :

PROPOSITION 10.12 *Soient f et g deux applications de classe C^q définie sur un ouvert U de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^p$ et λ et μ deux réels. L'application $\lambda f + \mu g$ est une application de classe C^q sur U.*

Cette proposition, comme les deux suivantes, résulte des propriétés des dérivées partielles et du lien existant entre la régularité des dérivées partielles d'une application f et la régularité de l'application f elle-même.

PROPOSITION 10.13 *Soient f une application définie sur un ouvert U de $\mathbb{R}^n$ à valeurs dans un ouvert V de $\mathbb{R}^p$, g une application de V à valeurs dans $\mathbb{R}^m$. Si f est une application de classe C^q sur U et si g une application de classe C^q sur V alors $g \circ f$ est une application de classe C^q sur U.*

PROPOSITION 10.14 *Soient f et g deux applications définies sur un ouvert U de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}$.*

✗ *Si f et g sont deux applications de classe C^q sur U alors le produit $f \times g$ est une application de classe C^q sur U.*

✗ *Si f et g sont deux applications de classe C^q sur U et que g ne s'annule pas sur U alors le quotient f/g est une application de classe C^q sur U.*

Remarques

1. La proposition 10.14 permet de prouver que le produit d'applications de classe C^q à valeurs réelles est une application de classe C^q. Ce résultat est valide dès que l'espace d'arrivée est muni d'une structure sur laquelle on peut effectuer un « produit ». Ainsi le produit de deux applications qui sont à valeurs dans une algèbre[53] et de classe C^q est aussi de classe C^q.

2. En pratique, on utilise les propositions 10.12, 10.13 et 10.14 pour établir qu'une application est de classe C^q comme cela est fait pour montrer qu'une application est continue, voir p. 549. En particulier, c'est à l'aide de ces propositions qu'on établit le plus souvent qu'une application est de classe C^2 afin d'appliquer le théorème de Schwarz. ◇

[53] La définition d'une algèbre est donnée à la p. 544.

10.3.4 Cas des applications à valeurs dans $\mathbb{R}$

Comme nous l'avons déjà indiqué, l'étude d'une fonction de $\mathbb{R}^n$ dans $\mathbb{R}^p$ se ramène le plus souvent à l'étude de chacune de ses applications coordonnées qui sont des fonctions à valeurs réelles. Cela justifie l'intérêt particulier porté à l'étude des fonctions à valeurs réelles.

DÉFINITION 10.9 *Soit f une application définie sur un ouvert U de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}$, supposée deux fois différentiable en $\mathbf{a} \in U$. La* matrice hessienne[(54)] *de f en $\mathbf{a}$ est la matrice de $\mathcal{M}_n(\mathbb{R})$ définie par :*

$$\mathrm{H}_f(\mathbf{a}) = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1 \partial x_1}(\mathbf{a}) & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n}(\mathbf{a}) \\ \vdots & & \vdots \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1}(\mathbf{a}) & \cdots & \dfrac{\partial^2 f}{\partial x_n \partial x_n}(\mathbf{a}) \end{pmatrix}.$$

Remarques

1. On appelle « matrice hessienne » d'une application f de U dans $\mathbb{R}$ la « matrice » dont les éléments sont les applications partielles de f :

$$\mathrm{H}_f = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1 \partial x_1} & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & & \vdots \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \dfrac{\partial^2 f}{\partial x_n \partial x_n} \end{pmatrix}.$$

On sera vigilant au fait que ce tableau n'est pas une matrice au sens défini dans le *Cours de première année* car les éléments sont des applications de U dans $\mathbb{R}$ et l'ensemble des applications de U dans $\mathbb{R}$ n'est pas un corps.

2. Considérons une application $f \in \mathcal{C}^2(U, \mathbb{R})$ où U est un ouvert de $\mathbb{R}^n$ et $\mathbf{a} \in U$. D'après la relation (16) p. 649, la forme bilinéaire $\mathrm{d}^2 f_{\mathbf{a}}$ a pour matrice dans la base canonique de $\mathbb{R}^n$, la matrice hessienne de f en $\mathbf{a}$ ce qui signifie que pour tous $\mathbf{h} = (h_1, \ldots, h_n) \in \mathbb{R}^n$ et $\mathbf{k} = (k_1, \ldots, k_n) \in \mathbb{R}^n$ on a

$$\boxed{\mathrm{d}^2 f_a(\mathbf{h}, \mathbf{k}) = \mathbf{h}\ \mathrm{H}_f(\mathbf{a})\ {}^t\mathbf{k}.} \tag{17}$$

3. Le théorème de Schwarz implique que si f est une application de classe $\mathcal{C}^2$ sur U alors sa matrice hessienne en $\mathbf{a}$ est symétrique pour tout $\mathbf{a} \in U$. ◇

(54) HESSE, Ludwig-Otto (1811, Kaliningrad - 1874, Munich). Hesse travailla essentiellement sur le développement de la théorie des fonctions algébriques. Il introduisit la matrice hessienne en 1842 lors de l'étude de courbes quadratiques et cubiques.

Exemple Considérons l'application f de $U =]-1,1[^2$ dans $\mathbb{R}$ définie par

$$f : (x_1, x_2) \in U \longmapsto -\frac{1+x_1}{((1+x_1)^2+x_2^2)^{3/2}} \in \mathbb{R}.$$

Cette application est de classe $\mathcal{C}^\infty$ car c'est le quotient de la fonction $(x_1, x_2) \mapsto 1 + x_1$ qui est de classe $\mathcal{C}^\infty$ sur U par la fonction $(x_1, x_2) \mapsto \left((1+x_1)^2+x_2^2\right)^{3/2}$ qui ne s'annule pas sur U et est de classe $\mathcal{C}^\infty$ sur U. Les deux dérivées partielles de f en $(x_1, x_2) \in U$ sont

$$\partial_1 f(x_1,x_2) = \frac{2(1+x_1)^2 - x_2^2}{((1+x_1)^2+x_2^2)^{5/2}} \quad \text{et} \quad \partial_2 f(x_1,x_2) = \frac{3(1+x_1)x_2}{((1+x_1)^2+x_2^2)^{5/2}}.$$

Pour évaluer les dérivées partielles secondes, on dérive les dérivées partielles premières. L'application f étant de classe $\mathcal{C}^2$ sur U, d'après le théorème de Schwarz pour calculer la dérivée partielle seconde $\partial^2_{12} f$, on peut soit dériver l'application $\partial_1 f$ par rapport à x_2, soit[(55)] dériver l'application $\partial_2 f$ par rapport à x_1. On obtient

$$\partial^2_{11} f(x_1,x_2) = 3(1+x_1)\frac{3x_2^2 - 2(1+x_1)^2}{((1+x_1)^2+x_2^2)^{7/2}},$$

$$\partial^2_{21} f(x_1,x_2) = \partial^2_{12} f(x_1,x_2) = 3x_2\frac{x_2 - 4(1+x_1)^2}{((1+x_1)^2+x_2^2)^{7/2}},$$

$$\partial^2_{22} f(x_1,x_2) = 3(1+x_1)\frac{(1+x_1)^2 - 4x_2^2}{((1+x_1)^2+x_2^2)^{7/2}}.$$

On en déduit que la matrice hessienne de l'application f en $(0,0)$ est :

$$\mathrm{H}_f(0,0) = \begin{pmatrix} \partial^2_{11} f(0,0) & \partial^2_{12} f(0,0) \\ \partial^2_{21} f(0,0) & \partial^2_{22} f(0,0) \end{pmatrix} = \begin{pmatrix} -6 & 0 \\ 0 & 3 \end{pmatrix}.$$

10.3.5 Opérateurs différentiels

La notion de dérivée partielle intervient très fréquemment dans la formulation mathématique de nombreux phénomènes physiques. Ces formulations relient souvent les dérivées partielles d'une quantité physique par le biais d'« opérateurs différentiels ». Certains de ces opérateurs (le gradient, la divergence, le rotationnel et le laplacien) sont présentés dans cette partie.

(55) Utiliser les deux méthodes peut être un bon moyen de vérification du calcul effectué.

Définition 10.10 *Soit f une application d'un ouvert U de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}$, différentiable en $\mathbf{a} \in U$. On appelle* gradient *de f en $\mathbf{a}$, le vecteur de $\mathbb{R}^n$ défini par*

$$\nabla f(\mathbf{a}) = \begin{pmatrix} \dfrac{\partial f}{\partial x_1}(\mathbf{a}) \\ \vdots \\ \dfrac{\partial f}{\partial x_n}(\mathbf{a}) \end{pmatrix}.$$

Remarques

1. On rencontre parfois la notation $\operatorname{grad} f(\mathbf{a})$ au lieu de $\nabla f(\mathbf{a})$.

2. Le gradient de f en $\mathbf{a}$ n'est rien d'autre que la transposée de la matrice jacobienne de f en $\mathbf{a}$. D'ailleurs, on utilise parfois la notation $\nabla f(\mathbf{a})$ pour des applications à valeurs dans $\mathbb{R}^p$ avec $p \geqslant 2$. Dans ce cas, la définition de ce « gradient » est donnée par $\nabla f(\mathbf{a}) = {}^t\mathrm{J}_f(\mathbf{a})$.

3. Si l'application f est différentiable sur l'ouvert U alors on définit l'application $\nabla f : \mathbf{a} \in U \mapsto \nabla f(\mathbf{a}) \in \mathbb{R}^n$ appelée le *gradient* de f que l'on note

$$\nabla f = \begin{pmatrix} \dfrac{\partial f}{\partial x_1} \\ \vdots \\ \dfrac{\partial f}{\partial x_n} \end{pmatrix}.$$

4. Le symbole ∇, qui se lit *nabla*, tire son nom d'une lyre antique dont la forme est celle d'un triangle pointant vers le bas. Il contient toutes les informations sur la différentielle. On peut voir cet opérateur comme un opérateur formel, vecteur des dérivées partielles :

$$\nabla = \begin{pmatrix} \dfrac{\partial}{\partial x_1} \\ \vdots \\ \dfrac{\partial}{\partial x_n} \end{pmatrix}.$$

Attention, ce « vecteur » n'a rien d'un vecteur classique puisque ces composantes ne sont pas dans un espace que l'on peut décrire dans le cadre de ce cours. Il ne contient pas de valeurs scalaires, mais on va utiliser ses éléments constitutifs (que l'on peut voir comme des opérateurs en attente d'argument) exactement comme on aurait utilisé les valeurs scalaires composant un vecteur. ◇

Exemple Considérons l'application

$$f : (x_1, x_2, x_3) \in \mathbb{R}^3_* \longmapsto \frac{1}{\sqrt{x_1^2 + x_2^2 + x_3^2}} \in \mathbb{R}.$$

Si on note $\mathbf{r} = (x_1, x_2, x_3)$ le vecteur « position » et $\|\mathbf{r}\| = \sqrt{x_1^2 + x_2^2 + x_3^2}$ sa norme euclidienne canonique alors $f(x_1, x_2, x_3) = 1/\|\mathbf{r}\|$ correspond à l'inverse de la distance du point (x_1, x_2, x_3) à l'origine. Pour calculer le gradient de cette application, on calcule les dérivées partielles selon les trois vecteurs de la base canonique de $\mathbb{R}^3$. Pour $(x_1, x_2, x_3) \in \mathbb{R}^3_*$, on a

$$\begin{aligned}
\frac{\partial f}{\partial x_1}(x_1, x_2, x_3) &= -\frac{x_1}{(x_1^2 + x_2^2 + x_3^2)^{3/2}} = -\frac{x_1}{\|\mathbf{r}\|^3} \\
\frac{\partial f}{\partial x_2}(x_1, x_2, x_3) &= -\frac{x_2}{(x_1^2 + x_2^2 + x_3^2)^{3/2}} = -\frac{x_2}{\|\mathbf{r}\|^3} \\
\frac{\partial f}{\partial x_3}(x_1, x_2, x_3) &= -\frac{x_3}{(x_1^2 + x_2^2 + x_3^2)^{3/2}} = -\frac{x_3}{\|\mathbf{r}\|^3}
\end{aligned}$$

et on obtient $\nabla f(x_1, x_2, x_3) = -\mathbf{r}/\|\mathbf{r}\|^3$.

DÉFINITION 10.11 *Soit f une application d'un ouvert U de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^n$ différentiable en $\mathbf{a} \in U$. On appelle* divergence *de f en $\mathbf{a}$, le réel*

$$\operatorname{div} f(\mathbf{a}) = \frac{\partial f_1}{\partial x_1}(\mathbf{a}) + \frac{\partial f_2}{\partial x_2}(\mathbf{a}) + \cdots + \frac{\partial f_n}{\partial x_n}(\mathbf{a}),$$

où les applications f_1, f_2,..., f_n désignent les n applications coordonnées de f.

Remarques

1. La divergence est un opérateur différentiel qui agit uniquement sur les fonctions de $\mathbb{R}^n$ dans $\mathbb{R}^n$. Autrement dit, il faut que les espaces vectoriels de départ et d'arrivée aient la même dimension.

2. Si l'application f est différentiable sur l'ouvert U alors on définit l'application $\operatorname{div} f : \mathbf{a} \in U \mapsto \operatorname{div} f(\mathbf{a}) \in \mathbb{R}$ appelée la *divergence* de f.

3. La notation ∇ introduite pour noter le gradient d'une application fournit un moyen commode pour exprimer les opérateurs différentiels, et en particulier la divergence. Formellement, on a $\operatorname{div} f = \nabla \cdot f$: on fait comme un « produit scalaire » entre le « vecteur » ∇ et le « vecteur » dont les composantes sont les applications f_1, f_2,..., f_n.

4. La divergence d'une application f d'un ouvert U de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^n$ en $\mathbf{a} \in U$ peut aussi être définie comme la trace de la matrice jacobienne de f ; on a $\operatorname{div} f(\mathbf{a}) = \operatorname{Tr}(\mathrm{J}_f(\mathbf{a}))$. ◇

Exemple Considérons l'application

$$f : (x_1, x_2, x_3) \in \mathbb{R}^3_* \longmapsto \frac{-1}{(x_1^2 + x_2^2 + x_3^2)^{3/2}} \ (x_1, x_2, x_3) \in \mathbb{R}^3.$$

Notons $\mathbf{r} = (x_1, x_2, x_3)$ le vecteur « position » ; on a donc $f(x_1, x_2, x_3) = -\mathbf{r}/\|\mathbf{r}\|^3$. Pour déterminer la divergence de f, on calcule d'abord la dérivée partielle selon le i–ième vecteur de la base canonique de $\mathbb{R}^3$ de la i–ième composante de f, pour $i \in \{1, 2, 3\}$. La première dérivée partielle de f_1 vaut

$$\frac{\partial f_1}{\partial x_1}(x_1, x_2, x_3) = -\frac{1}{(x_1^2 + x_2^2 + x_3^2)^{3/2}} + \frac{3x_1^2}{(x_1^2 + x_2^2 + x_3^2)^{5/2}} = -\frac{1}{\|\mathbf{r}\|^3} + \frac{3x_1^2}{\|\mathbf{r}\|^5}.$$

De même, les variables x_1, x_2 et x_3 ayant un rôle symétrique, on obtient

$$\frac{\partial f_2}{\partial x_2}(x_1, x_2, x_3) = -\frac{1}{\|\mathbf{r}\|^3} + \frac{3x_2^2}{\|\mathbf{r}\|^5} \quad \text{et} \quad \frac{\partial f_3}{\partial x_3}(x_1, x_2, x_3) = -\frac{1}{\|\mathbf{r}\|^3} + \frac{3x_3^2}{\|\mathbf{r}\|^5}.$$

En additionnant ces trois dérivées partielles, on en déduit que la divergence de f est donnée par

$$\operatorname{div} f(x_1, x_2, x_3) = -\frac{3}{\|\mathbf{r}\|^3} + \frac{3\|\mathbf{r}\|^2}{\|\mathbf{r}\|^5} = 0.$$

Le troisième opérateur différentiel que nous allons introduire est le rotationnel. On sera attentif au fait qu'il n'est défini que pour des fonctions de $\mathbb{R}^3$ dans $\mathbb{R}^3$.

DÉFINITION 10.12 *Soit f une application d'un ouvert U de $\mathbb{R}^3$ à valeurs dans $\mathbb{R}^3$, différentiable en $\mathbf{a} \in U$. On appelle* rotationnel *de f en $\mathbf{a}$ le vecteur de $\mathbb{R}^3$ défini par*

$$\operatorname{rot} f(\mathbf{a}) = \begin{pmatrix} \dfrac{\partial f_3}{\partial x_2}(\mathbf{a}) - \dfrac{\partial f_2}{\partial x_3}(\mathbf{a}) \\ \dfrac{\partial f_1}{\partial x_3}(\mathbf{a}) - \dfrac{\partial f_3}{\partial x_1}(\mathbf{a}) \\ \dfrac{\partial f_2}{\partial x_1}(\mathbf{a}) - \dfrac{\partial f_1}{\partial x_2}(\mathbf{a}) \end{pmatrix},$$

où les applications f_1, f_2 et f_3 désignent les trois fonctions coordonnées de f.

Remarques

1. Le rotationnel défini ici est un opérateur différentiel qui agit uniquement sur des fonctions de $\mathbb{R}^3$ dans $\mathbb{R}^3$. Il existe néanmoins une définition de rotationnel pour les applications d'un ouvert U de $\mathbb{R}^2$ dans $\mathbb{R}^2$ différentiables qui est défini en $\mathbf{a} \in U$ par

$$\operatorname{rot} f(\mathbf{a}) = \frac{\partial f_2}{\partial x_1}(\mathbf{a}) - \frac{\partial f_1}{\partial x_2}(\mathbf{a}) \in \mathbb{R}.$$

Il n'existe pas de moyen simple de généraliser la notion de rotationnel à des fonctions de n variables pour $n \geqslant 4$.

2. Si l'application f est différentiable sur un ouvert U de $\mathbb{R}^3$ et à valeurs dans $\mathbb{R}^3$ alors on définit l'application $\operatorname{rot} f : \mathbf{a} \in U \mapsto \operatorname{rot} f(\mathbf{a}) \in \mathbb{R}^3$ appelée le *rotationnel* de f.

3. Comme pour la divergence, la notation ∇ introduite pour noter la gradient d'une application fournit un moyen commode pour exprimer le rotationnel. Formellement, on a $\operatorname{rot} f = \nabla \wedge f$: on fait comme un « produit vectoriel » entre le « vecteur » ∇ et le « vecteur » dont les composantes sont les applications f_1, f_2 et f_3. ◇

Exemple Considérons à nouveau l'application

$$f : (x_1, x_2, x_3) \in \mathbb{R}^3_* \longmapsto -\frac{1}{(x_1^2 + x_2^2 + x_3^2)^{3/2}} (x_1, x_2, x_3) \in \mathbb{R}^3.$$

La seconde dérivée partielle d'ordre 1 de la première application coordonnée en $(x_1, x_2, x_3) \in \mathbb{R}^3_*$ est donnée par

$$\frac{\partial f_1}{\partial x_2}(x_1, x_2, x_3) = \frac{3x_1x_2}{(x_1^2 + x_2^2 + x_3^2)^{5/2}}.$$

Par raison de symétrie, on obtient pour tout $(i, j) \in \{1, 2, 3\}^2$ avec $i \neq j$

$$\frac{\partial f_i}{\partial x_j}(x_1, x_2, x_3) = \frac{3x_ix_j}{(x_1^2 + x_2^2 + x_3^2)^{5/2}}.$$

On en déduit que $\operatorname{rot} f(x_1, x_2, x_3) = \mathbf{0}$.

Un moyen de se faire une idée de ce que représente le gradient, la divergence et le rotationnel est d'imaginer un banc de poissons qui se laisse porter par les courants marins. Quand un poisson se trouve en un point (x_1, x_2, x_3) de l'océan, il se déplace à la vitesse $\mathbf{v}(x_1, x_2, x_3)$ du courant à cet endroit. Si nous suivons la trajectoire d'un poisson en particulier, nous verrons sa vitesse changer au cours du temps : chaque fois qu'il se trouve en un endroit, il adopte comme vitesse la valeur du courant à cet endroit. Ce faisant, il se déplace légèrement, et doit alors prendre pour vitesse la valeur de $\mathbf{v}$ au nouveau point où il se trouve. Si nous observons maintenant le banc d'une façon plus globale, nous constatons que la densité ρ des poissons, c'est-à-dire leur nombre par unité de volume, n'est pas constante.

– Le gradient de ρ en un point de l'océan est un vecteur dirigé vers l'endroit où il y a le plus de poissons à côté de ce point. La longueur de ce vecteur est d'autant plus grande que le nombre de poissons dans cette direction varie rapidement.

– Les poissons ont tendance à se concentrer à certains endroits, et en revanche à déserter d'autres endroits. La divergence de $\mathbf{v}$ est la quantité qui mesure cette tendance : les poissons se concentrent aux endroits où la divergence de $\mathbf{v}$ est négative et désertent les endroits où elle est positive. Là où la divergence de $\mathbf{v}$ est nulle, la densité des poissons reste constante.

– Comme son nom l'indique, le rotationnel de $\mathbf{v}$ a quelque chose à voir avec une rotation. Il exprime dans quelle mesure, lorsqu'on regarde un petit secteur du banc de poissons, celui-ci tourne au cours du temps. Le rotationnel de $\mathbf{v}$ en un point est un vecteur dirigé selon l'axe de la rotation que subit le banc de poissons en ce point au cours de son mouvement, et il est de longueur proportionnelle à la vitesse de cette rotation.

Terminons cette partie consacrée aux opérateurs différentiels par le laplacien qui contrairement au gradient, à la divergence et au rotationnel est un opérateur d'ordre 2, c'est-à-dire un opérateur qui fait intervenir les dérivées partielles secondes.

DÉFINITION 10.13 *Soit f une application d'un ouvert U de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^p$ deux fois différentiable en $\mathbf{a} \in U$. On appelle* laplacien *de f en $\mathbf{a}$, le vecteur de $\mathbb{R}^p$ défini par*

$$\Delta f(\mathbf{a}) = \frac{\partial^2 f}{\partial x_1^2}(\mathbf{a}) + \frac{\partial^2 f}{\partial x_2^2}(\mathbf{a}) + \cdots + \frac{\partial^2 f}{\partial x_n^2}(\mathbf{a}).$$

Remarques

1. Si l'application f est deux fois différentiable sur l'ouvert U de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^p$ alors on définit l'application $\Delta f : \mathbf{a} \in U \mapsto \Delta f(\mathbf{a}) \in \mathbb{R}^p$ appelée le *laplacien* de f.

2. L'opérateur laplacien Δ peut s'exprimer de façon formelle à l'aide de la notation ∇ introduite pour noter la gradient. On écrit $\Delta = \nabla \cdot \nabla$: on fait comme un « produit scalaire » entre le « vecteur » ∇ et lui-même.

3. Le laplacien d'une application f d'un ouvert U de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}$ en $\mathbf{a} \in U$ peut aussi être défini comme la trace de la matrice hessienne de f en $\mathbf{a}$; on a $\Delta f(\mathbf{a}) = \mathrm{Tr}(\mathrm{H}_f(\mathbf{a}))$. ◇

Exemple Considérons l'application

$$f : (x_1, x_2, x_3) \in \mathbb{R}^3_* \longmapsto \frac{1}{\sqrt{x_1^2 + x_2^2 + x_3^2}} \in \mathbb{R}.$$

Pour obtenir le laplacien de f, il faut calculer ses dérivées partielles secondes selon les trois vecteurs de la base canonique de $\mathbb{R}^3$. On a[(56)]

$$\frac{\partial^2 f}{\partial x_1^2}(x_1, x_2, x_3) = -\frac{1}{(x_1^2 + x_2^2 + x_3^2)^{3/2}} + \frac{3x_1^2}{(x_1^2 + x_2^2 + x_3^2)^{5/2}} = -\frac{1}{\|\mathbf{r}\|^3} + \frac{3x_1^2}{\|\mathbf{r}\|^5}.$$

Compte tenu des symétries dans l'expression de $f(x_1, x_2, x_3)$, on obtient

$$\frac{\partial^2 f_2}{\partial x_2^2}(x_1, x_2, x_3) = -\frac{1}{\|\mathbf{r}\|^3} + \frac{3x_2^2}{\|\mathbf{r}\|^5} \quad \text{et} \quad \frac{\partial^2 f_3}{\partial x_3^2}(x_1, x_2, x_3) = -\frac{1}{\|\mathbf{r}\|^3} + \frac{3x_3^2}{\|\mathbf{r}\|^5}.$$

(56) Les dérivées partielles premières de cette application ont été calculées p. 656.

En additionnant ces trois dérivées partielles secondes on en déduit que pour tout $(x_1, x_2, x_3) \in \mathbb{R}^3_*$

$$\Delta f(x_1, x_2, x_3) = -\frac{3}{\|\mathbf{r}\|^3} + \frac{3\|\mathbf{r}\|^2}{\|\mathbf{r}\|^5} = 0.$$

En observant les exemples détaillés dans les pages précédentes, on se rend compte qu'il existe des liens entre les différents opérateurs introduits auparavant. Ces liens apparents sont en fait le résultat de relations générales existant entre ces différents opérateurs. Nous allons en donner les principales.

PROPOSITION 10.15 *Soient f une application d'un ouvert U de $\mathbb{R}^3$ à valeurs dans $\mathbb{R}$ et $\mathbf{F}$ une application de U dans $\mathbb{R}^3$. Si f et $\mathbf{F}$ sont de classe C^2 sur U alors*

✗ $\text{rot}(\nabla f) = \mathbf{0}$,

✗ $\text{div}(\text{rot } \mathbf{F}) = 0$.

Démonstration ⊵ L'application $G = \nabla f$ étant une application de U dans $\mathbb{R}^3$ de classe C^1 son rotationnel est bien défini sur U. Désignons par g_1, g_2 et g_3 les trois applications coordonnées de G. Le rotationnel de G vaut

$$\text{rot } G = \begin{pmatrix} \dfrac{\partial g_3}{\partial x_2} - \dfrac{\partial g_2}{\partial x_3} \\ \dfrac{\partial g_1}{\partial x_3} - \dfrac{\partial g_3}{\partial x_1} \\ \dfrac{\partial g_2}{\partial x_1} - \dfrac{\partial g_1}{\partial x_2} \end{pmatrix} = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_2 \partial x_3} - \dfrac{\partial^2 f}{\partial x_3 \partial x_2} \\ \dfrac{\partial^2 f}{\partial x_3 \partial x_1} - \dfrac{\partial^2 f}{\partial x_1 \partial x_3} \\ \dfrac{\partial^2 f}{\partial x_1 \partial x_2} - \dfrac{\partial^2 f}{\partial x_2 \partial x_1} \end{pmatrix}.$$

Puisque l'application f est de classe C^2 sur U, le théorème de Schwarz indique que

$$\forall (i,j) \in \{1,2,3\}^2 \qquad \frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_j}.$$

On en conclut que rot G est l'application nulle.

⊵ Puisque $\mathbf{F}$ est une application de U dans $\mathbb{R}^3$ de classe C^2, rot $\mathbf{F}$ est une application de U dans $\mathbb{R}^3$ de classe C^1 et sa divergence est bien définie sur U. Notons F_1, F_2 et F_3 les trois applications coordonnées de $\mathbf{F}$ et notons $G = \text{rot } \mathbf{F}$. Les trois applications coordonnées de G sont

$$g_1 = \frac{\partial F_3}{\partial x_2} - \frac{\partial F_2}{\partial x_3}, \quad g_2 = \frac{\partial F_1}{\partial x_3} - \frac{\partial F_3}{\partial x_1} \quad \text{et} \quad g_3 = \frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2}.$$

La divergence de g vaut

$$\begin{aligned} \text{div } g &= \frac{\partial g_1}{\partial x_1} + \frac{\partial g_2}{\partial x_2} + \frac{\partial g_3}{\partial x_3} \\ &= \frac{\partial^2 F_3}{\partial x_1 \partial x_2} - \frac{\partial^2 F_2}{\partial x_1 \partial x_3} + \frac{\partial^2 F_1}{\partial x_2 \partial x_3} - \frac{\partial^2 F_3}{\partial x_2 \partial x_1} + \frac{\partial^2 F_2}{\partial x_3 \partial x_1} - \frac{\partial^2 F_1}{\partial x_3 \partial x_2}. \end{aligned}$$

Puisque l'application $\mathbf{F}$ est une application de classe $\mathcal{C}^2$, en utilisant le théorème de Schwarz, on obtient div g est l'application nulle. □

Remarque La proposition 10.15 est énoncée uniquement pour des applications f et $\mathbf{F}$ dont le domaine de définition est un ouvert de $\mathbb{R}^3$. Si le domaine de définition est un ouvert de $\mathbb{R}^2$ alors la proposition est toujours vraie en prenant cette fois-ci la définition adaptée pour le rotationnel d'une fonction de $\mathbb{R}^2$ dans $\mathbb{R}^2$, voir p. 657. ◇

PROPOSITION 10.16

✗ *Soit f une application d'un ouvert U de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}$. Si f est deux fois différentiable sur U alors*

$$\Delta f = \operatorname{div}(\nabla f).$$

✗ *Soit $\mathbf{F}$ une application d'un ouvert V de $\mathbb{R}^3$ à valeurs dans $\mathbb{R}^3$. Si $\mathbf{F}$ est deux fois différentiable sur V alors*

$$\operatorname{rot}(\operatorname{rot}\mathbf{F}) = \nabla(\operatorname{div}\ \mathbf{F}) - \Delta\mathbf{F}.$$

Démonstration ⊵ Puisque f est une application de U dans $\mathbb{R}$ deux fois différentiable, $g = \nabla f$ est une application de $U \subset \mathbb{R}^n$ dans $\mathbb{R}^n$ un fois différentiable et sa divergence est bien définie. Les trois applications coordonnées de l'application g sont $g_1 = \frac{\partial f}{\partial x_1}$, $g_2 = \frac{\partial f}{\partial x_2}$ et $g_3 = \frac{\partial f}{\partial x_3}$ et la divergence de g vaut

$$\operatorname{div}\ g = \frac{\partial^2 f}{\partial x_1^2} + \frac{\partial^2 f}{\partial x_2^2} + \frac{\partial^2 f}{\partial x_3^2} = \Delta f.$$

⊵ Puisque $\mathbf{F}$ est une application de U dans $\mathbb{R}^3$ deux fois différentiable, div $\mathbf{F}$ est une application de U dans $\mathbb{R}$ une fois différentiable dont le gradient est bien défini et rot $\mathbf{F}$ est une application de U dans $\mathbb{R}^3$ une fois différentiable dont le rotationnel est bien défini. En calculant les composantes de rot(rot $\mathbf{F}$) puis celles de ∇(div $\mathbf{F}$), on obtient $\nabla(\operatorname{div}\ \mathbf{F}) - \operatorname{rot}(\operatorname{rot}\mathbf{F}) = \Delta\mathbf{F}$. □

Remarque La seconde égalité, faisant intervenir l'opérateur rotationnel, est aussi vraie pour une application $\mathbf{F}$ définie sur un ouvert de $\mathbb{R}^2$, à valeurs dans $\mathbb{R}^2$ et qui est deux fois différentiable, à condition d'utiliser la définition adaptée pour le rotationnel d'une fonction de $\mathbb{R}^2$ dans $\mathbb{R}^2$. ◇

PROPOSITION 10.17 *Soient f, g deux applications d'un ouvert U de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}$ et $\mathbf{F}$ une application de U dans $\mathbb{R}^n$. Si f, g et $\mathbf{F}$ sont deux fois différentiable sur U alors :*

✗ $\operatorname{div}(f\mathbf{F}) = f \operatorname{div}\mathbf{F} + \mathbf{F} \cdot \nabla f$,

✗ $\Delta(fg) = f\Delta g + 2\nabla f \cdot \nabla g + g\Delta f$.

Démonstration Pour établir ces relations, nous allons avoir recours aux propriétés suivantes. Si g_1 et g_2 sont deux applications différentiables de U dans $\mathbb{R}$ alors leur produit $g_1 g_2$ est aussi une application différentiable sur U et pour tout $k \in \{1, \dots, n\}$, on a

$$\partial_k(g_1 g_2) = g_1\, \partial_k g_2 + g_2\, \partial_k g_1.$$

De plus, si g_1 et g_2 sont deux fois différentiable sur U, le produit $g_1 g_2$ est deux fois différentiable sur U et pour tout $k \in \{1, \dots, n\}$, on a

$$\partial^2_{kk}(g_1 g_2) = g_1\, \partial^2_{kk} g_2 + 2\, \partial_k g_1\, \partial_k g_2 + g_2\, \partial^2_{kk} g_1.$$

$\trianglerighteq$ Notons F_1, F_2,..., F_n les applications coordonnées de $\mathbf{F}$. L'application $f\mathbf{F}$ est une applications de U dans $\mathbb{R}^n$ dont les applications coordonnées sont fF_1, fF_2,..., fF_n. La divergence de $f\mathbf{F}$ vaut donc

$$\operatorname{div}(f\mathbf{F}) = \sum_{k=1}^{n} \partial_k(fF_k) = \sum_{k=1}^{n} (f\, \partial_k F_k + F_k\, \partial_k f) = f \sum_{k=1}^{n} \partial_k F_k + \sum_{k=1}^{n} F_k\, \partial_k f.$$

La première somme dans la dernière égalité est égale à $\operatorname{div} \mathbf{F}$. Quant à la seconde somme, elle s'interprète comme le « produit scalaire » de $\mathbf{F}$ par ∇f. On a donc

$$\operatorname{div}(f\mathbf{F}) = f \operatorname{div} \mathbf{F} + \mathbf{F} \cdot \nabla f.$$

$\trianglerighteq$ Intéressons-nous à la seconde relation. On a

$$\begin{aligned} \Delta(fg) &= \sum_{k=1}^{n} \partial^2_{kk}(fg) = \sum_{k=1}^{n} \left(f\, \partial^2_{kk} g + 2\, \partial_k f\, \partial_k g + g\, \partial^2_{kk} f\right) \\ &= f \sum_{k=1}^{n} \partial^2_{kk} g + 2 \sum_{k=1}^{n} \partial_k f\, \partial_k g + g \sum_{k=1}^{n} \partial^2_{kk} f. \end{aligned}$$

La première somme dans la dernière égalité est égale à Δg et la dernière somme vaut Δf. Quant à la seconde du milieu, elle s'interprète comme le « produit scalaire » de ∇f et de ∇g. On a donc

$$\Delta(fg) = f\, \Delta g + 2\, \nabla f \cdot \nabla g + g\, \Delta f.$$

La seconde relation de la proposition est démontrée. □

Les opérateurs introduits dans cette partie permettent d'écrire les équations régissant de nombreux phénomènes physiques dont voici quelques exemples.

1. La température θ dans un corps dépend de la position $\mathbf{x} \in \mathbb{R}^3$ et évolue au cours du temps $t \in \mathbb{R}^+$ selon l' « équation de la chaleur » :

$$\frac{\partial \theta}{\partial t}(t, \mathbf{x}) - \kappa \Delta \theta(t, \mathbf{x}) = 0,$$

où la constante $\kappa \in \mathbb{R}^*_+$ est liée à la capacité calorifique du corps.

2. Le champ magnétique $\mathbf{H}$ dans un milieu de perméabilité magnétique $\mu \in \mathbb{R}_+^*$ vérifie les « équations de la magnétostatique » :

$$\operatorname{rot} \mathbf{H}(\mathbf{x}) = \mathbf{0} \quad \text{et} \quad \operatorname{div}(\mu \mathbf{H})(\mathbf{x}) = 0.$$

3. Le vecteur vitesse $\mathbf{U}$ d'un fluide et la pression p au sein de ce fluide sont régis au cours du temps par les « équations de Stokes » :

$$\frac{\partial \mathbf{U}}{\partial t}(t, \mathbf{x}) + \nabla p(t, \mathbf{x}) - \eta \Delta \mathbf{U}(t, \mathbf{x}) = 0 \quad \text{et} \quad \operatorname{div}\ \mathbf{U}(\mathbf{x}) = 0,$$

où $\eta \in \mathbb{R}_+^*$ désigne la viscosité du fluide.

Considérons par exemple une pièce rectangulaire, assimilée au sous-ensemble de $\mathbb{R}^2$ défini par $P = \{(x_1, x_2) \in \mathbb{R}^2 \ ; \ 0 \leqslant x_1 \leqslant 2 \quad \text{et} \quad 0 \leqslant x_2 \leqslant 1\}$ dont le mur $\{(x_1, x_2) \in \mathbb{R}^2 \ ; \ x_1 = 0\}$ est constamment chauffé à 100°C, le mur $\{(x_1, x_2) \in \mathbb{R}^2 \ ; \ x_1 = 2\}$ est refroidi à 0°C alors que les autres murs, représentés par les ensembles $\{(x_1, x_2) \in \mathbb{R}^2 \ ; \ x_2 = 0\}$ et $\{(x_1, x_2) \in \mathbb{R}^2 \ ; \ x_2 = 1\}$, sont supposés isolés thermiquement. La température θ de la pièce, que l'on suppose indépendante du temps, satisfait l'équation de la chaleur stationnaire[57] et les conditions imposées au bord de la pièce. Elle est solution du « problème aux dérivées partielle » :

$$\begin{cases} \Delta\theta(x_1, x_2) = 0 & & \text{pour tout } (x_1, x_2) \in P, \\ \theta(0, x_2) = 100, & \theta(2, x_2) = 0 & \text{pour tout } 0 \leqslant x_2 \leqslant 1, \\ \dfrac{\partial \theta}{\partial x_2}(x_1, 0) = 0, & \dfrac{\partial \theta}{\partial x_2}(x_1, 1) = 0 & \text{pour tout } 0 \leqslant x_1 \leqslant 2. \end{cases}$$

Il arrive fréquemment que contrairement à l'exemple précédent, le domaine physique ne soit pas « adapté » aux coordonnées cartésiennes. Par exemple lorsque le problème de la chaleur est posé dans une boule, ou lorsqu'un fluide satisfaisant aux équations de Stokes s'écoule dans un cylindre. L'outil permettant de « changer de domaine » est présenté dans la partie suivante.

10.3.6 Théorèmes d'inversion locale et globale

Définition 10.14 *Soient U et V deux ouverts de $\mathbb{R}^n$. On dit qu'une application f de U dans V est un $\mathcal{C}^1$-difféomorphisme si f est bijective et si les applications f et f^{-1} sont de classe $\mathcal{C}^1$.*

Exemple Considérons l'application f définie de $\mathbb{R}^2$ dans $\mathbb{R}^2$ par

$$f : (x_1, x_2) \in \mathbb{R}^2 \longmapsto (x_1 \cos(x_2), x_1 \sin(x_2)) \in \mathbb{R}^2.$$

[57] On dit qu'un phénomène est stationnaire lorsqu'on fait l'hypothèse qu'il ne dépend pas du temps.

Cette application est une application de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$ (chacune des applications coordonnées étant le produit de deux applications de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$). Le calcul des dérivées partielles de f a été effectué à la p. 630. Remarquons que pour tout $(x_1, x_2) \in \mathbb{R}^2$, on a $f(x_1, x_2) = f(x_1, x_2 + 2\pi)$ ce qui prouve que l'application f n'est pas injective, ce n'est donc pas un $\mathcal{C}^1$-difféomorphisme. Par contre, si on restreint l'application f à $\mathbb{R}_+^* \times [0, 2\pi[$ alors on obtient une application injective. Toutefois cet ensemble n'étant pas un ouvert, on est amené à considérer son intérieur $U = \mathbb{R}_+^* \times]0, 2\pi[$. La fonction f est donc une bijection de U sur $V = f(U)$. On vérifie que $V = \mathbb{R}^2 \setminus \{(y_1, 0) \in \mathbb{R}^2 \ ; \ y_1 \geqslant 0\}$. Pour montrer que $f : U \to V$ est un $\mathcal{C}^1$-difféomorphisme en utilisant la définition 10.14, on doit déterminer l'application réciproque f^{-1}. Pour $(x_1, x_2) \in U$ et $(y_1, y_2) \in V$ avec $y_1 \neq 0$, on a

$$\begin{aligned}
(y_1, y_2) = f(x_1, x_2) &\iff y_1 = x_1 \cos(x_2) \quad \text{et} \quad y_2 = x_1 \sin(x_2) \\
&\iff y_1^2 + y_2^2 = x_1^2 \quad \text{et} \quad \frac{y_2}{y_1} = \tan(x_2) \\
&\iff x_1 = (y_1^2 + y_2^2)^{\frac{1}{2}} \quad \text{et} \quad x_2 = \arctan(y_2/y_1) .
\end{aligned}$$

Si $y_1 = 0$ alors on obtient

$$(y_1, y_2) = f(x_1, x_2) \iff \begin{cases} x_1 = y_2 \quad \text{et} \quad x_2 = \pi/2 \quad \text{si} \quad y_2 > 0, \\ x_1 = -y_2 \quad \text{et} \quad x_2 = -\pi/2 \quad \text{si} \quad y_2 > 0. \end{cases}$$

On en déduit que l'application réciproque de f est donnée par

$$f^{-1} : (y_1, y_2) \in V \longmapsto \begin{cases} \left((y_1^2 + y_2^2)^{\frac{1}{2}} \ , \ \arctan(y_2/y_1)\right) & \text{si } y_1 \neq 0 \\ (y_2, \pi/2) & \text{si } y_1 = 0 \text{ et } y_2 > 0 \\ (-y_2, -\pi/2) & \text{si } y_1 = 0 \text{ et } y_2 < 0 \end{cases} .$$

On peut vérifier que cette application f^{-1} est une application de classe $\mathcal{C}^1$ sur V ce qui prouve que l'application f est un $\mathcal{C}^1$-difféomorphisme de U sur V.

À l'instar des $\mathcal{C}^1$-difféomorphismes, on peut définir les *$\mathcal{C}^q$-difféomorphismes* : ce sont les applications bijectives de classe $\mathcal{C}^q$ ($q \in \mathbb{N}^* \cup \{\infty\}$) dont la réciproque est aussi une application de classe $\mathcal{C}^q$.

Définition 10.15 *Soient U et V deux ouverts de $\mathbb{R}^n$ et $q \in \mathbb{N}^* \cup \{\infty\}$. On dit qu'une application f de U sur V est un $\mathcal{C}^q$-difféomorphisme si f est bijective et si les applications f et f^{-1} sont de classe $\mathcal{C}^q$.*

Remarque Les applications continues (c'est-à-dire de classe $\mathcal{C}^0$) bijectives dont la réciproque est aussi continue sont appelées des *homéomorphismes.* L'étude des homéomorphismes est différente de celle que l'on va présenter ici pour les $\mathcal{C}^q$-difféomorphisme où $q \in \mathbb{N}^* \cup \{\infty\}$. Cette différence résulte principalement de la caractérisation qui sera donnée pour les $\mathcal{C}^q$-difféomorphismes au théorème 10.4. ◇

D'un point de vue pratique, montrer qu'une application f est un $\mathcal{C}^q$-difféomorphisme à l'aide de la définition n'est en général pas simple. En effet, elle nécessite d'avoir une expression de la réciproque de f. Celle-ci n'est pas toujours facile à expliciter. Le théorème suivant, que nous admettrons, permet de savoir si une application est un $\mathcal{C}^q$-difféomorphisme sans avoir besoin d'exhiber sa réciproque.

THÉORÈME 10.4 (Théorème d'inversion locale)
Soient U un ouvert de $\mathbb{R}^n$, f une application de U dans $\mathbb{R}^n$ de classe $\mathcal{C}^q$, $q \in \mathbb{N}^* \cup \{\infty\}$, et $\mathbf{a} \in U$. Si le jacobien de f ne s'annule pas en $\mathbf{a}$ alors il existe un voisinage $U_\mathbf{a}$ de $\mathbf{a}$ dans U tel que l'application f soit un $\mathcal{C}^q$-difféomorphisme de $U_\mathbf{a}$ sur son image $f(U_\mathbf{a})$.

Exemple Considérons à nouveau l'application f définie de $\mathbb{R}^2$ dans $\mathbb{R}^2$ par

$$f : (x_1, x_2) \in \mathbb{R}^2 \longmapsto (x_1 \cos(x_2), x_1 \sin(x_2)) \in \mathbb{R}^2$$

et le point $\mathbf{a} = (1,1) \in \mathbb{R}^2$. L'application f est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}^2$ et le jacobien[58] de f en $\mathbf{a} = (1,1)$ vaut $\mathcal{J}_f(1,1) = 1 \neq 0$. Le jacobien étant non nul, on peut appliquer le théorème d'inversion locale qui assure l'existence d'un voisinage $U_\mathbf{a}$ de $\mathbf{a}$ dans $\mathbb{R}^2$ sur lequel f est un $\mathcal{C}^\infty$-difféomorphisme.

Le théorème d'inversion local permet de déterminer si une application est un difféomorphisme au voisinage d'un point donné. Le théorème suivant, que nous admettons également, permet avec une hypothèse supplémentaire de déterminer le voisinage « maximal » sur lequel l'application est un difféomorphisme.

THÉORÈME 10.5 (Théorème d'inversion globale)
Soient U un ouvert de $\mathbb{R}^n$, $q \in \mathbb{N}^* \cup \{\infty\}$ et f une application de U dans $\mathbb{R}^n$. L'application $f : U \to f(U)$ est un $\mathcal{C}^q$-difféomorphisme si et seulement si

1. *l'application f est de classe $\mathcal{C}^q$ sur U,*
2. *l'application f est injective sur U,*
3. *le jacobien de f ne s'annule pas sur U.*

Exemple On peut ainsi compléter l'exemple précédent : on a montré, grâce au théorème d'inversion locale, que l'application

$$f : (x_1, x_2) \in \mathbb{R}^2 \longmapsto (x_1 \cos(x_2), x_1 \sin(x_2)) \in \mathbb{R}^2$$

[58] Voir le calcul du jacobien de cette application à la p. 636.

est un $\mathcal{C}^\infty$-difféomorphisme d'un voisinage $U_{\mathbf{a}}$ de $\mathbf{a} = (1,1)$ sur $f(U_{\mathbf{a}})$. En remarquant[59] que l'application f est injective sur l'ouvert $U = \mathbb{R}_+^* \times]0, 2\pi[$ et que sur cet ouvert, le jacobien de f ne s'annule pas[60], on établit à l'aide du théorème d'inversion globale que f est un $\mathcal{C}^\infty$-difféomorphisme de U sur son image $f(U)$ de façon plus rapide qu'en ayant recours à la définition 10.15 comme dans l'exemple de la p. 663.

D'un point de vue pratique, pour montrer qu'une application f définie sur un ouvert U de $\mathbb{R}^n$ et à valeurs dans un ouvert V de $\mathbb{R}^n$ est un $\mathcal{C}^q$-difféomorphisme, on vérifie que :

- l'application f est de classe $\mathcal{C}^q$ sur U,
- le jacobien de f ne s'annule pas sur U,
- l'application f est injective sur U,
- l'image de U par f est l'ouvert V.

Exercice 10.11 Considérons l'application de $\mathbb{R}^3$ dans $\mathbb{R}^3$ définie par

$$f : (x_1, x_2, x_3) \in \mathbb{R}^3 \mapsto (x_1 \cos(x_2)\sin(x_3), x_1 \sin(x_2)\sin(x_3), x_1 \cos(x_3)) \in \mathbb{R}^3.$$

1 - Montrer que f n'est pas un C^1-difféomorphisme de $\mathbb{R}^3$ sur $\mathbb{R}^3$.
2 - Montrer[61] que f est un C^1-difféomorphisme de $\mathbb{R}_+^* \times]0, 2\pi[\times \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$ sur son image.

10.3.7 Changements de coordonnées

De nombreux problèmes physiques sont posés dans des domaines qui ne sont pas « adaptés » aux coordonnées cartésiennes. On utilise alors des « changements de coordonnées » pour exprimer mathématiquement ces problèmes dans des systèmes mieux adaptés où leur résolution s'en trouvera facilitée. Les changements de coordonnées les plus courants sont les changements en coordonnées polaires, en coordonnées cylindriques et en coordonnées sphériques.

Changement de coordonnées en polaire

La fonction de changement de coordonnées

$$\Phi : (r, \theta) \longmapsto (r\cos(\theta), r\sin(\theta))$$

est un $\mathcal{C}^\infty$-difféomorphisme de U sur V où

$$U = \mathbb{R}_+^* \times]0, 2\pi[\qquad \text{et} \qquad V = \mathbb{R}^2 \setminus \{(x, 0) \in \mathbb{R}^2 \ ; \ x \geqslant 0\}.$$

---

[59] Voir à ce sujet l'exemple de la p. 663.

[60] Le jacobien de f en (x_1, x_2) a été calculé à la p. 636 ; il vaut $\mathcal{J}_f(x_1, x_2) = x_1$.

[61] On pourra consulter les exercices 10.4 et 10.7 au cours desquels ont été calculées les dérivées partielles et le jacobien de l'application f.

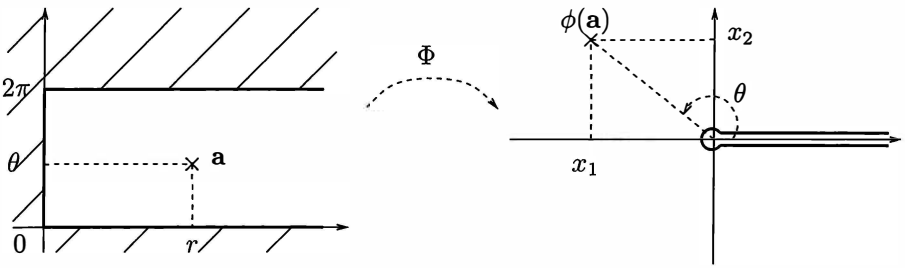


Fig. 2 Illustration du changement de coordonnées en polaire

Considérons un ouvert $V' \subset V$ et une application f de V' dans $\mathbb{R}$ de classe $\mathcal{C}^q$ sur V' avec $q \in \mathbb{N}^* \cup \{\infty\}$. Pour tout $(x,y) \in V'$, il existe un unique couple (r,θ) dans $U' = \Phi^{-1}(V')$ tel que $x = r\cos(\theta)$ et $y = r\sin(\theta)$ et

$$f(x,y) = f(r\cos(\theta), r\sin(\theta)) = g(r,\theta)$$

où $g = f \circ \Phi$. L'application g est de classe $\mathcal{C}^q$ de U' sur $f(V')$.

Remarque La fonction Φ est un $\mathcal{C}^\infty$-difféomorphisme de U sur V. On peut aussi montrer que c'est un $\mathcal{C}^\infty$-difféomorphisme de $U' = \mathbb{R}_+^* \times]-\pi,\pi[$ sur $V' = \mathbb{R}^2 \setminus \{(x,0) \in \mathbb{R}^2 \ ; \ x \leqslant 0\}$. Ainsi, si l'application f à laquelle on veut appliquer le changement de coordonnées n'est pas définie sur un ouvert de V, mais sur un ouvert de V' alors on pourra encore utiliser le même changement de coordonnées. ◇

Le principe du changement de coordonnées est de traduire toutes les informations connues sur l'application g en informations sur l'application f et inversement. Ainsi, les applications Φ et Φ^{-1} étant de classe $\mathcal{C}^\infty$, d'après la proposition 10.13, pour tout $q \in \mathbb{N} \cup \{\infty\}$, on a

$$g \in \mathcal{C}^q(U,\mathbb{R}) \iff f \in \mathcal{C}^q(V,\mathbb{R}).$$

On peut exprimer les dérivées partielles de l'application g en fonction des dérivées partielles de l'application f et inversement. Ainsi, les dérivées partielles de g s'expriment en fonction des dérivées partielles de f de la manière suivante :

$$\frac{\partial g}{\partial r}(r,\theta) = \cos(\theta)\frac{\partial f}{\partial x}(r\cos(\theta), r\sin(\theta)) + \sin(\theta)\frac{\partial f}{\partial y}(r\cos(\theta), r\sin(\theta))$$

et

$$\frac{\partial g}{\partial \theta}(r,\theta) = -r\sin(\theta)\frac{\partial f}{\partial x}(r\cos(\theta), r\sin(\theta)) + r\cos(\theta)\frac{\partial f}{\partial y}(r\cos(\theta), r\sin(\theta)).$$

On en déduit l'expression des dérivées partielles de f en fonction des dérivées partielles de g :

$$\frac{\partial f}{\partial x}(r\cos(\theta), r\sin(\theta)) = \cos(\theta)\frac{\partial g}{\partial r}(r,\theta) - \frac{\sin(\theta)}{r}\frac{\partial g}{\partial \theta}(r,\theta)$$

et
$$\frac{\partial f}{\partial y}(r\cos(\theta), r\sin(\theta)) = \sin(\theta)\frac{\partial g}{\partial r}(r,\theta) + \frac{\cos(\theta)}{r}\frac{\partial g}{\partial \theta}(r,\theta).$$

Par conséquent, le gradient de f en fonction des dérivées partielles de g a pour expression dans la base canonique $(\mathbf{e}_1, \mathbf{e}_2)$ de $\mathbb{R}^2$:

$$\nabla f(x,y) = \frac{\partial g}{\partial r}(r,\theta)(\cos(\theta)\,\mathbf{e}_1 + \sin(\theta)\,\mathbf{e}_2) + \frac{1}{r}\frac{\partial g}{\partial \theta}(r,\theta)(-\sin(\theta)\,\mathbf{e}_1 + \cos(\theta)\,\mathbf{e}_2).$$

Notons $\mathbf{e}_r = \cos(\theta)\,\mathbf{e}_1 + \sin(\theta)\,\mathbf{e}_2$ et $\mathbf{e}_\theta = -\sin(\theta)\,\mathbf{e}_1 + \cos(\theta)\,\mathbf{e}_2$. On vérifie que $(\mathbf{e}_r, \mathbf{e}_\theta)$ est une base de $\mathbb{R}^2$ et que cette base est orthonormée pour le produit scalaire usuel. Elle est appelée *base des coordonnées polaires*. Le gradient de f dans cette base s'exprime sous la forme

$$\nabla f(x,y) = \frac{\partial g}{\partial r}(r,\theta)\,\mathbf{e}_r + \frac{1}{r}\frac{\partial g}{\partial \theta}(r,\theta)\,\mathbf{e}_\theta.$$

Autrement dit, dans la base des coordonnées polaires, l'opérateur gradient s'exprime sous la forme

$$\boxed{\nabla f(x,y) = \begin{pmatrix} \dfrac{\partial g}{\partial r}(r,\theta) \\ \dfrac{1}{r}\dfrac{\partial g}{\partial \theta}(r,\theta) \end{pmatrix}.}$$

On peut, sur le même principe, exprimer l'opérateur divergence en coordonnées polaires. Considérons une application $\mathbf{F}$ d'un ouvert V' de V dans $\mathbb{R}^2$ de classe C^q, $q \in \mathbb{N}^* \cup \{\infty\}$. Pour tout $(x,y) \in V'$ il existe un unique couple (r,θ) dans $U' = \Phi^{-1}(V')$ tel que $x = r\cos(\theta)$, $y = r\sin(\theta)$ et on a

$$\mathbf{F}(x,y) = \big(f_1(r\cos(\theta), r\sin(\theta)), f_2(r\cos(\theta), r\sin(\theta))\big) = \big(g_1(r,\theta), g_2(r,\theta)\big)$$

où $g_1 = f_1 \circ \Phi$ et $g_2 = f_2 \circ \Phi$ sont les applications coordonnées de $\mathbf{G} = \mathbf{F} \circ \Phi$ dans la base canonique $(\mathbf{e}_1, \mathbf{e}_2)$. L'application $\mathbf{G}$ est de classe C^q de U' sur $\Phi(V')$ et la divergence de $\mathbf{F}$ s'exprime à l'aide des dérivées partielles premières de $\mathbf{G}$ de la façon suivante.

$$\begin{aligned} \operatorname{div}\,\mathbf{F}(x,y) &= \frac{\partial f_1}{\partial x}(x,y) + \frac{\partial f_2}{\partial y}(x,y) \\ &= \cos(\theta)\,\frac{\partial g_1}{\partial r}(r,\theta) - \frac{\sin(\theta)}{r}\,\frac{\partial g_1}{\partial \theta}(r,\theta) + \sin(\theta)\,\frac{\partial g_2}{\partial r}(r,\theta) + \frac{\cos(\theta)}{r}\,\frac{\partial g_2}{\partial \theta}(r,\theta). \end{aligned}$$

Cette relation constitue l'expression de la divergence en coordonnées polaires dans la base canonique $(\mathbf{e}_1, \mathbf{e}_2)$ de $\mathbb{R}^2$. Notons g_r et g_θ les applications coordonnées de $\mathbf{G}$ dans la base des coordonnées polaires $(\mathbf{e}_r, \mathbf{e}_\theta)$. On a

$$\begin{aligned} \mathbf{G}(r,\theta) &= g_r(r,\theta)\;\mathbf{e}_r + g_\theta(r,\theta)\;\mathbf{e}_\theta \\ &= g_r(r,\theta)\;(\cos(\theta)\,\mathbf{e}_1 + \sin(\theta)\,\mathbf{e}_2) + g_\theta(r,\theta)\;(-\sin(\theta)\,\mathbf{e}_1 + \cos(\theta)\,\mathbf{e}_2) \\ &= (g_r(r,\theta)\,\cos(\theta) - g_\theta(r,\theta)\,\sin(\theta))\,\mathbf{e}_1 + (g_r(r,\theta)\,\sin(\theta) + g_\theta(r,\theta)\,\cos(\theta))\,\mathbf{e}_2. \end{aligned}$$

On en déduit que $g_1(r,\theta) = g_r\cos(\theta) - g_\theta\sin(\theta)$ et $g_2(r,\theta) = g_r\sin(\theta) + g_\theta\cos(\theta)$. On a donc

$$\frac{\partial g_1}{\partial r}(r,\theta) = \cos(\theta)\,\frac{\partial g_r}{\partial r}(r,\theta) - \sin(\theta)\,\frac{\partial g_\theta}{\partial r}(r,\theta),$$
$$\frac{\partial g_1}{\partial \theta}(r,\theta) = -\sin(\theta)\,g_r(r,\theta) + \cos(\theta)\,\frac{\partial g_r}{\partial \theta}(r,\theta) + \cos(\theta)\,g_\theta(r,\theta) - \sin(\theta)\,\frac{\partial g_\theta}{\partial \theta}(r,\theta),$$
$$\frac{\partial g_2}{\partial r}(r,\theta) = \sin(\theta)\,\frac{\partial g_r}{\partial r}(r,\theta) + \cos(\theta)\,\frac{\partial g_\theta}{\partial r}(r,\theta),$$
$$\frac{\partial g_2}{\partial \theta}(r,\theta) = \cos(\theta)\,g_r(r,\theta) + \sin(\theta)\,\frac{\partial g_r}{\partial \theta}(r,\theta) - \sin(\theta)\,g_\theta(r,\theta) + \cos(\theta)\,\frac{\partial g_\theta}{\partial \theta}(r,\theta).$$

On obtient donc en coordonnées polaires l'expression suivante de la divergence dans la base des coordonnées polaires :

$$\operatorname{div}\ \mathbf{F}(x,y) = \frac{1}{r}\frac{\partial g_r}{\partial r}(r,\theta) + \frac{\partial g_r}{\partial r}(r,\theta) + \frac{1}{r}\frac{\partial g_\theta}{\partial \theta}(r,\theta)$$

$$\boxed{\operatorname{div}\ \mathbf{F}(x,y) = \frac{1}{r}\frac{\partial (r g_r)}{\partial r}(r,\theta) + \frac{1}{r}\frac{\partial g_\theta}{\partial \theta}(r,\theta).}$$

On peut déterminer l'expression du laplacien en coordonnées polaires dans la base des coordonnées polaires : pour une application f d'un ouvert V' de V dans $\mathbb{R}$ de classe $\mathcal{C}^2$ on a

$$\Delta f(x,y) = \operatorname{div}\,(\nabla f(x,y)) = \operatorname{div}\,(\nabla g(r,\theta)) = \operatorname{div}\begin{pmatrix} \dfrac{\partial g}{\partial r}(r,\theta) \\ \dfrac{1}{r}\dfrac{\partial g}{\partial \theta}(r,\theta) \end{pmatrix}$$

$$\boxed{\Delta f(x,y) = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial g}{\partial r}(r,\theta)\right) + \frac{1}{r^2}\frac{\partial^2 g}{\partial \theta^2}(r,\theta).}$$

Changement de coordonnées en cylindrique

La fonction de changement de coordonnées

$$\Phi : (r,\theta,z) \longmapsto (r\cos(\theta), r\sin(\theta), z)$$

est un $\mathcal{C}^\infty$-difféomorphisme de l'ensemble $U = \mathbb{R}_+^* \times]0, 2\pi[\times \mathbb{R}$ sur l'ensemble $V = V_1 \times \mathbb{R}$ où $V_1 = \mathbb{R}^2 \setminus \{(x,0) \in \mathbb{R}^2\ ;\ x \geqslant 0\}$.

Considérons une application f définie sur un ouvert V' de V, à valeurs dans $\mathbb{R}$ et de classe $\mathcal{C}^q$ sur V' avec $q \in \mathbb{N}^* \cup \{\infty\}$. Pour tout $(x,y,z) \in V'$, il existe un unique couple de réels (r,θ) tel que $(r,\theta,z) \in U' = \Phi^{-1}(V')$, $x = r\cos(\theta)$, $y = r\sin(\theta)$ et

$$f(x,y,z) = f(r\cos(\theta), r\sin(\theta), z) = g(r,\theta,z)$$

où $g = f \circ \Phi$. L'application g est de classe $\mathcal{C}^q$ de U' sur $f(V')$.

Le principe du changement de coordonnées est de traduire toutes les informations connues sur l'application g en informations sur l'application f et inversement. Ainsi, les applications Φ et Φ^{-1} étant de classe C^∞, d'après la proposition 10.13, pour tout $q \in \mathbb{N} \cup \{\infty\}$, on a

$$g \in C^q(U, \mathbb{R}) \iff f \in C^q(V, \mathbb{R}).$$

On peut exprimer les dérivées partielles de l'application g en fonction des dérivées partielles de l'application f et inversement. Ainsi, les dérivées partielles de g s'expriment en fonction des dérivées partielles de f de la manière suivante :

$$\begin{aligned}
\frac{\partial g}{\partial r}(r,\theta,z) &= \cos(\theta)\,\frac{\partial f}{\partial x}(r\cos(\theta), r\sin(\theta), z) + \sin(\theta)\,\frac{\partial f}{\partial y}(r\cos(\theta), r\sin(\theta), z),\\
\frac{\partial g}{\partial \theta}(r,\theta,z) &= -r\sin(\theta)\,\frac{\partial f}{\partial x}(r\cos(\theta), r\sin(\theta), z) + r\cos(\theta)\,\frac{\partial f}{\partial y}(r\cos(\theta), r\sin(\theta), z),\\
\frac{\partial g}{\partial z}(r,\theta,z) &= \frac{\partial f}{\partial z}(r\cos(\theta), r\sin(\theta), z).
\end{aligned}$$

On en déduit l'expression des dérivées partielles de f en fonction des dérivées partielles de g :

$$\begin{aligned}
\frac{\partial f}{\partial x}(r\cos(\theta), r\sin(\theta), z) &= \cos(\theta)\,\frac{\partial g}{\partial r}(r,\theta,z) - \frac{1}{r}\sin(\theta)\,\frac{\partial g}{\partial \theta}(r,\theta,z),\\
\frac{\partial f}{\partial y}(r\cos(\theta), r\sin(\theta), z) &= \sin(\theta)\,\frac{\partial g}{\partial r}(r,\theta,z) + \frac{1}{r}\cos(\theta)\,\frac{\partial g}{\partial \theta}(r,\theta,z),\\
\frac{\partial f}{\partial z}(r\cos(\theta), r\sin(\theta), z) &= \frac{\partial g}{\partial z}(r,\theta,z).
\end{aligned}$$

Ainsi, on exprime le gradient de f en fonction des dérivées partielles de g dans la base canonique $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ de $\mathbb{R}^3$:

$$\begin{aligned}
\nabla f(x,y,z) = &\frac{\partial g}{\partial r}(r,\theta,z)\ (\cos(\theta)\,\mathbf{e}_1 + \sin(\theta)\,\mathbf{e}_2)\\
&+ \frac{1}{r}\frac{\partial g}{\partial \theta}(r,\theta,z)\ (-\sin(\theta)\,\mathbf{e}_1 + \cos(\theta)\,\mathbf{e}_2) + \frac{\partial g}{\partial z}(r,\theta,z)\ \mathbf{e}_3.
\end{aligned}$$

Notons $\mathbf{e}_r = \cos(\theta)\,\mathbf{e}_1 + \sin(\theta)\,\mathbf{e}_2$, $\mathbf{e}_\theta = -\sin(\theta)\,\mathbf{e}_1 + \cos(\theta)\,\mathbf{e}_2$ et $\mathbf{e}_z = \mathbf{e}_3$. On vérifie que $(\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_z)$ est une base de $\mathbb{R}^3$ et que cette base est orthonormée. Elle est appelée *base des coordonnées cylindriques*. Le gradient de f dans cette base s'exprime sous la forme

$$\nabla f(x,y,z) = \frac{\partial g}{\partial r}(r,\theta,z)\ \mathbf{e}_r + \frac{1}{r}\frac{\partial g}{\partial \theta}(r,\theta,z)\ \mathbf{e}_\theta + \frac{\partial g}{\partial z}(r,\theta,z)\ \mathbf{e}_z.$$

Autrement dit, dans la base des coordonnées cylindriques, l'opérateur gradient s'exprime sous la forme

$$\boxed{\nabla f(x,y,z) = \begin{pmatrix} \dfrac{\partial g}{\partial r}(r,\theta,z) \\ \dfrac{1}{r}\dfrac{\partial g}{\partial \theta}(r,\theta,z) \\ \dfrac{\partial g}{\partial z}(r,\theta,z) \end{pmatrix}.}$$

On peut, sur le même principe, exprimer l'opérateur divergence en coordonnées cylindriques. Considérons une application $\mathbf{F}$ d'un ouvert V' de V dans $\mathbb{R}^3$ de classe $\mathcal{C}^q$, $q \in \mathbb{N}^* \cup \{\infty\}$. Pour tout $(x, y, z) \in V'$ il existe un unique couple de réels (r, θ) tel que $(r, \theta, z) \in U' = \Phi^{-1}(V')$, $x = r\cos(\theta)$, $y = r\sin(\theta)$ et

$$\mathbf{F}(x, y, z) = (g_1(r, \theta, z), g_2(r, \theta, z), g_3(r, \theta, z))$$

où $g_1 = f_1 \circ \Phi$, $g_2 = f_2 \circ \Phi$ et $g_3 = f_3 \circ \Phi$ sont les applications coordonnées de $\mathbf{G} = \mathbf{F} \circ \Phi$ dans la base canonique $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$. L'application $\mathbf{G}$ est de classe $\mathcal{C}^q$ de U' sur $\Phi(V')$ et la divergence de $\mathbf{F}$ s'exprime à l'aide des dérivées partielles premières de $\mathbf{G}$ de la façon suivante.

$$\begin{aligned} \operatorname{div} \mathbf{F}(x, y, z) &= \frac{\partial f_1}{\partial x}(x, y, z) + \frac{\partial f_2}{\partial y}(x, y, z) + \frac{\partial f_3}{\partial z}(x, y, z) \\ &= \cos(\theta)\frac{\partial g_1}{\partial r}(r, \theta, z) - \frac{\sin(\theta)}{r}\frac{\partial g_1}{\partial \theta}(r, \theta, z) \\ &\quad + \sin(\theta)\frac{\partial g_2}{\partial r}(r, \theta, z) + \frac{\cos(\theta)}{r}\frac{\partial g_2}{\partial \theta}(r, \theta, z) + \frac{\partial g_3}{\partial z}(r, \theta, z). \end{aligned} \tag{18}$$

Cette expression constitue l'expression de la divergence en coordonnées cylindriques dans la base canonique $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ de $\mathbb{R}^3$. Notons g_r, g_θ et g_z les applications coordonnées de $\mathbf{G}$ dans la base des coordonnées cylindriques $(\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_z)$. On a

$$\begin{aligned} \mathbf{G}(r, \theta, z) &= g_r(r, \theta, z)\ \mathbf{e}_r + g_\theta(r, \theta, z)\ \mathbf{e}_\theta + g_z(r, \theta, z)\ \mathbf{e}_z \\ &= g_r(r, \theta, z)\ (\cos(\theta)\,\mathbf{e}_1 + \sin(\theta)\,\mathbf{e}_2) \\ &\qquad + g_\theta(r, \theta, z)\ (-\sin(\theta)\,\mathbf{e}_1 + \cos(\theta)\,\mathbf{e}_2) + g_z(r, \theta, z)\ \mathbf{e}_3 \\ &= (g_r(r, \theta, z)\cos(\theta) - g_\theta(r, \theta, z)\sin(\theta))\ \mathbf{e}_1 \\ &\qquad + (g_r(r, \theta, z)\sin(\theta) + g_\theta(r, \theta, z)\cos(\theta))\ \mathbf{e}_2 + g_z(r, \theta, z)\ \mathbf{e}_3. \end{aligned}$$

On en déduit que $g_1(r, \theta, z) = g_r(r, \theta, z)\cos(\theta) - g_\theta(r, \theta, z)\sin(\theta)$, que $g_2(r, \theta, z) = g_r(r, \theta, z)\sin(\theta) + g_\theta(r, \theta, z)\cos(\theta)$ et que $g_3(r, \theta, z) = g_z(r, \theta, z)$. On a donc

$$\begin{aligned} \frac{\partial g_1}{\partial r}(r, \theta, z) &= \cos(\theta)\,\frac{\partial g_r}{\partial r}(r, \theta, z) - \sin(\theta)\,\frac{\partial g_\theta}{\partial r}(r, \theta, z), \\ \frac{\partial g_1}{\partial \theta}(r, \theta, z) &= -\sin(\theta)\,g_r(r, \theta, z) + \cos(\theta)\,\frac{\partial g_r}{\partial \theta}(r, \theta, z) + \cos(\theta)\,g_\theta(r, \theta, z) \\ &\qquad - \sin(\theta)\,\frac{\partial g_\theta}{\partial \theta}(r, \theta, z), \\ \frac{\partial g_2}{\partial r}(r, \theta, z) &= \sin(\theta)\,\frac{\partial g_r}{\partial r}(r, \theta, z) + \cos(\theta)\,\frac{\partial g_\theta}{\partial r}(r, \theta, z), \\ \frac{\partial g_2}{\partial \theta}(r, \theta, z) &= \cos(\theta)\,g_r(r, \theta, z) + \sin(\theta)\,\frac{\partial g_r}{\partial \theta}(r, \theta, z) - \sin(\theta)\,g_\theta(r, \theta, z) \\ &\qquad + \cos(\theta)\,\frac{\partial g_\theta}{\partial \theta}(r, \theta, z), \end{aligned}$$

et

$$\frac{\partial g_1}{\partial z}(r, \theta, z) = 0, \quad \frac{\partial g_2}{\partial z}(r, \theta, z) = 0.$$

De plus,

$$\frac{\partial g_3}{\partial r}(r,\theta,z)=0,\quad \frac{\partial g_3}{\partial \theta}(r,\theta,z)=0\quad\text{et}\quad \frac{\partial g_3}{\partial z}(r,\theta,z)=\frac{\partial g_z}{\partial z}(r,\theta,z).$$

De l'équation (18), on déduit l'expression en coordonnées cylindriques de la divergence dans la base des coordonnées cylindriques :

$$\text{div}\ \mathbf{F}(x,y,z)=\frac{1}{r}\frac{\partial g_r}{\partial r}(r,\theta,z)+\frac{\partial g_r}{\partial r}(r,\theta,z)+\frac{1}{r}\frac{\partial g_\theta}{\partial \theta}(r,\theta,z)+\frac{\partial g_z}{\partial z}(r,\theta,z),$$

ce que l'on peut encore écrire :

$$\boxed{\text{div}\ \mathbf{F}(x,y,z)=\frac{1}{r}\frac{\partial (rg_r)}{\partial r}(r,\theta,z)+\frac{1}{r}\frac{\partial g_\theta}{\partial \theta}(r,\theta,z)+\frac{\partial g_z}{\partial z}(r,\theta,z).}$$

On peut aussi déterminer l'expression du laplacien en coordonnées cylindriques dans la base des coordonnées cylindriques : pour une application f d'un ouvert V' de V dans $\mathbb{R}$ de classe $\mathcal{C}^2$, on a

$$\Delta f(x,y,z)=\text{div}\,(\nabla f(x,y,z))=\text{div}\,(\nabla g(r,\theta,z))=\text{div}\begin{pmatrix}\dfrac{\partial g}{\partial r}(r,\theta,z)\\ \dfrac{1}{r}\dfrac{\partial g}{\partial \theta}(r,\theta,z)\\ \dfrac{\partial g}{\partial z}(r,\theta,z).\end{pmatrix}.$$

On en déduit que

$$\boxed{\Delta f(x,y,z)=\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial g}{\partial r}(r,\theta,z)\right)+\frac{1}{r^2}\frac{\partial^2 g}{\partial \theta^2}(r,\theta,z)+\frac{\partial^2 g}{\partial z^2}(r,\theta,z).}$$

Changement de coordonnées en sphérique

La fonction de changement de coordonnées

$$\Phi:(r,\theta,\psi)\longmapsto(r\cos(\theta)\sin(\psi),r\sin(\theta)\sin(\psi),r\cos(\psi))$$

est un $\mathcal{C}^\infty$-difféomorphisme de l'ensemble $U=\mathbb{R}_+^*\times]0,2\pi[\times]-\pi/2,\pi/2[$ sur l'ensemble $V=\mathbb{R}^3\setminus\{(x,0,z)\in\mathbb{R}^3\ ;\ x\geqslant 0\}$.

Considérons une application f définie sur un ouvert V' inclus dans V, à valeurs dans $\mathbb{R}$ et de classe $\mathcal{C}^q$ sur V' avec $q\in\mathbb{N}^*\cup\{\infty\}$. Pour tout $(x,y,z)\in V'$ il existe un unique triplet $(r,\theta,\psi)\in U'=\Phi^{-1}(V')$ tel que $x=r\cos(\theta)\sin(\psi)$, $y=r\sin(\theta)\sin(\psi)$, $z=r\cos(\psi)$ et

$$f(x,y,z)=f(r\cos(\theta)\sin(\psi),r\sin(\theta)\sin(\psi),r\cos(\psi))=g(r,\theta,\psi)$$

où $g=f\circ\Phi$. L'application g est de classe $\mathcal{C}^q$ de U' sur $f(V')$.

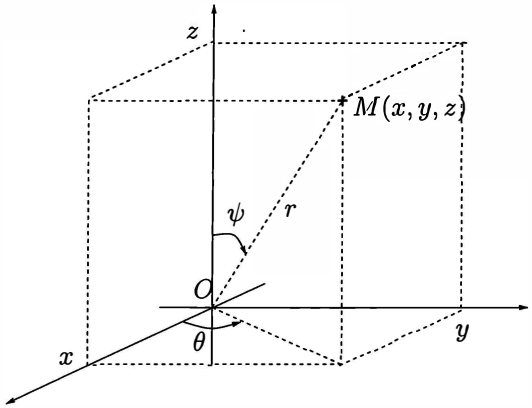


Fig. 3 Illustration du changement de coordonnées en sphérique.

Le principe du changement de coordonnées est de traduire toutes les informations connues sur l'application g en informations sur l'application f et inversement. Ainsi, les applications Φ et Φ^{-1} étant de classe $\mathcal{C}^\infty$, d'après la proposition 10.13, pour tout $q \in \mathbb{N} \cup \{\infty\}$, on a

$$g \in \mathcal{C}^q(U, \mathbb{R}) \iff f \in \mathcal{C}^q(V, \mathbb{R}).$$

On peut exprimer les dérivées partielles de l'application g en fonction des dérivées partielles de l'application f et inversement. Ainsi, les dérivées partielles de g s'expriment en fonction des dérivées partielles de f de la manière suivante :

$$\begin{aligned}
\frac{\partial g}{\partial r}(r,\theta,\psi) =& \cos(\theta)\sin(\psi)\,\frac{\partial f}{\partial x}(r\cos(\theta)\sin(\psi), r\sin(\theta)\sin(\psi), r\cos(\psi)) \\
&+ \sin(\theta)\sin(\psi)\,\frac{\partial f}{\partial y}(r\cos(\theta)\sin(\psi), r\sin(\theta)\sin(\psi), r\cos(\psi)) \\
&+ \cos(\psi)\,\frac{\partial f}{\partial z}(r\cos(\theta)\sin(\psi), r\sin(\theta)\sin(\psi), r\cos(\psi)), \\
\frac{\partial g}{\partial \theta}(r,\theta,\psi) =& -r\sin(\theta)\sin(\psi)\,\frac{\partial f}{\partial x}(r\cos(\theta)\sin(\psi), r\sin(\theta)\sin(\psi), r\cos(\psi)) \\
&+ r\cos(\theta)\sin(\psi)\,\frac{\partial f}{\partial y}(r\cos(\theta)\sin(\psi), r\sin(\theta)\sin(\psi), r\cos(\psi)), \\
\frac{\partial g}{\partial \psi}(r,\theta,\psi) =& r\cos(\theta)\cos(\psi)\,\frac{\partial f}{\partial x}(r\cos(\theta)\sin(\psi), r\sin(\theta)\sin(\psi), r\cos(\psi)) \\
&+ r\sin(\theta)\cos(\psi)\,\frac{\partial f}{\partial y}(r\cos(\theta)\sin(\psi), r\sin(\theta)\sin(\psi), r\cos(\psi)) \\
&- r\sin(\psi)\frac{\partial f}{\partial z}(r\cos(\theta)\sin(\psi), r\sin(\theta)\sin(\psi), r\cos(\psi)).
\end{aligned}$$

On en déduit l'expression suivante des dérivées partielles de f en fonction des dérivées partielles de g :

$$\frac{\partial f}{\partial x}(r\cos(\theta)\sin(\psi), r\sin(\theta)\sin(\psi), r\cos(\psi)) =$$
$$\cos(\theta)\sin(\psi)\,\frac{\partial g}{\partial r}(r,\theta,\psi) - \frac{\sin(\theta)}{r\sin(\psi)}\,\frac{\partial g}{\partial \theta}(r,\theta,\psi) + \frac{1}{r}\cos(\theta)\cos(\psi)\,\frac{\partial g}{\partial \psi}(r,\theta,\psi),$$
$$\frac{\partial f}{\partial y}(r\cos(\theta)\sin(\psi), r\sin(\theta)\sin(\psi), r\cos(\psi)) =$$
$$\sin(\theta)\sin(\psi)\,\frac{\partial g}{\partial r}(r,\theta,\psi) + \frac{\cos(\theta)}{r\sin(\psi)}\,\frac{\partial g}{\partial \theta}(r,\theta,\psi) + \frac{1}{r}\sin(\theta)\cos(\psi)\,\frac{\partial g}{\partial \psi}(r,\theta,\psi),$$
$$\frac{\partial f}{\partial z}(r\cos(\theta)\sin(\psi), r\sin(\theta)\sin(\psi), r\cos(\psi)) =$$
$$\cos(\psi)\,\frac{\partial g}{\partial r}(r,\theta,\psi) - \frac{1}{r}\sin(\psi)\,\frac{\partial g}{\partial \psi}(r,\theta,\psi).$$

On définit les vecteurs $\mathbf{e}_r$, $\mathbf{e}_\theta$ et $\mathbf{e}_\psi$ de la façon suivante :

$$\mathbf{e}_r = \cos(\theta)\sin(\psi)\ \mathbf{e}_1 + \sin(\theta)\sin(\psi)\ \mathbf{e}_2 + \cos(\psi)\ \mathbf{e}_3,$$
$$\mathbf{e}_\theta = -\sin(\theta)\ \mathbf{e}_1 + \cos(\theta)\ \mathbf{e}_2,$$
$$\mathbf{e}_\psi = \cos(\theta)\cos(\psi)\ \mathbf{e}_1 + \sin(\theta)\cos(\psi)\ \mathbf{e}_2 - \sin(\psi)\ \mathbf{e}_3.$$

On vérifie que $(\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_\psi)$ est une base de $\mathbb{R}^3$ et que cette base est orthonormée. Elle est appelée *base des coordonnées sphériques*. Le gradient de f dans cette base s'exprime sous la forme

$$\nabla f(x,y,z) = \frac{\partial g}{\partial r}(r,\theta,\psi)\ \mathbf{e}_r + \frac{1}{r\sin(\psi)}\,\frac{\partial g}{\partial \theta}(r,\theta,\psi)\ \mathbf{e}_\theta + \frac{1}{r}\,\frac{\partial g}{\partial \psi}(r,\theta,\psi)\ \mathbf{e}_\psi.$$

Autrement dit, dans la base des coordonnées sphériques, l'opérateur gradient s'exprime sous la forme

$$\boxed{\nabla f(x,y,z) = \begin{pmatrix} \dfrac{\partial g}{\partial r}(r,\theta,\psi) \\ \dfrac{1}{r\sin(\psi)}\,\dfrac{\partial g}{\partial \theta}(r,\theta,\psi) \\ \dfrac{1}{r}\,\dfrac{\partial g}{\partial \psi}(r,\theta,\psi) \end{pmatrix}.}$$

On peut, sur le même principe, exprimer l'opérateur divergence en coordonnées sphériques. Considérons une application $\mathbf{F}$ d'un ouvert V' de V dans $\mathbb{R}^3$ de classe C^q, $q \in \mathbb{N}^* \cup \{\infty\}$. Pour tout $(x,y,z) \in V'$, il existe un unique triplet $(r,\theta,\psi) \in U' = \Phi^{-1}(V')$ tel que $x = r\cos(\theta)\sin(\psi)$, $y = r\sin(\theta)\sin(\psi)$, $z = r\cos(\psi)$ et

$$\mathbf{F}(x,y,z) = (g_1(r,\theta,\psi), g_2(r,\theta,\psi), g_3(r,\theta,\psi))$$

où $g_1 = f_1 \circ \Phi$, $g_2 = f_2 \circ \Phi$ et $g_3 = f_3 \circ \Phi$ sont les applications coordonnées de $\mathbf{G} = \mathbf{F} \circ \Phi$ dans la base canonique $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$. L'application $\mathbf{G}$ est de classe C^q de U' sur $\Phi(V')$ et la divergence de $\mathbf{F}$ donnée par

$$\text{div } \mathbf{F}(x,y,z) = \frac{\partial f_1}{\partial x}(x,y,z) + \frac{\partial f_2}{\partial y}(x,y,z) + \frac{\partial f_3}{\partial z}(x,y,z)$$

s'exprime à l'aide des dérivées partielles premières de $\mathbf{G}$ de la façon suivante :

$$\begin{aligned}\text{div } \mathbf{F}(x,y,z) \;=\; & \cos(\theta)\sin(\psi)\,\frac{\partial g_1}{\partial r}(r,\theta,\psi) - \frac{\sin(\theta)}{r\sin(\psi)}\,\frac{\partial g_1}{\partial \theta}(r,\theta,\psi) \\ & + \frac{1}{r}\cos(\theta)\cos(\psi)\,\frac{\partial g_1}{\partial \psi}(r,\theta,\psi) + \sin(\theta)\sin(\psi)\,\frac{\partial g_2}{\partial r}(r,\theta,\psi) \qquad (19) \\ & + \frac{\cos(\theta)}{r\sin(\psi)}\,\frac{\partial g_2}{\partial \theta}(r,\theta,\psi) + \frac{1}{r}\sin(\theta)\cos(\psi)\,\frac{\partial g_2}{\partial \psi}(r,\theta,\psi) \\ & + \cos(\psi)\,\frac{\partial g_3}{\partial r}(r,\theta,\psi) - \frac{1}{r}\sin(\psi)\,\frac{\partial g_3}{\partial \psi}(r,\theta,\psi).\end{aligned}$$

Cette expression constitue l'expression de la divergence en coordonnées sphériques dans la base canonique $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ de $\mathbb{R}^3$. Notons g_r, g_θ et g_ψ les composantes de $\mathbf{G}$ dans la base des coordonnées sphériques $(\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_\psi)$. On a[62]

$$\begin{aligned}\mathbf{G} =& g_r\,\mathbf{e}_r + g_\theta\,\mathbf{e}_\theta + g_\psi\,\mathbf{e}_\psi \\ =& (g_r\cos(\theta)\sin(\psi) - g_\theta\sin(\theta) + g_\psi\cos(\theta)\cos(\psi))\,\mathbf{e}_1 \\ &+ (g_r\sin(\theta)\sin(\psi) + g_\theta\cos(\theta) + g_\psi\sin(\theta)\cos(\psi))\,\mathbf{e}_2 + (g_r\cos(\psi) - g_\psi\sin(\psi))\,\mathbf{e}_3.\end{aligned}$$

On en déduit que

$$\begin{aligned} g_1(r,\theta,\psi) &= g_r(r,\theta,\psi)\cos(\theta)\sin(\psi) - g_\theta(r,\theta,\psi)\sin(\theta) + g_\psi(r,\theta,\psi)\cos(\theta)\cos(\psi) \\ g_2(r,\theta,\psi) &= g_r(r,\theta,\psi)\sin(\theta)\sin(\psi) + g_\theta(r,\theta,\psi)\cos(\theta) + g_\psi(r,\theta,\psi)\sin(\theta)\cos(\psi) \\ g_3(r,\theta,\psi) &= g_r(r,\theta,\psi)\cos(\psi) - g_\psi(r,\theta,\psi)\sin(\psi)\end{aligned}$$

On peut exprimer les dérivées partielles des applications g_1, g_2 et g_3 et par conséquent obtenir à partir de l'équation (19) l'expression en coordonnées sphériques de la divergence dans la base des coordonnées sphériques :

$$\boxed{\begin{aligned}\text{div } \mathbf{F}(x,y,z) = \frac{1}{r^2}\frac{\partial(r^2 g_r)}{\partial r}(r,\theta,\psi) + \frac{1}{r\sin(\psi)}\frac{\partial g_\theta}{\partial \theta}(r,\theta,\psi) \\ + \frac{1}{r\sin(\psi)}\frac{\partial(\sin(\psi)g_\psi)}{\partial \psi}(r,\theta,\psi).\end{aligned}}$$

On peut déterminer l'expression du laplacien en coordonnées sphériques dans la base des coordonnées sphériques : pour une application f d'un ouvert V' de V dans $\mathbb{R}$ de classe C^2 on a

$$\Delta f(x,y,z) = \text{div}\,(\nabla f(x,y,\psi)) = \text{div}\,(\nabla g(r,\theta,\psi))\,.$$

[62] On omet exceptionnellement les variables (r,θ,ψ) pour $\mathbf{G}, g_r, g_\theta, g_\psi$ dans un soucis de lisibilité.

D'après les relations obtenues pour le gradient et la divergence en coordonnées sphériques, on en déduit que :

$$\Delta f(x,y,z) = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial g}{\partial r}(r,\theta,\psi)\right) + \frac{1}{r^2\sin^2(\psi)}\frac{\partial^2 g}{\partial\theta^2}(r,\theta,\psi) + \frac{1}{r\sin(\psi)}\frac{\partial}{\partial\psi}\left(\frac{\sin(\psi)}{r}\frac{\partial g}{\partial\psi}(r,\theta,\psi)\right).$$

Application : comportement de la pression dans un coin

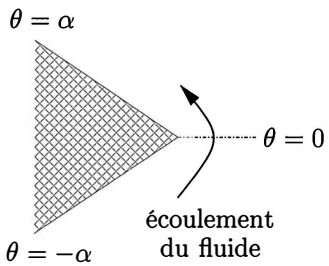
La pression p d'un fluide newtonien, comme par exemple de l'eau, dans un domaine $\Omega \subset \mathbb{R}^2$ est régie par l'équation

$$\Delta p = 0 \quad \text{dans } \Omega \tag{20}$$

à laquelle on adjoint des conditions au bord $\partial\Omega$ du domaine, par exemple[63]

$$\frac{\partial p}{\partial n} = 0 \quad \text{sur } \partial\Omega. \tag{21}$$

Pour résoudre ces équations dans le cas où le domaine Ω présente un coin caractérisé par un angle $\alpha \in]0,\pi[$ comme représenté sur la figure ci-contre, il peut être utile d'écrire les équations (20) et (21) en coordonnées polaires, en prenant comme origine le sommet du coin : $p(x,y) = q(r,\theta)$.



Les équations (20) et (21) sont équivalentes aux équations suivantes

$$\begin{cases} \dfrac{\partial}{\partial r}\left(r\dfrac{\partial q}{\partial r}(r,\theta)\right) + \dfrac{1}{r}\dfrac{\partial^2 q}{\partial\theta^2}(r,\theta) = 0 & \forall (r,\theta) \in]0,+\infty[\times]-\alpha,\alpha[\\ \dfrac{\partial q}{\partial\theta}(r,\theta) = 0 & \forall (r,\theta) \in]0,+\infty[\times\{-\alpha,\alpha\} \end{cases} \tag{22}$$

Afin de connaître le comportement de la pression au voisinage du coin, déterminons le plus petit réel $m > 0$ pour lequel il existe une solution de la forme

$$q(r,\theta) = r^m\varphi(\theta) \quad \text{avec } \varphi \in \mathcal{C}^\infty(]-\alpha,\alpha[,\mathbb{R}).$$

En utilisant les équations (22), on en déduit que φ doit vérifier

$$\begin{cases} \varphi''(\theta) + m^2\varphi(\theta) = 0 \quad \text{pour } \theta \in]-\alpha,\alpha[, \\ \varphi'(-\alpha) = 0 \quad \text{et} \quad \varphi'(\alpha) = 0. \end{cases} \tag{23}$$

[63] La notation $\frac{\partial}{\partial n}$ représente la dérivée normale, c'est la dérivée directionnelle selon le vecteur normal sortant à Ω.

Les solutions de cette équation différentielle linéaire du second ordre à coefficients constants[64] sont de la forme $\varphi(\theta) = A\sin(m\theta) + B\cos(m\theta)$ où A et B sont deux constantes réelles. On montre aisément qu'une telle application φ satisfaisant aux conditions $\varphi'(-\alpha) = 0$ et $\varphi'(\alpha) = 0$ est nécessairement de la forme

$$\varphi(\theta) = A\sin(m\theta), \quad m = \frac{k\pi}{\alpha} \quad \text{avec } A \in \mathbb{R} \text{ et } k \in \mathbb{Z},$$

ou bien de la forme

$$\varphi(\theta) = B\cos(m\theta), \quad m = \frac{\pi}{2\alpha} + \frac{k\pi}{\alpha} \quad \text{avec } B \in \mathbb{R} \text{ et } k \in \mathbb{Z}.$$

Ainsi, le plus petit réel $m > 0$ qui convient vaut

$$m = \frac{\pi}{2\alpha}.$$

En particulier, lorsque $\alpha = 3\pi/4$, le réel m correspondant vaut $m = 2/3$ alors que lorsque $\alpha = \pi/4$ on obtient $m = 2$. Les comportements obtenus pour la pression au voisinage du coin dans ces deux cas sont décrits sur la figure 4.

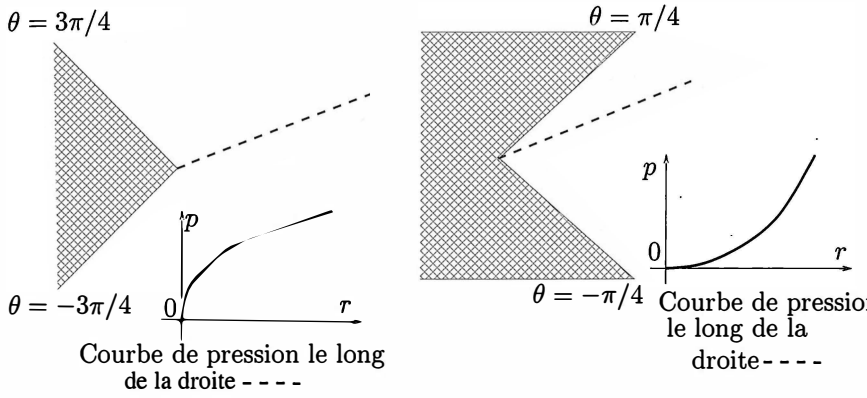


Fig. 4 Pression au voisinage d'un coin sortant (à gauche) et d'un coin rentrant (à droite).

10.4 Inégalité des accroissements finis ; formules de Taylor

10.4.1 Inégalité des accroissements finis

Pour les fonctions réelles de la variable réelle, le théorème des accroissements finis[65] exprime le fait que, pour tous réels a et b, si f est continue sur $[a, b]$ et dérivable sur $]a, b[$ alors on a une *égalité* de la forme

$$f(b) - f(a) = f'(c)(b - a) \tag{24}$$

[64] Voir le *Cours de première année*, chap. 20.
[65] Voir le *Cours de première année*, p. 731.

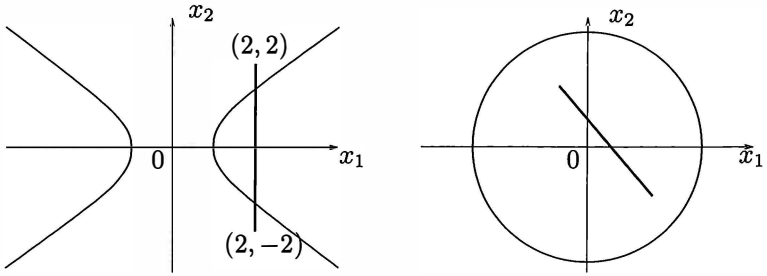


Fig. 5 Ensemble non convexe et ensemble convexe.

où c est un réel compris strictement entre a et b. Pour les fonctions de la variable complexe, un résultat analogue[66] a été démontré où l'égalité dans la relation (24) « devient » une inégalité. Pour des fonctions de $\mathbb{R}^n$ dans $\mathbb{R}^p$, le théorème des accroissements finis s'exprime également sous forme d'une inégalité. Nous allons, dans un premier temps, définir la notion de segment $[\mathbf{a}, \mathbf{b}]$ dans $\mathbb{R}^n$.

DÉFINITION 10.16 *Soient* $\mathbf{a}$ *et* $\mathbf{b}$ *sont deux éléments de l'espace vectoriel* $\mathbb{R}^n$. *On appelle* segment *d'extrémités* $\mathbf{a}$ *et* $\mathbf{b}$ *l'ensemble noté* $[\mathbf{a}, \mathbf{b}]$ *défini par*

$$[\mathbf{a}, \mathbf{b}] = \left\{\mathbf{c} \in \mathbb{R}^n \ ; \ \exists t \in [0,1] \quad \mathbf{c} = t\mathbf{a} + (1-t)\mathbf{b}\right\}.$$

Un sous-ensemble X *de* $\mathbb{R}^n$ *est dit* convexe *si*

$$\forall (\mathbf{a}, \mathbf{b}) \in X^2 \quad [\mathbf{a}, \mathbf{b}] \subset X.$$

Exemples

1. La boule unité de l'espace euclidien $\mathbb{R}^2$, c'est-à-dire l'ensemble défini par $\{(x_1, x_2) \in \mathbb{R}^2 \ ; \ x_1^2 + x_2^2 < 1\}$ est un sous-ensemble convexe de $\mathbb{R}^2$. Plus généralement, toute boule d'un espace vectoriel normé $(E, \|\cdot\|)$ est convexe. En effet, notons $B = \{\mathbf{x} \in E \ ; \ \|\mathbf{x}\| < 1\}$ la boule ouverte[67] de centre $\mathbf{0}_E$ et de rayon 1 de l'espace vectoriel $(E, \|\cdot\|)$. Si $\mathbf{a}$ et $\mathbf{b}$ sont deux éléments de B alors pour tout $t \in [0,1]$ on a

$$\|t\mathbf{a} + (1-t)\mathbf{b}\| \leqslant t\|\mathbf{a}\| + (1-t)\|\mathbf{b}\| < t + (1-t) = 1.$$

Ainsi, le segment $[\mathbf{a}, \mathbf{b}]$ est inclus dans la boule B. Ceci étant vrai pour tout couple $(\mathbf{a}, \mathbf{b}) \in B^2$, on en conclut que la boule B est convexe.

---

[66] Voir le théorème 4.2 p. 201.

[67] Un raisonnement semblable est possible avec une boule fermée.

2. L'ensemble $\{(x_1, x_2) \in \mathbb{R}^2 \ ; \ x_1^2 - x_2^2 < 1\}$ n'est pas un ensemble convexe. En effet, les deux vecteurs $\mathbf{a} = (2, 2)$ et $\mathbf{b} = (2, -2)$ appartiennent à cet ensemble alors que le milieu du segment $[\mathbf{a}, \mathbf{b}]$, c'est-à-dire le point $\mathbf{c} = t\mathbf{a} + (1-t)\mathbf{b}$ avec $t = 1/2$ n'appartient pas à cet ensemble.

THÉORÈME 10.6 (Inégalité des accroissements finis)
Soient f une application définie sur un ouvert U de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^p$ de classe $\mathcal{C}^1$ et K un sous-ensemble fermé et borné de U. Pour tout $(\mathbf{a}, \mathbf{b}) \in K^2$ on a

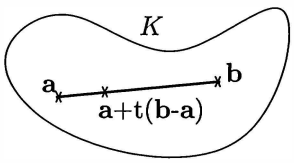
$$[\mathbf{a}, \mathbf{b}] \subset K \quad \Longrightarrow \quad \|f(\mathbf{b}) - f(\mathbf{a})\|_{\mathbb{R}^p} \leqslant M_K \|\mathbf{b} - \mathbf{a}\|_{\mathbb{R}^n},$$

où $M_K = \sup\limits_{\mathbf{x} \in K} |||\mathrm{d}f_{\mathbf{x}}|||_{\mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)}$.

Démonstration Tout d'abord, on peut noter que f est de classe $\mathcal{C}^1$ donc $\mathrm{d}f$ est continue ce qui justifie que $M_K = \sup_{\mathbf{x} \in K} |||\mathrm{d}f_{\mathbf{x}}|||_{\mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)}$ existe[68].

Soient $\mathbf{a}$ et $\mathbf{b}$ deux éléments de K tels que $[\mathbf{a}, \mathbf{b}] \subset K$ et soit φ l'application définie par[69]

$$\begin{aligned} \varphi \ : [0,1] &\longrightarrow \mathbb{R}^p \\ t &\longmapsto f(\mathbf{a} + t(\mathbf{b} - \mathbf{a})) \end{aligned}$$



L'application φ est de classe $\mathcal{C}^1$ (en tant que composée d'applications de classe $\mathcal{C}^1$, voir la proposition 10.6) et on a

$$\forall i \in \{1, \ldots, n\} \qquad \varphi_i(1) - \varphi_i(0) = \int_0^1 \varphi_i'(t) \, \mathrm{d}t.$$

Puisque $\varphi(1) = f(\mathbf{b})$, que $\varphi(0) = f(\mathbf{a})$ et que pour tout $t \in [0,1]$ on a $\varphi'(t) = \mathrm{d}f_{\mathbf{a}+t(\mathbf{b}-\mathbf{a})}(\mathbf{b} - \mathbf{a})$, on en déduit que

$$\begin{aligned} \|f(\mathbf{b}) - f(\mathbf{a})\|_{\mathbb{R}^p} &= \left\| \int_0^1 \mathrm{d}f_{\mathbf{a}+t(\mathbf{b}-\mathbf{a})}(\mathbf{b} - \mathbf{a}) \, \mathrm{d}t \right\|_{\mathbb{R}^p} \\ &\leqslant \int_0^1 \|\mathrm{d}f_{\mathbf{a}+t(\mathbf{b}-\mathbf{a})}(\mathbf{b} - \mathbf{a})\|_{\mathbb{R}^p} \, \mathrm{d}t \\ &\leqslant \left(\int_0^1 |||\mathrm{d}f_{\mathbf{a}+t(\mathbf{b}-\mathbf{a})}|||_{\mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)} \, \mathrm{d}t \right) \|(\mathbf{b} - \mathbf{a})\|_{\mathbb{R}^n} \\ &\leqslant M_K \|(\mathbf{b} - \mathbf{a})\|_{\mathbb{R}^n}. \end{aligned}$$

L'inégalité des accroissements finis est démontrée. □

[68] La borne supérieure d'une fonction continue sur un ensemble fermé et borné existe, voir le théorème 9.2 p. 564.

[69] Cette application est bien définie sur $[0,1]$ puisque $\mathbf{a} + t(\mathbf{b} - \mathbf{a}) \in [\mathbf{a}, \mathbf{b}]$ pour $t \in [0,1]$ et par conséquent $\mathbf{a} + t(\mathbf{b} - \mathbf{a}) \in K$.

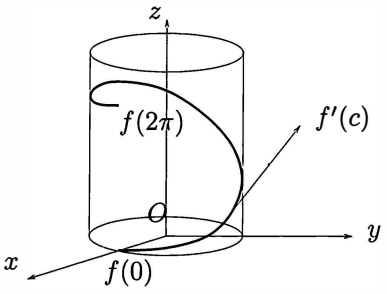
Remarques

1. Dans le théorème 10.6, l'hypothèse « f de classe $\mathcal{C}^1$ » est une condition suffisante. Le théorème est encore vrai si f est seulement différentiable, mais dans ce cas rien n'assure que M_K soit fini (car l'application $\mathrm{d}f$ n'est plus continue).

2. Contrairement au cas des fonctions d'une variable réelle, il n'y a en général pas égalité entre $\|f(\mathbf{b}) - f(\mathbf{a})\|_{\mathbb{R}^p}$ et $M_K\|\mathbf{a} - \mathbf{b}\|_{\mathbb{R}^n}$. Par exemple pour l'application

$$f : t \in \mathbb{R} \longmapsto \big(\cos(t), \sin(t), t\big) \in \mathbb{R}^3$$

il n'existe aucun réel $c \in [0, 2\pi]$ tel que $f(2\pi) - f(0) = 2\pi f'(c)$ car $f(2\pi) - f(0) = (0, 0, 2\pi)$ alors que $f'(c) = (-\sin(c), \cos(c), 1)$. Géométriquement, cela se traduit par le fait que le vecteur tangent à la courbe paramétrée par f n'est jamais colinéaire au vecteur $(0, 0, 2\pi)$.



◇

L'inégalité des accroissements finis prend la forme suivante lorsque l'ensemble K est convexe.

Corollaire 10.2 *Soient U un ouvert de $\mathbb{R}^n$, f une application de U dans $\mathbb{R}^p$ de classe $\mathcal{C}^1$ et K un sous-ensemble fermé, borné et* convexe *de U. Pour tout $(\mathbf{a}, \mathbf{b}) \in K^2$ on a*

$$\|f(\mathbf{b}) - f(\mathbf{a})\|_{\mathbb{R}^p} \leqslant M_K\|\mathbf{b} - \mathbf{a}\|_{\mathbb{R}^n},$$

où $M_K = \sup_{\mathbf{x}\in K} |||\mathrm{d}f_{\mathbf{x}}|||_{\mathcal{L}(\mathbb{R}^n,\mathbb{R}^p)}$.

Le résultat suivant qui découle du théorème des accroissements finis permet de caractériser les applications lipschitziennes sur un ensemble convexe U parmi les applications de classe $\mathcal{C}^1$ sur U.

Proposition 10.18 *Soient U un ouvert* convexe *de $\mathbb{R}^n$, f une application de U dans $\mathbb{R}^p$ de classe $\mathcal{C}^1$ et $k \in \mathbb{R}^+$. L'application f est k-lipschitzienne*[70] *si et seulement si*

$$\forall \mathbf{x} \in U \quad |||\,\mathrm{d}f_{\mathbf{x}}|||_{\mathcal{L}(\mathbb{R}^n,\mathbb{R}^p)} \leqslant k.$$

[70] La définition d'application k-lipschitzienne est donnée p. 558.

Démonstration $\trianglerighteq$ Supposons que pour tout $\mathbf{x} \in U$ on ait $|||\mathrm{d}f_{\mathbf{x}}|||_{\mathcal{L}(\mathbb{R}^n,\mathbb{R}^p)} \leqslant k$ et montrons que l'application f est k-lipschitzienne, c'est-à-dire montrons que :

$$\forall (\mathbf{a},\mathbf{b}) \in U^2 \quad \|f(\mathbf{b}) - f(\mathbf{a})\|_{\mathbb{R}^p} \leqslant k\|\mathbf{b}-\mathbf{a}\|_{\mathbb{R}^n}.$$

Considérons deux éléments $\mathbf{a}$ et $\mathbf{b}$ de l'ouvert U. Puisque U est un ensemble convexe, il s'ensuit que le segment $[\mathbf{a},\mathbf{b}]$ est inclus dans U. En considérant l'inégalité des accroissements finis avec $K = [\mathbf{a},\mathbf{b}]$ qui est un sous-ensemble fermé et borné de U, on a

$$\|f(\mathbf{b}) - f(\mathbf{a})\|_{\mathbb{R}^p} \leqslant M_K\|\mathbf{b}-\mathbf{a}\|_{\mathbb{R}^n},$$

où $M_K = \sup_{\mathbf{x}\in K} |||\mathrm{d}f_{\mathbf{x}}|||_{\mathcal{L}(\mathbb{R}^n,\mathbb{R}^p)}$. Puisque $|||\mathrm{d}f_{\mathbf{x}}|||_{\mathcal{L}(\mathbb{R}^n,\mathbb{R}^p)} \leqslant k$ pour tout $\mathbf{x} \in U$, on a donc $M_K \leqslant k$ ce qui montre que l'application f est k-lipschitzienne.

$\trianglelefteq$ Supposons que l'application f soit k-lipschitzienne et montrons que pour tout $\mathbf{a} \in U$ on a $|||\mathrm{d}f_{\mathbf{a}}|||_{\mathcal{L}(\mathbb{R}^n,\mathbb{R}^p)} \leqslant k$. Considérons un élément $\mathbf{a}$ de U ; l'application f étant k-lipschitzienne, pour $\mathbf{h} \in \mathbb{R}^n$ dans un voisinage de $\mathbf{0}$, on a

$$\|f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a})\|_{\mathbb{R}^p} \leqslant k\,\|\mathbf{h}\|_{\mathbb{R}^n}. \tag{25}$$

D'autre part, puisque l'application f est différentiable, on a aussi

$$f(\mathbf{a}+\mathbf{h}) = f(\mathbf{a}) + \mathrm{d}f_{\mathbf{a}}(\mathbf{h}) + \|\mathbf{h}\|_{\mathbb{R}^n}\varepsilon(\mathbf{h}) \quad \text{avec} \quad \lim_{\mathbf{h}\to\mathbf{0}} \varepsilon(\mathbf{h}) = \mathbf{0}.$$

On en déduit que

$$\mathrm{d}f_{\mathbf{a}}(\mathbf{h}) = f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) - \|\mathbf{h}\|_{\mathbb{R}^n}\varepsilon(\mathbf{h}),$$

puis par l'inégalité triangulaire, que

$$\|\mathrm{d}f_{\mathbf{a}}(\mathbf{h})\|_{\mathbb{R}^p} \leqslant \|f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a})\|_{\mathbb{R}^p} + \|\mathbf{h}\|_{\mathbb{R}^n}\|\varepsilon(\mathbf{h})\|_{\mathbb{R}^p}. \tag{26}$$

En combinant les équations (25) et (26), on obtient

$$\|\mathrm{d}f_{\mathbf{a}}(\mathbf{h})\|_{\mathbb{R}^p} \leqslant \|\mathbf{h}\|_{\mathbb{R}^n}(k + \|\varepsilon(\mathbf{h})\|_{\mathbb{R}^p}).$$

Comme $\lim_{\mathbf{h}\to\mathbf{0}} \varepsilon(\mathbf{h}) = \mathbf{0}$, on a[71]

$$\forall \xi \in \mathbb{R}_+^* \quad \exists \eta \in \mathbb{R}_+^* \quad (\ \|\mathbf{h}\|_{\mathbb{R}^n} \leqslant \eta \implies \|\varepsilon(\mathbf{h})\|_{\mathbb{R}^p} \leqslant \xi\ ).$$

On en déduit que pour tout $\xi \in \mathbb{R}_+^*$, il existe $\eta \in \mathbb{R}_+^*$ tel que si $\|\mathbf{h}\|_{\mathbb{R}^n} \leqslant \eta$ alors

$$\|\mathrm{d}f_{\mathbf{a}}(\mathbf{h})\|_{\mathbb{R}^p} \leqslant \|\mathbf{h}\|_{\mathbb{R}^n}(k+\xi). \tag{27}$$

Pour $\mathbf{H} \in \mathbb{R}_*^n$, l'élément $\mathbf{h} = \dfrac{\eta\mathbf{H}}{\|\mathbf{H}\|_{\mathbb{R}^n}}$ vérifie $\|\mathbf{h}\|_{\mathbb{R}^n} = \eta$ et d'après la relation (27) on a

$$\left\|\mathrm{d}f_{\mathbf{a}}\left(\frac{\eta\mathbf{H}}{\|\mathbf{H}\|_{\mathbb{R}^n}}\right)\right\|_{\mathbb{R}^p} \leqslant \left\|\frac{\eta\mathbf{H}}{\|\mathbf{H}\|_{\mathbb{R}^n}}\right\|_{\mathbb{R}^n}(k+\xi). \tag{28}$$

[71] Voir la définition 9.1 p. 536.

En utilisant le fait que la différentielle $\mathrm{d}f_a$ est une application linéaire, on a

$$\left\| \mathrm{d}f_{\mathbf{a}} \left(\frac{\eta \mathbf{H}}{\|\mathbf{H}\|_{\mathbb{R}^n}} \right) \right\|_{\mathbb{R}^p} = \frac{\eta}{\|\mathbf{H}\|_{\mathbb{R}^n}} \|\mathrm{d}f_{\mathbf{a}}(\mathbf{H})\|_{\mathbb{R}^p},$$

et la relation (28) s'écrit

$$\|\mathrm{d}f_{\mathbf{a}}(\mathbf{H})\|_{\mathbb{R}^p} \leqslant (k + \xi) \, \|\mathbf{H}\|_{\mathbb{R}^n}.$$

On a ainsi établi que, pour tout $\xi \in \mathbb{R}_+^*$ et pour tout $\mathbf{H} \in \mathbb{R}_*^n$ on a $\|\mathrm{d}f_{\mathbf{a}}(\mathbf{H})\|_{\mathbb{R}^p} \leqslant (k + \xi) \, \|\mathbf{H}\|_{\mathbb{R}^n}$. En considérant la limite à droite lorsque ξ tend vers 0, on obtient :

$$\forall \mathbf{H} \in \mathbb{R}_*^n \quad \|\mathrm{d}f_{\mathbf{a}}(\mathbf{H})\|_{\mathbb{R}^p} \leqslant k \|\mathbf{H}\|_{\mathbb{R}^n},$$

ce qui implique que :

$$|||\mathrm{d}f_{\mathbf{a}}|||_{\mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)} = \sup_{\mathbf{H} \in \mathbb{R}^n \setminus \{0\}} \frac{\|\mathrm{d}f_{\mathbf{a}}(\mathbf{H})\|_{\mathbb{R}^p}}{\|\mathbf{H}\|_{\mathbb{R}^n}} \leqslant k.$$

La proposition est démontrée. □

Le théorème des accroissements finis permet également d'établir le résultat suivant.

PROPOSITION 10.19 *Soient U un ouvert* convexe *de $\mathbb{R}^n$ et f une application de U dans $\mathbb{R}^p$ de classe $\mathcal{C}^1$. L'application f est constante sur U si et seulement si la différentielle de f est nulle sur U.*

Démonstration ⊵ Si l'application f est constante sur un ouvert U alors elle est différentiable sur U et sa différentielle est nulle, voir p. 610.

⊵ Réciproquement, quel que soit $(\mathbf{a}, \mathbf{b}) \in U^2$, on a toujours $[\mathbf{a}, \mathbf{b}] \subset U$ car U est convexe. Si $\mathrm{d}f = 0$ alors l'inégalité des accroissements finis implique que :

$$\|f(\mathbf{a}) - f(\mathbf{b})\|_{\mathbb{R}^p} \leqslant 0.$$

On en déduit que pour tout $(\mathbf{a}, \mathbf{b}) \in U^2$, on a $\|f(\mathbf{a}) - f(\mathbf{b})\|_{\mathbb{R}^p} = 0$ c'est-à-dire $f(\mathbf{a}) = f(\mathbf{b})$: l'application f est constante sur U. □

Remarque Dans la proposition 10.19, l'hypothèse de convexité sert à appliquer l'inégalité des accroissements finis. Le résultat « f est constante si et seulement si $\mathrm{d}f$ est nulle » est vrai plus généralement sur tout ouvert U *connexe*[72]. Notons qu'un ensemble convexe est toujours connexe. ◇

[72] Voir la définition 7.12 p. 384.

10.4.2 Formule de Taylor

L'objet de cette partie est de généraliser aux cas de fonctions de plusieurs variables la formule de Taylor vue dans le cas des applications réelles de la variable réelle[73]. Rappelons la formule de Taylor-Young pour les fonctions réelles de la variable réelle : étant donné une application f définie sur un intervalle ouvert I de $\mathbb{R}$, à valeurs dans $\mathbb{R}$ et de classe $\mathcal{C}^q$, on a pour tout $a \in I$ et pour tout $h \in \mathbb{R}$ tel que $a+h \in I$,

$$\begin{aligned} f(a+h) &= f(a) + hf'(a) + \frac{h^2}{2} f''(a) + \cdots + \frac{h^q}{q!} f^{(q)}(a) + \mathcal{O}_0(h^q) \\ &= f(a) + \sum_{j=1}^{q} \frac{h^j}{j!} f^{(j)}(a) + \mathcal{O}_0(h^q). \end{aligned}$$

La formule de Taylor-Young pour les fonctions de $\mathbb{R}^n$ dans $\mathbb{R}^p$, que nous admettrons, s'énonce de la manière suivante.

THÉORÈME 10.7 (Formule de Taylor)
Soient U un ouvert de $\mathbb{R}^n$ et f une application de U dans $\mathbb{R}^p$ de classe $\mathcal{C}^q$. Pour tout $\mathbf{a} \in U$ et $\mathbf{h} \in U$ tel que $\mathbf{a}+\mathbf{h} \in U$, on a

$$\begin{aligned} f(\mathbf{a}+\mathbf{h}) &= f(\mathbf{a}) + \mathrm{d}f_{\mathbf{a}}(\mathbf{h}) + \frac{1}{2!}\mathrm{d}^2 f_{\mathbf{a}}(\mathbf{h},\mathbf{h}) + \cdots \\ &\qquad + \frac{1}{q!}\mathrm{d}^q f_{\mathbf{a}}(\underbrace{\mathbf{h},\ldots,\mathbf{h}}_{q \text{ fois}}) + \mathcal{O}_{\mathbf{0}}(\|\mathbf{h}\|_{\mathbb{R}^n}^q) \\ &= f(\mathbf{a}) + \sum_{j=1}^{q} \frac{1}{j!}\mathrm{d}^j f_{\mathbf{a}}(\underbrace{\mathbf{h},\ldots,\mathbf{h}}_{j \text{ fois}}) + \mathcal{O}_{\mathbf{0}}(\|\mathbf{h}\|_{\mathbb{R}^n}^q). \end{aligned}$$

Rappelons[74] que la différentielle d'ordre j où $j \in \mathbb{N}^*$ d'une fonction de $\mathbb{R}^n$ dans $\mathbb{R}^p$ peut être vue comme une application j-linéaire de $(\mathbb{R}^n)^j$ dans $\mathbb{R}^p$ ce qui justifie la notation $\mathrm{d}^j f_{\mathbf{a}}(\mathbf{h},\ldots,\mathbf{h})$. Les différentielles première et seconde d'une application f définie sur un ouvert U de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}$ s'écrivent en $\mathbf{a} \in U$ de façon matricielle respectivement

$$\mathrm{d}f_{\mathbf{a}}(\mathbf{h}) = \mathrm{J}_f(\mathbf{a})\ {}^t\mathbf{h} \quad \text{et} \quad \mathrm{d}^2 f_{\mathbf{a}}(\mathbf{h},\mathbf{h}) = \mathbf{h}\ \mathrm{H}_f(\mathbf{a})\ {}^t\mathbf{h}.$$

On peut ainsi énoncer la formule de Taylor à l'ordre 2 pour les applications à valeurs réelles sous forme matricielle.

---

[73] Voir le *Cours de première année*, p. 778.
[74] Voir la p. 645.

TAYLOR, Brook (1685, Edmonton - 1731, Londres).



En 1715, Brook Taylor publie son œuvre principale *Methodus incrementorum directa et inversa* dans laquelle figure la formule qui porte son nom, sans mention du reste ni étude de convergence toutefois. Son intérêt pour le problème des cordes vibrantes l'amène par ailleurs à l'étude des équations différentielles du second ordre et à l'existence de solutions singulières pour ces équations. En 1717, il publie également un traité sur les problèmes de la perspective géométrique.

THÉORÈME 10.8 (Formule de Taylor à l'ordre 2)
Soient U un ouvert de $\mathbb{R}^n$ et f une application de U dans $\mathbb{R}$ de classe $\mathcal{C}^2$. Pour tout $\mathbf{a} \in U$ et $\mathbf{h} \in U$ tel que $\mathbf{a}+\mathbf{h} \in U$, on a

$$f(\mathbf{a}+\mathbf{h}) = f(\mathbf{a}) + \mathrm{J}_f(\mathbf{a})\ {}^t\mathbf{h} + \frac{1}{2}\mathbf{h}\ \mathrm{H}_f(\mathbf{a})\ {}^t\mathbf{h} + o_{\mathbf{0}}(\|\mathbf{h}\|^2_{\mathbb{R}^n}),$$

où J_f et H_f désignent respectivement la matrice jacobienne et la matrice hessienne de l'application f.

Exemple Considérons l'application f de $]-1,1[^2$ dans $\mathbb{R}$ définie par

$$f : (x_1, x_2) \in]-1,1[^2 \longmapsto \frac{-(1+x_1)}{((1+x_1)^2+x_2^2)^{3/2}} \in \mathbb{R}.$$

L'application f est de classe $\mathcal{C}^\infty$ car c'est un quotient de deux applications de classe $\mathcal{C}^\infty$ dont la fonction au dénominateur ne s'annule pas sur $]-1,1[^2$. La matrice jacobienne et la matrice hessienne de f au point $(0,0)$ sont[(75)]

$$\mathrm{J}_f(0,0) = \begin{pmatrix} 2 \\ 0 \end{pmatrix} \quad \text{et} \quad \mathrm{H}_f(0,0) = \begin{pmatrix} -6 & 0 \\ 0 & 3 \end{pmatrix}.$$

Puisque $f(0,0) = -1$, la formule de Taylor à l'ordre 2 au voisinage de $(0,0)$ pour f s'écrit

$$\begin{aligned} f(h_1,h_2) &= -1 + \begin{pmatrix} 2 \\ 0 \end{pmatrix} (h_1\ h_2) + \frac{1}{2}(h_1\ h_2) \begin{pmatrix} -6 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} + o_{\mathbf{0}}(\|\mathbf{h}\|^2_{\mathbb{R}^2}) \\ &= -1 + 2h_1 - 3h_1^2 + \frac{3}{2}h_2^2 + o_{(0,0)}(h_1^2+h_2^2). \end{aligned}$$

[(75)] Les dérivées partielles de f ont été calculées à la p. 654.

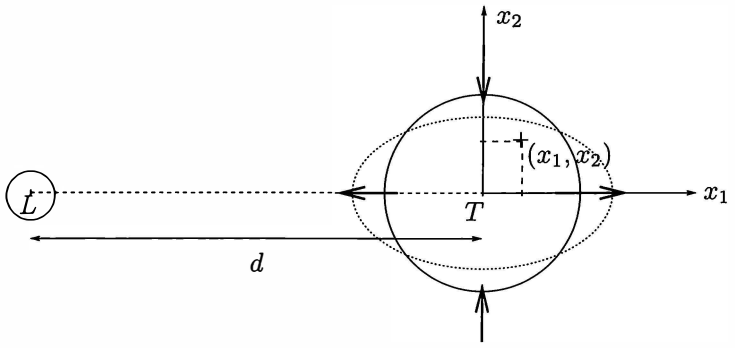


Fig. 6 Illustration du phénomène de marée.

Exercice 10.12 Soit f l'application de $]-1,1[^2$ dans $\mathbb{R}$ définie par

$$f : (x_1, x_2) \in]-1,1[^2 \longmapsto \frac{-x_2}{((1+x_1)^2 + x_2^2)^{3/2}} \in \mathbb{R}.$$

Écrire le développement de Taylor à l'ordre 2 de cette application au voisinage de $(0,0)$.

Application de la formule de Taylor aux phénomènes des marées

Le potentiel d'interaction $\mathcal{P}$ entre deux corps de masse m et M est donné par $\mathcal{P} = \mathcal{G}mM/r$ où $\mathcal{G}$ est la constante de gravitation et r la distance entre les deux corps. Intéressons-nous à l'interaction de la lune sur un point au voisinage de la Terre. Notons T le centre de la terre, L celui de la lune et d la distance terre-lune. Considérons un point, de masse m situé non loin de la terre. Notons (x_1, x_2) ses coordonnées dans le repère centré en T et d'axe des abscisses selon la droite (TL), voir la figure 6.

Le potentiel d'interaction de la lune (de masse M) sur ce point est donné par :

$$\mathcal{P}(x_1, x_2) = \frac{\mathcal{G}mM}{\sqrt{(d+x_1)^2 + x_2^2}}.$$

La force de gravité exercée par la lune est donnée par le gradient du potentiel d'interaction $\mathcal{P}$:

$$F(x_1, x_2) = \nabla\mathcal{P}(x_1, x_2) = \begin{pmatrix} -\dfrac{\mathcal{G}mM(d+x_1)}{((d+x_1)^2 + x_2^2)^{3/2}} \\ -\dfrac{\mathcal{G}mMx_2}{((d+x_1)^2 + x_2^2)^{3/2}} \end{pmatrix}.$$

Essayons d'analyser grâce à la formule de Taylor l'action de la force de gravité dans le cas où les grandeurs x_1 et x_2 sont petites devant la distance d. On fait apparaître les quantités x_1/d et x_2/d qui sont « petites devant 1 » de la façon suivante :

$$F(x_1,x_2)=\begin{pmatrix} -\dfrac{\mathcal{G}mM(1+x_1/d)}{\mathrm{d}^2\left((1+x_1/d)^2+(x_2/d)^2\right)^{3/2}} \\ -\dfrac{\mathcal{G}mMx_2/d}{\mathrm{d}^2\left((1+x_1/d)^2+(x_2/d)^2\right)^{3/2}} \end{pmatrix} = \frac{\mathcal{G}mM}{d^2}\, f(x_1/d,x_2/d)$$

où

$$f : (x_1,x_2)\in]-1,1[^2 \longmapsto \left(\frac{-(1+x_1)}{((1+x_1)^2+x_2^2)^{3/2}}, \frac{-x_2}{((1+x_1)^2+x_2^2)^{3/2}}\right)\in\mathbb{R}^2.$$

Les formules de Taylor à l'ordre 2 au voisinage de $(0,0)$ pour chacune des composantes de l'application f ont été établies p. 684 et au cours de l'exercice 10.12. On en déduit le développement de Taylor à l'ordre 1 en $(0,0)$ suivant pour F

$$F(x_1,x_2)=F(0,0)+\begin{pmatrix} \dfrac{2\mathcal{G}mM}{d^3}x_1 \\ -\dfrac{\mathcal{G}mM}{d^3}x_2 \end{pmatrix} + \mathcal{O}_{(0,0)}\left(\left(\frac{x_1}{d}\right)^2+\left(\frac{x_2}{d}\right)^2\right).$$

Cette relation indique que lorsque x_1 et x_2 sont petits devant d, les forces de gravité tendent à allonger le corps T dans l'axe (LT) (c'est-à-dire augmenter la première composante de F lorsque x_1 augmente) et à le contracter dans la direction perpendiculaire (c'est-à-dire diminuer la seconde composante de F lorsque x_2 augmente), voir la figure 6.

C'est principalement ce phénomène qui est à l'origine des marées sur les océans. On explique ainsi qu'il y a une marée « haute » toutes les 12h30 environ, la lune faisant le tour de la terre en 25h environ.

Tab. 10.1 Horaires des marées à Trébeurden (22) en janvier 2008.

Date	Matin			Après-midi		
	Coef.	Pleine mer	Basse mer	Coef.	Pleine mer	Basse mer
1	45	0h35	6h56	41	12h56	19h27
2	38	1h33	7h56	37	14h00	20h32
3	36	2h42	9h08	38	15h13	21h43
4	40	3h54	10h19	43	16h23	22h48
5	47	4h54	11h18	51	17h20	23h41
6	55	5h43	–	59	18h06	12h07
7	63	6h25	0h28	67	18h47	12h50

10.5 Extrema d'une fonction réelle de plusieurs variables

Dans cette partie on cherche à étudier l'existence d'extrema pour une fonction de $\mathbb{R}^n$ dans $\mathbb{R}$, puis à calculer ces éventuels extrema.

DÉFINITION 10.17 (Extrema locaux)
Soit f une application définie sur un sous-ensemble X de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}$.

✗ *On dit que f présente un* maximum local *en $\mathbf{a} \in X$ s'il existe une boule ouverte $B(\mathbf{a}, r)$ de centre $\mathbf{a}$ et de rayon $r \in \mathbb{R}_+^*$ tel que*

$$\forall \mathbf{x} \in B(\mathbf{a}, r) \cap X \quad f(\mathbf{x}) \leqslant f(\mathbf{a}).$$

✗ *On dit que ce maximum est* strict *si l'inégalité est stricte pour $\mathbf{x} \neq \mathbf{a}$.*

✗ *On dit que f présente un* minimum local *en $\mathbf{a} \in X$ s'il existe une boule ouverte $B(\mathbf{a}, r)$ de centre $\mathbf{a}$ et de rayon $r \in \mathbb{R}_+^*$ tel que*

$$\forall \mathbf{x} \in B(\mathbf{a}, r) \cap X \quad f(\mathbf{x}) \geqslant f(\mathbf{a}).$$

✗ *On dit que ce minimum est* strict *si l'inégalité est stricte pour $\mathbf{x} \neq \mathbf{a}$.*

✗ *On appelle* extremum local *un minimum ou un maximum local.*

DÉFINITION 10.18 (Extrema globaux)
Soit f une application définie sur un sous-ensemble X de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}$.

✗ *On dit que f présente un* maximum global *en $\mathbf{a} \in X$ si*

$$\forall \mathbf{x} \in X \quad f(\mathbf{x}) \leqslant f(\mathbf{a}).$$

✗ *On dit que f présente un* minimum global *en $\mathbf{a} \in X$ si*

$$\forall \mathbf{x} \in X \quad f(\mathbf{x}) \geqslant f(\mathbf{a}).$$

✗ *On appelle* extremum global *un minimum ou un maximum global.*

Dans le cas de fonctions de plusieurs variables, contrairement au cas de fonctions d'une variable réelle, la visualisation des extrema est délicate. Considérons l'application $f : (x_1, x_2) \in X \mapsto 1 + 30x_1^2 + 20x_1^3 - 15x_1^4 - 12x_1^5 + 8x_2^2$ où X est l'ensemble ouvert $X =]-2, 2[\times]-1, 1[$ dont la représentation graphique est donnée à la figure 7. On observe que sur l'ensemble X, l'application f admet deux extrema :un minimum local et un minimum global. Sur l'ensemble fermé $\overline{X} = [-2, 2] \times [-1, 1]$ l'application f possède, en plus des deux minima sur X, six maxima et six autre minima (situés sur la frontière $\overline{X} \setminus X$ de X).

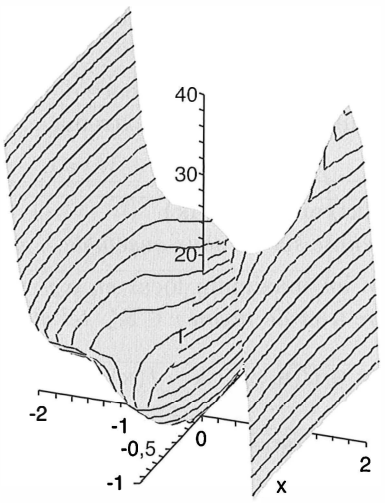


Fig. 7 Représentation graphique des extrema d'une application de $\mathbb{R}^2$ dans $\mathbb{R}$.

Remarque Généralement les deux problèmes mathématiques « montrer qu'il existe un objet » et « trouver un tel objet » sont assez différents. Le second problème entraînant le premier (car si on a trouvé un objet c'est bien qu'il existe !), il est en général plus difficile à résoudre. Par exemple, nous avons déjà à notre disposition un résultat d'existence d'extrema[(76)] :

« Toute application continue sur un ensemble fermé et borné est une application bornée qui atteint ses bornes. »

Ce résultat d'existence d'extrema globaux ne donne toutefois aucune indication sur la valeur de ces extrema. Dans ce qui suit, nous allons voir que le cadre pour obtenir des informations sur les valeurs des extrema s'applique à des fonctions définies sur des ensembles ouverts. ◇

10.5.1 Conditions nécessaires d'existence d'un extremum local

L'étude des extrema de fonctions de $\mathbb{R}$ dans $\mathbb{R}$ a été effectuée dans le *Cours de première année* et on a vu que les extrema d'une application f d'un intervalle ouvert $I \subset \mathbb{R}$ à valeurs dans $\mathbb{R}$ qui est dérivable sur I sont atteints en des points où la dérivée f' s'annule. Comme pour les fonctions de $\mathbb{R}$ dans $\mathbb{R}$, nous allons voir qu'une condition nécessaire pour qu'une application différentiable sur un ouvert U de $\mathbb{R}^n$ admette un extrema local en $\mathbf{a} \in U$ est que sa différentielle soit nulle en $\mathbf{a}$.

(76) Voir le théorème 9.2 p. 564

DÉFINITION 10.19 *Soient U un ouvert de $\mathbb{R}^n$ et f une application de U dans $\mathbb{R}$, différentiable sur U. On appelle* point critique *de f un élément $\mathbf{a} \in U$ tel que $\mathrm{d}f_{\mathbf{a}} = 0$.*

Remarque La condition $\mathrm{d}f_{\mathbf{a}} = 0$ peut se réinterpréter en termes matriciels par la relation $\mathrm{J}_f(\mathbf{a}) = 0$ et en termes vectoriels par $\nabla f(\mathbf{a}) = 0$. Par la suite, on utilisera alternativement ces trois aspects selon le contexte. ◇

THÉORÈME 10.9 *Soit f une application définie sur un ouvert U de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}$ qui est différentiable sur U. Si f présente un extremum local en $\mathbf{a} \in U$ alors $\mathbf{a}$ est un point critique de f.*

Démonstration Si f admet un extremum local en $\mathbf{a} = (a_1, \ldots, a_n)$ alors la j-ième fonction partielle $\varphi_j : x_j \in \mathbb{R} \mapsto f(a_1, \ldots, a_{j-1}, x_j, a_{j+1}, \ldots, a_n)$ définie sur un voisinage V_{a_j} de a_j présente un extremum local en a_j. Or, pour tout $j \in \{1, \ldots, n\}$ la fonction φ_j est une fonction réelle de la variable réelle qui est dérivable. Puisqu'elle admet un extremum local en a_j on a[(77)] $\varphi_j'(a_j) = 0$. Comme $\varphi_j'(a_j) = \frac{\partial f}{\partial x_j}(\mathbf{a})$ on en déduit que pour tout $j \in \{1, \ldots, n\}$ la dérivée partielle $\frac{\partial f}{\partial x_j}(\mathbf{a})$ est nulle et donc que

$$\forall \mathbf{h} \in \mathbb{R}^n \quad \mathrm{d}f_{\mathbf{a}}(\mathbf{h}) = \sum_{j=1}^{n} h_j \frac{\partial f}{\partial x_j}(\mathbf{a}) = 0.$$

On en conclut que $\mathrm{d}f_{\mathbf{a}} = 0$, c'est-à-dire que $\mathbf{a}$ est un point critique de f. □

La condition donnée par ce théorème 10.9 n'est pas suffisante : il existe des points critiques qui ne sont pas des extrema. D'un point de vue pratique, après avoir calculé les points critiques, il convient d'établir s'il s'agit ou non d'un extrema.

Exemple Considérons l'application

$$f : (x_1, x_2) \in]-2, 2[\times]-1, 1[\longmapsto 1 + 30x_1^2 + 20x_1^3 - 15x_1^4 - 12x_1^5 + 8x_2^2.$$

La représentation graphique de cette application est donnée à la figure 7. Les deux dérivées partielles de f en (x_1, x_2) valent

$$\frac{\partial f}{\partial x_1}(x_1, x_2) = 60x_1(1 - x_1)(1 + x_1)^2 \quad \text{et} \quad \frac{\partial f}{\partial x_2}(x_1, x_2) = 16x_2.$$

(77) Voir le *Cours de première année*, p. 736.

On en déduit que f admet les trois points critiques suivants : $(-1,0)$, $(0,0)$ et $(1,0)$. Le point $(-1,0)$ n'est pas un extrema de f puisque pour $\delta \in \mathbb{R}$ on a $f(-1+\delta,0)-f(-1,0) = -\delta^3(40-45\delta+12\delta^2)$ et cette quantité change de signe lorsque δ change de signe. Autrement dit, pour tout $\delta > 0$ assez petit on a $f(-1+\delta,0) < f(-1,0)$ ce qui implique $(-1,0)$ ne peut pas être un minimum, et pour tout $\delta < 0$ assez petit on a $f(-1+\delta,0) > f(-1,0)$ ce qui implique $(-1,0)$ ne peut pas être un maximum.

On sera vigilant au fait que le théorème 10.9 s'applique à des ensembles ouverts. Une application définie sur un ensemble non ouvert peut admettre un extrema sans que sa différentielle s'annule. Par exemple, l'application $(x_1,x_2) \in \overline{B}(0,1) \mapsto x_1^2+x_2^2$ présente des maxima en tout point du bord du disque $\overline{B}(0,1)$ alors que sa différentielle ne s'annule qu'en $(0,0)$.

Remarque Comme dans le cas des fonctions de $\mathbb{R}$ dans $\mathbb{R}$, la méthodologie pour déterminer les extrema globaux d'une fonction sur un ensemble X quelconque consiste à déterminer les extrema locaux dans l'intérieur $\mathrm{Int}\,(X)$ de l'ensemble, puis à comparer ces extrema entre eux ainsi qu'aux valeurs prises par la fonction sur la frontière de X. ◇

Exemple Déterminons quels sont les triangles d'aire maximale inscrits dans le cercle de centre $(0,0)$ et de rayon 1. On suppose[78] que l'un des sommets du triangle est le point C de coordonnées $(1,0)$, que les deux autres sommets A et B du triangles sont déterminés par les angles $\alpha = (\widehat{OC,OA})$ et $\beta = (\widehat{OC,OB})$ avec[79] $0 \leqslant \alpha \leqslant \beta \leqslant 2\pi$, voir la figure 8.

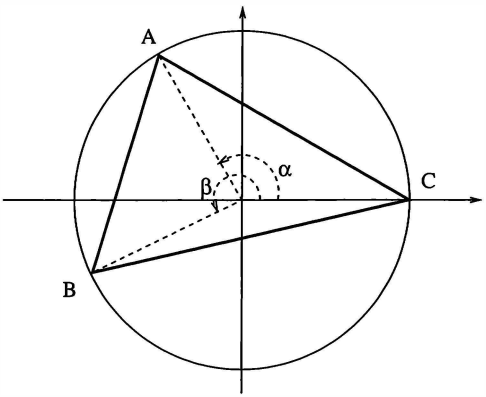


Fig. 8 Situation considérée : un triangle inscrit dans un cercle.

[78] L'aire d'un triangle ne change pas si on fait faire une rotation à ce triangle. On peut toujours supposer, quitte à effectuer une telle rotation, que l'un des sommets du triangle a pour coordonnées $(1,0)$.

[79] Quitte à échanger A et B, on peut supposer que $0 \leqslant \alpha \leqslant \beta \leqslant 2\pi$.

L'aire $\mathcal{A}(\alpha,\beta)$ du triangle ABC est donnée par

$$\begin{aligned}\mathcal{A}(\alpha,\beta) &= \frac{1}{2}\|\mathbf{CA}\wedge\mathbf{CB}\| = \frac{1}{2}|(\cos\alpha-1)\sin\beta-(\cos\beta-1)\sin\alpha| \\ &= 2\sin\left(\frac{\alpha}{2}\right)\sin\left(\frac{\beta}{2}\right)\sin\left(\frac{\beta-\alpha}{2}\right).\end{aligned}$$

Il s'agit de chercher le maximum de la fonction $\mathcal{A}$ sur l'ensemble

$$K=\{(\alpha,\beta)\in\mathbb{R}^2\ ;\quad 0\leqslant\alpha\leqslant\beta\leqslant 2\pi\}.$$

Remarquons que $\mathcal{A}$ étant une fonction continue sur l'ensemble K qui est fermé et borné dans $\mathbb{R}^2$, la fonction $\mathcal{A}$ admet[80] un maximum global sur K. Le théorème 10.9 qui fournit un moyen d'obtenir d'éventuels maxima ne s'applique que sur des ensembles ouverts. Il faut donc étudier séparément les maxima situés sur la frontière de K puis ceux situés sur l'intérieur de K (qui est un ouvert). Sur la frontière de K, on a soit $\alpha=0$, soit $\alpha=\beta$, soit $\beta=2\pi$ et dans tous les cas on a $\mathcal{A}=0$, ce qui représente un minimum pour $\mathcal{A}$. L'intérieur de K est l'ouvert $U=\{(\alpha,\beta)\in\mathbb{R}^2\ ;\ 0<\alpha<\beta<2\pi\}$. Les points critiques sur U sont obtenus en résolvant les équations

$$\frac{\partial\mathcal{A}}{\partial\alpha}(\alpha,\beta)=0\quad\text{et}\quad\frac{\partial\mathcal{A}}{\partial\beta}(\alpha,\beta)=0.$$

Puisque $\dfrac{\partial\mathcal{A}}{\partial\alpha}(\alpha,\beta)=\sin\left(\dfrac{\beta}{2}\right)\sin(2\alpha-\beta)$ et $\dfrac{\partial\mathcal{A}}{\partial\beta}(\alpha,\beta)=\sin\left(\dfrac{\alpha}{2}\right)\sin(2\beta-\alpha)$, on en déduit que (α,β) est un point critique si et seulement si

$$0<\alpha<\beta<2\pi,\quad 2\alpha-\beta=0\ (\mathrm{mod}\,\pi)\quad\text{et}\quad 2\beta-\alpha=0\ (\mathrm{mod}\,\pi).$$

Autrement dit, (α,β) est un point critique si et seulement si $0<\alpha<\beta<2\pi$ et si il existe deux entiers k et k' tels que $2\alpha-\beta=k\pi$ et $2\beta-\alpha=k'\pi$. Ces deux dernières relations impliquent que α et β s'écrivent sous la forme

$$\alpha=\frac{2k+k'}{3}\pi\quad\text{et}\quad\beta=\frac{2k'+k}{3}\pi.$$

Pour assurer les conditions $0<\alpha<\beta<2\pi$ il faut que les entiers k et k' satisfassent $0<2k+k'$, $k<k'$ et $2k'+k>6$ ce qui impose pour valeurs de (k,k') les valeurs suivantes :

$$(k,k')\in\{(-1,3),(0,1),(0,2),(1,2)\}.$$

On en déduit qu'il existe exactement quatre points critiques pour l'application $\mathcal{A}$; ils correspondent aux valeurs de (α,β) obtenues à partir des couples (k,k') précédents, *i.e.*

$$(\alpha,\beta)\in\left\{\left(\frac{\pi}{3},\frac{5\pi}{3}\right),\left(\frac{\pi}{3},\frac{2\pi}{3}\right),\left(\frac{2\pi}{3},\frac{4\pi}{3}\right),\left(\frac{4\pi}{3},\frac{5\pi}{3}\right)\right\}.$$

[80] Voir le théorème 9.2 p. 564.

Parmi ces quatre points critiques, on sait qu'il y a au moins un point en lequel l'application $\mathcal{A}$ atteint son maximum. On peut calculer « à la main » les valeurs de $\mathcal{A}$ en chacun de ces points (ou observer sur les quatre figures lequel des triangles a la plus grande aire). Le maximum est atteint au point $(\alpha, \beta) = (2\pi/3, 4\pi/3)$. L'aire maximale d'un triangle inscrit dans le cercle vaut $\mathcal{A}(2\pi/3, 4\pi/3) = 3\sqrt{3}/4$ et correspond aux points $C = (1, 0)$, $A = (-1/2, \sqrt{3}/2)$ et $B = (-1/2, -\sqrt{3}/2)$. Cette configuration est celle d'un triangle équilatéral.

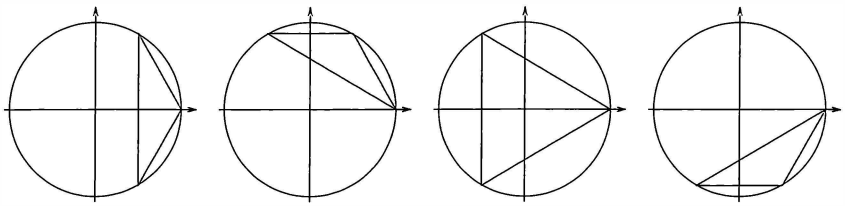


Fig. 9 Les quatre triangles correspondants aux quatre points critiques.

Dans cet exemple, nous avons réussi à trouver le maximum global grâce à deux propriétés : d'une part il n'y avait qu'un petit nombre de points critiques et d'autre part on était certain d'avoir un maximum global. On ne peut pas toujours conclure ainsi. En général il est nécessaire d'étudier la fonction au voisinage des points critiques pour savoir s'il s'agit d'un extremum.

10.5.2 Conditions suffisantes d'existence d'un extremum local

Dans le cas d'une application de classe $\mathcal{C}^2$, il est possible de déterminer à partir de l'expression de la différentielle seconde si un point critique correspond à un minimum, à un maximum ou si ce n'est pas un extremum.

On a vu dans le *Cours de première année*, p. 736, qu'une condition suffisante pour qu'une application f d'un ouvert I de $\mathbb{R}$ à valeurs dans $\mathbb{R}$ atteigne un extremum en $x_0 \in I$ est que $f'(x_0) = 0$ (*i.e.* que x_0 est un point critique de f) et que f' change de signe en x_0. Dans le cas d'applications deux fois dérivables, cette dernière condition peut s'exprimer sous la forme : f'' est de signe constant au voisinage de x_0. Nous allons chercher à caractériser les points critiques d'une manière analogue dans le cas de fonctions de plusieurs variables réelles.

Remarquons d'abord que la notion de dérivée seconde pour une application d'un ouvert U de $\mathbb{R}^n$ dans $\mathbb{R}$ de classe $\mathcal{C}^2$ se généralise par la notion de différentielle seconde en x_0, autrement dit *via* une application bilinéaire de $\mathbb{R}^n \times \mathbb{R}^n$ dans $\mathbb{R}$ pour laquelle être de signe constant est ambigu. La différentielle seconde peut s'exprimer à l'aide de la matrice hessienne et compte tenu du théorème de Schwarz, cette matrice est symétrique. Elle admet donc des valeurs propres qui sont toutes réelles. La condition « la dérivée est de signe constant » servant à caractériser les extrema dans le cas des fonctions d'une variable réelle devient pour les fonctions de $\mathbb{R}^n$ dans $\mathbb{R}$ de classe $\mathcal{C}^2$ une condition de signe sur les valeurs propres de la matrice hessienne.

THÉORÈME 10.10 *Soit f une application définie sur un ouvert U de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}$, de classe C^2 admettant $\mathbf{x}_0 \in U$ comme point critique. On suppose que les valeurs propres de la matrice hessienne $\mathrm{H}_f(\mathbf{x}_0)$ sont toutes non nulles.*

1. *Si toutes ses valeurs propres sont strictement positives alors $\mathbf{x}_0$ est un minimum local de f.*
2. *Si toutes ses valeurs propres sont strictement négatives alors $\mathbf{x}_0$ est un maximum local de f.*
3. *Si certaines valeurs propres sont strictement positives et d'autres sont strictement négatives alors $\mathbf{x}_0$ n'est ni un minimum ni un maximum (on dit que $\mathbf{x}_0$ est un* point col *ou* point selle*).*

Démonstration Le point $\mathbf{x}_0$ étant un point critique pour f, la formule de Taylor à l'ordre 2 appliquée à la fonction f en $\mathbf{x}_0$ indique que pour tout $\mathbf{h} \in \mathbb{R}^n$ tel que $\mathbf{x}_0 + \mathbf{h} \in U$, on a

$$f(\mathbf{x}_0 + \mathbf{h}) - f(\mathbf{x}_0) = \frac{1}{2}\mathbf{h}\ \mathrm{H}_f(\mathbf{x}_0)\ {}^t\mathbf{h} + \|\mathbf{h}\|^2\varepsilon(\mathbf{h}), \tag{29}$$

où ε est une fonction de $\mathbb{R}^n$ dans $\mathbb{R}$ de limite nulle en zéro. La matrice $\mathrm{H}_f(\mathbf{x}_0)$ est une matrice de $\mathcal{M}_n(\mathbb{R})$ symétrique, elle est donc[(81)] diagonalisable et la matrice de passage P est orthogonale. En désignant par $\lambda_1, \ldots, \lambda_n$ les n valeurs propres de $\mathrm{H}_f(\mathbf{x}_0)$, on a

$$\mathrm{H}_f(\mathbf{x}_0) = {}^t\mathrm{PDP}$$

où D est la matrice diagonale dont les éléments diagonaux sont $\lambda_1, \ldots, \lambda_n$. Le terme $\mathbf{h}\ \mathrm{H}_f(\mathbf{x}_0)\ {}^t\mathbf{h}$ dans l'égalité (29) s'écrit

$$\mathbf{h}\ \mathrm{H}_f(\mathbf{x}_0)\ {}^t\mathbf{h} = {}^t(\mathrm{P}^t\mathbf{h})\ \mathrm{D}\ \mathrm{P}^t\mathbf{h}.$$

En désignant par $k_1, \ldots, k_n$ les composantes du vecteur $\mathrm{P}^t\mathbf{h}$, la relation (29) devient, pour tout $\mathbf{h} \in \mathbb{R}^n$ tel que $\mathbf{x}_0 + \mathbf{h} \in U$,

$$f(\mathbf{x}_0 + \mathbf{h}) - f(\mathbf{x}_0) = \frac{1}{2}\sum_{i=1}^{n} \lambda_i k_i^2 + \|\mathbf{h}\|^2\varepsilon(\mathbf{h}). \tag{30}$$

$\trianglerighteq$ Intéressons-nous au cas où toutes les valeurs propres de la matrice hessienne sont strictement positives et notons λ_1 la plus petite valeur propre. On a

$$\sum_{i=1}^{n} \lambda_i k_i^2 \geqslant \lambda_1 \sum_{i=1}^{n} k_i^2 = \lambda_1 \|\mathbf{k}\|_2^2 = \lambda_1 \|\mathrm{P}^t\mathbf{h}\|_2^2.$$

La matrice P étant orthogonale, on a[(82)] $\|\mathrm{P}^t\mathbf{h}\|_2 = \|\mathbf{h}\|_2$ et la relation (30) utilisée avec la norme euclidienne sur $\mathbb{R}^n$ implique

$$f(\mathbf{x}_0 + \mathbf{h}) - f(\mathbf{x}_0) \geqslant \left(\frac{\lambda_1}{2} + \varepsilon(\mathbf{h})\right) \|\mathbf{h}\|_2^2.$$

(81) Voir le théorème 8.10 p. 505.

(82) Rappelons qu'une matrice orthogonale est la matrice d'une isométrie, et donc qu'elle « conserve » la norme, voir p. 493 à 490.

Comme $\lambda_1 > 0$ et que $\lim_{\mathbf{h}\to\mathbf{0}} \varepsilon(\mathbf{h}) = 0$, pour $\mathbf{h}$ assez petit on a $\frac{1}{2}\lambda_1 + \varepsilon(\mathbf{h}) > 0$. On en déduit que pour $\mathbf{x} = \mathbf{x}_0 + \mathbf{h}$ dans un voisinage de $\mathbf{x}_0$ on a $f(\mathbf{x}) \geqslant f(\mathbf{x}_0)$, ce qui permet d'affirmer que $\mathbf{x}_0$ est un minimum local pour f.

$\trianglerighteq$ Dans le cas où toutes les valeurs propres de la matrice hessienne sont négatives, un raisonnement analogue en considérant la plus grande valeur propre fournit le résultat annoncé.

$\trianglerighteq$ Supposons que l'on ait deux valeurs propres de signe opposé. Notons λ_1 une valeur propre strictement positive, λ_2 une valeur propre strictement négative et désignons par $\mathbf{e}_1$ (resp. $\mathbf{e}_2$) un vecteur propre unitaire associé à λ_1 (resp. λ_2). Considérons l'équation (29) avec $\mathbf{h} = \delta^t \mathrm{P}\mathbf{e}_1$ où $\delta \in \mathbb{R}$; on a

$$f(\mathbf{x}_0 + \delta^t \mathrm{P}\mathbf{e}_1) - f(\mathbf{x}_0) = \delta^2 \left(\frac{\lambda_1}{2} + \varepsilon(\delta^t \mathrm{P}\mathbf{e}_1) \right).$$

Comme $\lambda_1 > 0$ et comme $\lim_{\delta\to 0} \varepsilon(\delta^t \mathrm{P}\mathbf{e}_1) = 0$, pour δ assez petit on a $\frac{1}{2}\lambda_1 + \varepsilon(\delta^t \mathrm{P}\mathbf{e}_1) > 0$ et par conséquent $f(\mathbf{x}_0 + \delta^t \mathrm{P}\mathbf{e}_1) > f(\mathbf{x}_0)$. On en déduit que $\mathbf{x}_0$ n'est pas un minimum local pour f.

Considérons à présent l'équation (29) avec $\mathbf{h} = \delta^t \mathrm{P}\mathbf{e}_2$ où $\delta \in \mathbb{R}$:

$$f(\mathbf{x}_0 + \delta^t \mathrm{P}\mathbf{e}_2) - f(\mathbf{x}_0) = \delta^2 \left(\frac{\lambda_2}{2} + \varepsilon(\delta^t \mathrm{P}\mathbf{e}_2) \right).$$

Comme $\lambda_2 < 0$ et comme $\lim_{\delta\to 0} \varepsilon(\delta^t \mathrm{P}\mathbf{e}_2) = 0$, pour δ assez petit on a $\frac{1}{2}\lambda_2 + \varepsilon(\delta^t \mathrm{P}\mathbf{e}_2) < 0$ et par conséquent $f(\mathbf{x}_0 + \delta^t \mathrm{P}\mathbf{e}_2) > f(\mathbf{x}_0)$. On en déduit que $\mathbf{x}_0$ n'est pas un maximum local pour f.

On en conclut que dans le cas où la matrice hessienne $\mathrm{H}_f(\mathbf{x}_0)$ a des valeurs propres de signe opposé, $\mathbf{x}_0$ n'est ni un minimum ni un maximum. □

Si la matrice hessienne de f admet une valeur propre nulle alors le théorème 10.10 ne permet pas de conclure. En effet, supposons que λ_1 soit une valeur propre nulle de $\mathrm{H}_f(\mathbf{x}_0)$ et désignons par $\mathbf{e}_1$ un vecteur propre unitaire associé. En considérant l'équation (29) avec $\mathbf{h} = \delta^t \mathrm{P}\mathbf{e}_1$ où $\delta \in \mathbb{R}$, on obtient :

$$f(\mathbf{x}_0 + \delta^t \mathrm{P}\mathbf{e}_1) - f(\mathbf{x}_0) = \delta^2 \varepsilon(\delta^t \mathrm{P}\mathbf{e}_1).$$

Il est impossible, avec cette démarche, de déterminer le signe de la différence $f(\mathbf{x}_0 + \delta^t \mathrm{P}\mathbf{e}_1) - f(\mathbf{x}_0)$ puisque la seule hypothèse $\lim_{\mathbf{h}\to\mathbf{0}} \varepsilon(\mathbf{h}) = 0$ sur la fonction ε ne permet pas de déterminer le signe de $\delta^2 \varepsilon(\delta^t \mathrm{P}\mathbf{e}_1)$.

Plus généralement, sans supposer que les valeurs propres de la hessienne sont non nulles, on a la disjonction de cas donnée dans le corollaire suivant.

Corollaire 10.3 *Soit f une application définie sur un ouvert U de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}$ de classe $\mathcal{C}^2$ et admettant $\mathbf{x}_0 \in U$ comme point critique. Considérons les valeurs propres de la matrice $\mathrm{H}_f(\mathbf{x}_0)$:*

1. *soit toutes ses valeurs propres sont strictement positives : $\mathbf{x}_0$ est un minimum local de f ;*
2. *soit toutes ses valeurs propres sont strictement négatives : $\mathbf{x}_0$ est un maximum local de f ;*
3. *soit il existe au moins une valeur propre strictement positive et une valeur propre strictement négative : $\mathbf{x}_0$ est un point col de f.*
4. *Dans les autres cas*[83]*, on ne peut pas conclure directement.*

Exemples

1. Considérons l'application[84]

$$f : (x_1, x_2) \in]-2, 2[\times]-1, 1[\mapsto 1 + 30x_1^2 + 20x_1^3 - 15x_1^4 - 12x_1^5 + 8x_2^2$$

dont les dérivées partielles sont :

$$\frac{\partial f}{\partial x_1}(x_1, x_2) = 60x_1(1 - x_1)(1 + x_1)^2 \quad \text{et} \quad \frac{\partial f}{\partial x_2}(x_1, x_2) = 16x_2.$$

Il existe trois points critiques : $(0, 0)$, $(1, 0)$ et $(-1, 0)$. Les matrices hessiennes de f en ces trois points valent :

$$\mathrm{H}_f(0,0) = \begin{pmatrix} 60 & 0 \\ 0 & 16 \end{pmatrix}, \quad \mathrm{H}_f(1,0) = \begin{pmatrix} -240 & 0 \\ 0 & 16 \end{pmatrix} \quad \text{et} \quad \mathrm{H}_f(-1,0) = \begin{pmatrix} 0 & 0 \\ 0 & 16 \end{pmatrix}.$$

On a vu lors de l'exemple de la p. 689 que le point $(-1, 0)$ n'est ni un maximum ni un minimum. On détermine la nature des deux autres points critiques à l'aide du corollaire 10.3 : le point $(0, 0)$ est un minimum et le point $(1, 0)$ est un point col. La représentation graphique de l'application f est donnée p. 690.

2. Considérons l'application $f : (x_1, x_2) \in \mathbb{R}^2 \mapsto x_1^2 - x_2^2$ qui est de classe $\mathcal{C}^\infty$ et possède un seul point critique : $(0, 0)$. La hessienne de f en $(0, 0)$ vaut :

$$\mathrm{H}_f(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}.$$

D'après le théorème 10.10, le point $(0, 0)$ est un point col et il n'y a donc pas d'extrema pour l'application f.

---

[83] Il s'agit des cas sont les cas où les valeurs propres de la hessienne sont toutes de même signe et au moins une des valeurs propres est nulle.

[84] Voir l'exemple de la p. 689.

La démonstration du théorème 10.10 permet d'obtenir des informations supplémentaires sur le comportement de f au voisinage du point critique $(0,0)$. En effet, considérons deux vecteurs propres unitaires $\mathbf{e}_1 = (1\ 0)$ et $\mathbf{e}_2 = (0\ 1)$ associés respectivement aux valeurs propres 2 et -2 de la matrice $\mathrm{H}_f(0,0)$. La démonstration du théorème 10.10 indique que l'application f atteint un minimum lorsqu'on la restreint à la droite engendrée par le vecteur $\mathbf{e}_1$ et qu'elle atteint un maximum lorsqu'on la restreint à la droite engendrée par le vecteur $\mathbf{e}_2$.

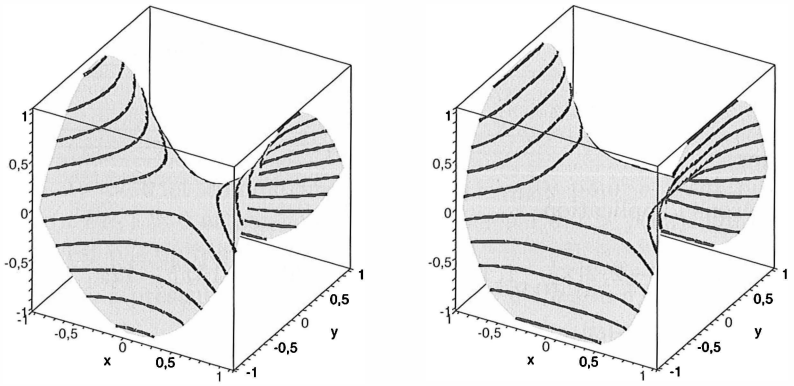


Fig. 10 Représentation des fonctions $(x_1, x_2) \mapsto x_1^2 - x_2^2$ et $(x_1, x_2) \mapsto x_1^4 - x_2^4$.

3. Considérons l'application $f : (x_1, x_2) \in \mathbb{R}^2 \mapsto x_1^4 - x_2^4$ qui est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}^2$ et possède un seul point critique : $(0,0)$. Remarquons que la hessienne de f en $(0,0)$ est la matrice nulle : ses deux valeurs propres sont nulles et le théorème 10.10 n'est pas applicable. Dans ce cas précis on peut facilement observer que f n'admet pas d'extremum en $(0,0)$: dans tout voisinage de $(0,0)$ il y a des points pour lesquels f est strictement négative, par exemple les points du type $(0,\delta)$ avec $\delta \in \mathbb{R}$ et d'autres pour lesquels f est strictement positive (les points du type $(\delta, 0)$) : l'application f n'admet pas d'extremum.

Exercice 10.13 Trouver les points critiques de l'application f définie sur $\mathbb{R}^2$ par $f(x_1, x_2) = x_1^2 + x_2^3$. Cette application admet-elle un extremum ?

10.6 Exercices de synthèse

Exercice 10.14 ***Étude des fonctions harmoniques du plan***

Soit f une application définie sur un ouvert U de $\mathbb{R}^2$ à valeurs dans $\mathbb{R}$ et de classe $\mathcal{C}^2$. Le laplacien de l'application f est, par définition, l'application Δf définie dans l'ouvert U par la relation suivante :

$$\Delta f(x,y) = \frac{\partial^2 f}{\partial x^2}(x,y) + \frac{\partial^2 f}{\partial y^2}(x,y).$$

Une telle application f est dite *harmonique* si son laplacien est nul sur U, i.e. si :

$$\forall (x,y) \in U \quad \Delta f(x,y) = 0.$$

PARTIE 1 - QUELQUES EXEMPLES DE FONCTIONS HARMONIQUES

1 - Trouver toutes les applications u définies sur $]0,+\infty[$ de classe $\mathcal{C}^2$ qui sont des solutions de l'équation différentielle[85]

$$\forall r \in]0,+\infty[\qquad \frac{1}{r}\,u'(r) + u''(r) = 0. \tag{31}$$

2 - En utilisant la question précédente, déterminer les applications u définies sur $]0,+\infty[$ à valeurs dans $\mathbb{R}$ de classe $\mathcal{C}^2$ telles que l'application h définie sur $\mathbb{R}^2 \setminus \{(0,0)\}$ par $h(x,y) = u(\sqrt{x^2+y^2})$ soit harmonique (on pourra poser $r = \sqrt{x^2+y^2}$).

3 - Résoudre l'équation différentielle suivante[86] sur $\mathbb{R}$:

$$2t\,v'(t) + (1+t^2)\,v''(t) = 0. \tag{32}$$

4 - En utilisant la question précédente, déterminer les applications v de $\mathbb{R}$ dans $\mathbb{R}$ de classe $\mathcal{C}^2$ telles que l'application k, définie sur $\mathbb{R}^2 \setminus \{(0,y),\ y \in \mathbb{R}\}$ par la relation $k(x,y) = v\,(y/x)$ soit harmonique.

PARTIE 2 - LE PRINCIPE DU MAXIMUM

Soit f une application harmonique définie dans $\mathbb{R}^2$. Soient D le disque fermé de centre $(0,0)$ et de rayon $r \in \mathbb{R}_+^*$ et C le cercle de centre $(0,0)$ et de rayon r :

$$D = \{(x,y) \in \mathbb{R}^2 \ ;\ x^2+y^2 \leqslant r^2\} \quad \text{et} \quad C = \{(x,y) \in \mathbb{R}^2 \ ;\ x^2+y^2 = r^2\}.$$

Étant donné un entier $n \in \mathbb{N}^*$, on note f_n de $\mathbb{R}^2$ dans $\mathbb{R}$ définie par la relation

$$f_n(x,y) = f(x,y) + \frac{x^2+y^2}{n}.$$

---

[85] Indication : penser à la relation $(ru'(r))' = u'(r) + ru''(r)$.

[86] Indication : comme pour la première question, faire apparaître la dérivée d'un produit.

5 - Démontrer qu'il existe un point (a_n, b_n), appartenant au disque fermé D, en lequel l'application f_n atteint son maximum.

6 - Démontrer que, si le point (a_n, b_n) appartient à l'intérieur du disque D alors[87]

$$\frac{\partial^2 f_n}{\partial x^2}(a_n, b_n) \leqslant 0 \quad \text{et} \quad \frac{\partial^2 f_n}{\partial y^2}(a_n, b_n) \leqslant 0.$$

7 - En calculant le laplacien de l'application f_n, en déduire que le point (a_n, b_n) est situé sur le cercle C.

8 - Démontrer qu'il existe un point (a, b) du cercle C en lequel la fonction f atteint son maximum sur D.

9 - En déduire que si deux applications sont harmoniques dans le plan $\mathbb{R}^2$ et égales le long d'un cercle $\widetilde{C}$ du plan alors ces deux applications sont égales dans tout le disque $\widetilde{D}$ de frontière $\widetilde{C}$.

Exercice 10.15

Sur l'espace vectoriel $\mathcal{M}_2(\mathbb{R})$ des matrices carrées d'ordre 2 à coefficients réels, on considère la norme $\|\cdot\|$ définie pour $\mathrm{A} = (a_{i,j})_{1 \leqslant i,j \leqslant 2}$ par $\|\mathrm{A}\| = \sum_{1 \leqslant i,j \leqslant 2} |a_{i,j}|$.
Étant donné les deux matrices

$$\mathrm{A}_0 = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \quad \text{et} \quad \mathrm{B}_0 = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix},$$

on se propose de résoudre dans $\mathcal{M}_2(\mathbb{R})$ l'équation $\mathrm{X}^2 = \mathrm{B}$ pour B proche de B_0.

1 - Calculer la différentielle de l'application $\mathrm{X} \in \mathcal{M}_2(\mathbb{R}) \mapsto \mathrm{X}^2 \in \mathcal{M}_2(\mathbb{R})$ au point A_0.

2 - Prouver qu'il existe $\alpha \in \mathbb{R}_+^*$ et $\beta \in \mathbb{R}_+^*$ tels que si $\|\mathrm{B} - \mathrm{B}_0\| \leqslant \alpha$ alors il existe une unique matrice $\mathrm{X} \in \mathcal{M}_2(\mathbb{R})$ vérifiant $\|\mathrm{X} - \mathrm{A}_0\| \leqslant \beta$ et $\mathrm{X}^2 = \mathrm{B}$.

3 - Calculer le développement limité à l'ordre 2 du coefficient x_{11} de X en fonction de ceux de B, supposée proche de B_0.

---

[87] Indication : on pourra utiliser le fait que si une application $g \in \mathcal{C}^2(\mathbb{R}, \mathbb{R})$ atteint un maximum en un point $x_0 \in \mathbb{R}$ alors on a $g''(x_0) \leqslant 0$.

10.7 Solutions des exercices

Solution de l'exercice 10.1

1 - Considérons une application bilinéaire $B : E \times E \to \mathbb{R}$. Pour $(\mathbf{a}, \mathbf{b}) \in E \times E$ et $(\mathbf{h}, \mathbf{k}) \in E \times E$, on obtient par bilinéarité de B :

$$B((\mathbf{a}, \mathbf{b}) + (\mathbf{h}, \mathbf{k})) = B(\mathbf{a} + \mathbf{h}, \mathbf{b} + \mathbf{k}) = B(\mathbf{a}, \mathbf{b}) + B(\mathbf{a}, \mathbf{k}) + B(\mathbf{h}, \mathbf{b}) + B(\mathbf{h}, \mathbf{k}).$$

Vérifions que la différentielle de l'application B au point $(\mathbf{a}, \mathbf{b})$ est donnée par

$$\mathrm{d}B_{(\mathbf{a},\mathbf{b})} : (\mathbf{h}, \mathbf{k}) \in E \times E \longmapsto B(\mathbf{a}, \mathbf{k}) + B(\mathbf{h}, \mathbf{b}) \in \mathbb{R}.$$

Cette application est linéaire car l'application bilinéaire B est linéaire par rapport à chacune de ses variables. Elle est par conséquent continue puisque $E \times E$ est de dimension finie. Il reste à montrer que le terme $B(\mathbf{h}, \mathbf{k})$ est négligeable devant $\|(\mathbf{h}, \mathbf{k})\|_{E\times E} = \sqrt{\|\mathbf{h}\|_E^2 + \|\mathbf{k}\|_E^2}$ au voisinage de $(\mathbf{0}, \mathbf{0})$. Pour cela, montrons qu'il existe une constante $C > 0$ telle que pour tout $(\mathbf{h}, \mathbf{k}) \in E^2$ on a

$$|B(\mathbf{h}, \mathbf{k})| \leqslant C\|\mathbf{h}\|_E \, \|\mathbf{k}\|_E. \tag{33}$$

En utilisant l'inégalité $ab \leqslant \frac{1}{2}(a^2 + b^2)$, on en déduira alors que

$$|B(\mathbf{h}, \mathbf{k})| \leqslant \tfrac{1}{2}C \, \|(\mathbf{h}, \mathbf{k})\|_{E\times E}^2 \quad \text{et comme} \quad \|(\mathbf{h}, \mathbf{k})\|_{E\times E}^2 = \mathcal{O}_{(\mathbf{0},\mathbf{0})}(\|(\mathbf{h}, \mathbf{k})\|_{E\times E})$$

on pourra conclure.

Considérons une base $\mathcal{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ de E et notons C' le réel

$$C' = \max_{(i,j)\in\{1,\ldots,n\}^2} |B(\mathbf{e}_i, \mathbf{e}_j)|.$$

Les vecteurs $\mathbf{h}$ et $\mathbf{k}$ se décomposent dans la base $\mathcal{B}$ sous la forme : $\mathbf{h} = \sum_{i=1}^n h_i \, \mathbf{e}_i$ et $\mathbf{k} = \sum_{j=1}^n k_j \, \mathbf{e}_j$. On peut choisir de munir E de la norme 2 et on a dans ce cas

$$\|\mathbf{h}\|_E = \Big(\sum_{i=1}^n |h_i|^2\Big)^{\frac{1}{2}} \quad \text{et} \quad \|\mathbf{k}\|_E = \Big(\sum_{i=1}^n |k_i|^2\Big)^{\frac{1}{2}}.$$

En utilisant le fait que B est linéaire par rapport à chacune de ses variables, on obtient :

$$|B(\mathbf{h}, \mathbf{k})| = \Big|\sum_{i=1}^n \sum_{j=1}^n h_i k_j B(\mathbf{e}_i, \mathbf{e}_j)\Big| \leqslant C' \Big|\sum_{i=1}^n \sum_{j=1}^n h_i k_j\Big|.$$

En utilisant l'inégalité de Cauchy-Schwarz[88], on en déduit que

$$|B(\mathbf{h}, \mathbf{k})| \leqslant C'\Big(\sum_{i=1}^n \sum_{j=1}^n h_i^2\Big)^{\frac{1}{2}}\Big(\sum_{i=1}^n \sum_{j=1}^n k_j^2\Big)^{\frac{1}{2}} = C'\Big(\sum_{i=1}^n n h_i^2\Big)^{\frac{1}{2}}\Big(n\sum_{j=1}^n k_j^2\Big)^{\frac{1}{2}}$$

$$\leqslant C' n\Big(\sum_{i=1}^n h_i^2\Big)^{\frac{1}{2}}\Big(\sum_{j=1}^n k_j^2\Big)^{\frac{1}{2}} = C' n \|\mathbf{h}\|_E \, \|\mathbf{k}\|_E.$$

[88] Voir le théorème 8.5 p. 481 et les exemples de la p. 482.

L'inégalité (33) est établie avec $C = C'n$.

2 - Soit q une forme quadratique sur E ; il existe une forme bilinéaire symétrique $B : E \times E \mapsto \mathbb{R}$ telle que

$$\forall \mathbf{a} \in E \quad q(\mathbf{a}) = B(\mathbf{a}, \mathbf{a}).$$

Considérons $\mathbf{a} \in E$ et $\mathbf{h} \in E$. On a

$$\begin{aligned} q(\mathbf{a}+\mathbf{h}) &= B(\mathbf{a}+\mathbf{h}, \mathbf{a}+\mathbf{h}) = B(\mathbf{a},\mathbf{a}) + B(\mathbf{a},\mathbf{h}) + B(\mathbf{h},\mathbf{a}) + B(\mathbf{h},\mathbf{h}) \\ &= q(\mathbf{a}) + 2B(\mathbf{a},\mathbf{h}) + q(\mathbf{h}). \end{aligned}$$

L'application $\mathbf{h} \in E \mapsto 2B(\mathbf{a},\mathbf{h})$ est linéaire et d'après l'inégalité (33), on a $q(\mathbf{h}) = \|\mathbf{h}\|_E\, \varepsilon(\mathbf{h})$ où $\lim_{\mathbf{h}\to\mathbf{0}} \varepsilon(\mathbf{h}) = 0$. Par conséquent la différentielle de la forme quadratique q en $\mathbf{a}$ est l'application linéaire $\mathrm{d}q_{\mathbf{a}} : \mathbf{h} \in E \mapsto 2B(\mathbf{a},\mathbf{h})$.

Solution de l'exercice 10.2

1 - L'application trace qui à une matrice de $\mathcal{M}_n(\mathbb{R})$ associe sa trace (i.e. la somme de ses éléments diagonaux) est une application continue de $\mathcal{M}_n(\mathbb{R})$ (muni d'une norme quelconque) dans $\mathbb{R}$ (muni de la norme définie par la valeur absolue), voir p. 576. L'ensemble $U = \{\mathrm{M} \in \mathcal{M}_n(\mathbb{R}) \ ; \ \mathrm{Tr}(\mathrm{M}) \neq 0\}$ est donc un ouvert de $\mathcal{M}_n(\mathbb{R})$ puisque[89] c'est l'image réciproque de l'ensemble $\mathbb{R}^*$ qui est ouvert dans $\mathbb{R}$ par l'application continue Tr.

2 - L'application ξ est la composée de l'application trace Tr et de la fonction $g : x \in \mathbb{R}^* \mapsto 1/x$; schématiquement on a :

$$\xi : \mathrm{M} \in U \xrightarrow{\ \mathrm{Tr}\ } \mathrm{Tr}(\mathrm{M}) \in \mathbb{R}^* \xrightarrow{\ g\ } \frac{1}{\mathrm{Tr}(\mathrm{M})} \in \mathbb{R}^*.$$

Comme g est une application réelle de la variable réelle, sa différentielle en un point $x \in \mathbb{R}^*$ est $\mathrm{d}g_x : h \mapsto g'(x)\,h$, soit $\mathrm{d}g_x : h \mapsto -h/x^2$, voir la proposition 10.3. Par ailleurs, l'application trace étant linéaire, d'après la proposition 10.5 sa différentielle en tout point $M \in \mathcal{M}_n(\mathbb{R})$ est égale à elle-même et vaut :

$$\mathrm{dTr}_{\mathrm{M}} : \mathrm{H} \in \mathcal{M}_n(\mathbb{R}) \longmapsto \mathrm{Tr}(H).$$

La proposition 10.6 indique que ξ est différentiable sur U et que pour tout $M \in U$ on a $\mathrm{d}\xi_M = \mathrm{d}g_{\mathrm{Tr}(M)} \circ \mathrm{d}f_M$, autrement dit pour tout $H \in \mathcal{M}_n(\mathbb{R})$ on a

$$\mathrm{d}\xi_M(H) = \mathrm{d}g_{\mathrm{Tr}(M)}\,(\mathrm{d}f_M(H)) = \mathrm{d}g_{\mathrm{Tr}(M)}(\mathrm{Tr}(H)) = -\frac{\mathrm{Tr}(H)}{\mathrm{Tr}(M)^2}.$$

On a donc $\mathrm{d}\xi : M \in U \longmapsto \mathrm{d}\xi_M \in \mathcal{L}(\mathcal{M}_n(\mathbb{R}), \mathbb{R})$ où

$$\mathrm{d}\xi_M : H \in \mathcal{M}_n(\mathbb{R}) \longmapsto -\frac{\mathrm{Tr}(H)}{\mathrm{Tr}(M)^2} \in \mathbb{R}.$$

---

[89] Voir la proposition 9.17 p. 561.

C'est l'expression souhaitée.

Solution de l'exercice 10.3

1 - Remarquons que pour tout $\mathbf{a} = (a_1, a_2) \in \mathbb{R}^2$ et pour tout réel t, on a

$$f(\mathbf{a} + t\mathbf{v}_1) = \sin(a_1 - t + a_2 + t) = \sin(a_1 + a_2) = f(\mathbf{a})$$

de sorte que le taux d'accroissement $(f(\mathbf{a} + t\mathbf{v}_1) - f(\mathbf{a}))/t$ tend vers 0 lorsque t tend vers 0. On a donc $\mathrm{d}_{\mathbf{v}_1} f_{\mathbf{a}} = 0$.
Pour calculer la dérivée de f dans la direction $\mathbf{v}_2$, calculons la limite du taux d'accroissement de f en $\mathbf{a}$ dans la direction de $\mathbf{v}_2$. On a :

$$\begin{aligned}\lim_{t\to 0} \frac{f(\mathbf{a}+t\mathbf{v}_2) - f(\mathbf{a})}{t} &= \lim_{t\to 0} \frac{\sin(a_1 + a_2 + 2t) - \sin(a_1 + a_2)}{t} \\ &= 2 \lim_{t\to 0} \frac{\sin(a_1 + a_2 + 2t) - \sin(a_1 + a_2)}{2t} \\ &= 2 \lim_{s\to 0} \frac{\sin(a_1 + a_2 + s) - \sin(a_1 + a_2)}{s}.\end{aligned}$$

Cette dernière limite s'interprète comme la limite du taux d'accroissement de la fonction sinus en $a_1 + a_2$. Elle est donc égale à la dérivée de la fonction sinus en $a_1 + a_2$, c'est-à-dire à $\cos(a_1 + a_2)$. On en déduit que la dérivée de l'application f dans la direction $\mathbf{v}_2$ existe en tout point $\mathbf{a} \in \mathbb{R}^2$ et on a $\mathrm{d}_{\mathbf{v}_2} f_{\mathbf{a}} = 2\cos(a_1 + a_2)$.

2 - Soient $\mathbf{a} = (a_1, a_2) \in \mathbb{R}^2$ et $\mathbf{v} = (v_1, v_2) \in \mathbb{R}^2 \setminus \{(0,0)\}$. Étudions la limite du taux d'accroissement de f en $\mathbf{a}$ dans la direction de $\mathbf{v}$:

$$\lim_{t\to 0} \frac{f(\mathbf{a}+t\mathbf{v}) - f(\mathbf{a})}{t} = \lim_{t\to 0} \frac{g(a_1 + a_2 + t(v_1 + v_2)) - g(a_1 + a_2)}{t}.$$

Lorsque $v_1 + v_2 = 0$ le taux d'accroissement est nul. Supposons que $v_1 + v_2 \neq 0$; on a

$$\begin{aligned}\lim_{t\to 0} \frac{f(\mathbf{a}+t\mathbf{v}) - f(\mathbf{a})}{t} &= (v_1 + v_2) \lim_{t\to 0} \frac{g(a_1 + a_2 + t(v_1 + v_2)) - g(a_1 + a_2)}{t(v_1 + v_2)} \\ &= (v_1 + v_2) \lim_{s\to 0} \frac{g(a_1 + a_2 + s) - g(a_1 + a_2)}{s}\end{aligned}$$

Cette dernière limite s'interprète comme la limite du taux d'accroissement de la fonction g en $a_1 + a_2$. Elle est donc égale à la dérivée de la fonction g en $a_1 + a_2$. On en déduit que la dérivée de f dans la direction $\mathbf{v}$ en $\mathbf{a} = (a_1, a_2) \in \mathbb{R}^2$ existe et vaut $\mathrm{d}_{\mathbf{v}} f_{\mathbf{a}} = (v_1 + v_2)\, g'(a_1 + a_2)$.

Solution de l'exercice 10.4

Comme dans l'exemple de la p. 630, pour calculer la dérivée selon le premier vecteur de la base canonique de $\mathbb{R}^3$, on calcule la limite lorsque t tend vers 0 du taux d'accroissement

$$\frac{f(x_1 + t, x_2, x_3) - f(x_1, x_2, x_3)}{t} = (\cos(x_2)\sin(x_3), \sin(x_2)\sin(x_3), \cos(x_3)).$$

Le taux d'accroissement étant constant, on obtient

$$\frac{\partial f}{\partial x_1}(x_1,x_2,x_3) = (\cos(x_2)\sin(x_3), \sin(x_2)\sin(x_3), \cos(x_3)).$$

Les deux autres dérivées partielles s'obtiennent sur le même principe ; elles valent

$$\frac{\partial f}{\partial x_2}(x_1,x_2,x_3) = (-x_1\sin(x_2)\sin(x_3), x_1\cos(x_2)\sin(x_3), 0),$$
$$\frac{\partial f}{\partial x_3}(x_1,x_2,x_3) = (x_1\cos(x_2)\cos(x_3), x_1\sin(x_2)\cos(x_3), -x_1\sin(x_3)).$$

Solution de l'exercice 10.5

1 - Pour calculer la première dérivée partielle de $f : (x_1,x_2) \in \mathbb{R}_*^2 \mapsto x_2\frac{x_1^2-x_2^2}{x_1^2+x_2^2}$, on dérive l'expression de $f(x_1,x_2)$ par rapport à x_1 en considérant que x_2 est une constante. On obtient

$$\frac{\partial f}{\partial x_1}(x_1,x_2) = \frac{4\,{x_2}^3 x_1}{\left({x_1}^2+{x_2}^2\right)^2}.$$

Pour calculer la seconde dérivée partielle de f, on dérive l'expression de $f(x_1,x_2)$ par rapport à x_2 en considérant que x_1 est une constante. On obtient

$$\frac{\partial f}{\partial x_2}(x_1,x_2) = \frac{x_1^2-x_2^2}{x_1^2+x_2^2} - \frac{4x_1^2x_2^2}{(x_1^2+x_2^2)^2} = \frac{{x_1}^4-{x_2}^4-4\,{x_1}^2{x_2}^2}{\left({x_1}^2+{x_2}^2\right)^2}.$$

2 - Pour calculer la première dérivée partielle de $f : (x_1,x_2) \in \mathbb{R}_*^2 \mapsto \frac{4x_1^2x_2^3}{(x_1^2+x_2^2)^2}$, on dérive l'expression de $f(x_1,x_2)$ par rapport à x_1 en considérant que x_2 est une constante. On obtient

$$\frac{\partial f}{\partial x_1}(x_1,x_2) = -\frac{8\,x_1\,{x_2}^3\left({x_1}^2-{x_2}^2\right)}{\left({x_1}^2+{x_2}^2\right)^3}.$$

Pour calculer la seconde dérivée partielle de f, on dérive l'expression de $f(x_1,x_2)$ par rapport à x_2 en considérant que x_1 est une constante. On obtient

$$\frac{\partial f}{\partial x_2}(x_1,x_2) = \frac{4\,{x_1}^2{x_2}^2\left(3\,{x_1}^2-{x_2}^2\right)}{\left({x_1}^2+{x_2}^2\right)^3}.$$

Solution de l'exercice 10.6

D'après la définition 10.6, les dérivées partielles dépendent de la base de l'espace de départ considérée. Si on a choisi pour base de $E = \mathbb{R}_2[X]$ la famille $\mathcal{B} = \{1, 1+X, 1+X^2\}$ alors les trois dérivées partielles de f en $Q = 1+X^2$ sont les limites lorsque t tend vers 0 des taux d'accroissement suivants :

$$\frac{f(Q+t)-f(Q)}{t}, \quad \frac{f(Q+t(1+X))-f(Q)}{t} \quad \text{et} \quad \frac{f(Q+t(1+X^2))-f(Q)}{t}.$$

On a $f(Q) = f(Q+t) = -\cos(1)X^2$, $f(Q+t(1+X)) = \sin(t(1+t))X - \cos(1)X^2$ et $f((Q+t(1+X^2)) = -\cos(1+t)X^2$. On en déduit que pour tout $t \in \mathbb{R}^*$,

$$\begin{aligned}
&\frac{f(Q+t)-f(Q)}{t} = 0,\\
&\frac{f(Q+t(1+X))-f(Q)}{t} = \frac{1}{t}\sin(t(1+t))\,X,\\
&\frac{f(Q+t(1+X^2))-f(Q)}{t} = \frac{1}{t}\big(\cos(1)-\cos(1+t)\big)\,X^2.
\end{aligned}$$

On a $\sin(t(1+t)) \underset{0}{\sim} \sin(t) \underset{0}{\sim} t$ et $\cos(1+t) = \cos(1) - t\sin(1) + o_0(t^2)$. On en déduit que

$$\lim_{t\to 0}\frac{1}{t}\sin(t(1+t)) = 1 \quad \text{et} \quad \lim_{t\to 0}\frac{1}{t}\big(\cos(1)-\cos(1+t)\big) = \sin(1).$$

On obtient finalement

$$\partial_1 f(Q) = 0, \quad \partial_2 f(Q) = X \quad \text{et} \quad \partial_3 f(Q) = \sin(1)\,X^2.$$

Solution de l'exercice 10.7

D'après l'exercice 10.4, les dérivées partielles de f en $(x_1,x_2,x_3) \in \mathbb{R}^3$ sont données par

$$\begin{cases}
\dfrac{\partial f}{\partial x_1}(x_1,x_2,x_3) = (\cos(x_2)\sin(x_3), \sin(x_2)\sin(x_3), \cos(x_3)),\\
\dfrac{\partial f}{\partial x_2}(x_1,x_2,x_3) = (-x_1\sin(x_2)\sin(x_3), x_1\cos(x_2)\sin(x_3), 0),\\
\dfrac{\partial f}{\partial x_3}(x_1,x_2,x_3) = (x_1\cos(x_2)\cos(x_3), x_1\sin(x_2)\cos(x_3), -x_1\sin(x_3)).
\end{cases}$$

La matrice jacobienne de f en $(x_1,x_2,x_3) \in \mathbb{R}^3$ est donc

$$\mathrm{J}_f(x_1,x_2,x_3) = \begin{pmatrix}
\cos(x_2)\sin(x_3) & -x_1\sin(x_2)\sin(x_3) & x_1\cos(x_2)\cos(x_3)\\
\sin(x_2)\sin(x_3) & x_1\cos(x_2)\sin(x_3) & x_1\sin(x_2)\cos(x_3)\\
\cos(x_3) & 0 & -x_1\sin(x_3)
\end{pmatrix}$$

et le jacobien de f en $(x_1,x_2,x_3) \in \mathbb{R}^3$ qui est le déterminant de cette matrice jacobienne vaut $\mathcal{J}_f(x_1,x_2,x_3) = -x_1\sin(x_3)$.

Solution de l'exercice 10.8

Les dérivées partielles de l'application composée $\varphi = g \circ f$ peuvent être obtenues en utilisant soit la formule (9) de la p. 637 soit la formule (10) de la p. 638. Cette dernière s'écrit

$$\begin{aligned}\left(\frac{\partial\varphi}{\partial x_1}(x_1,x_2,x_3) \quad \frac{\partial\varphi}{\partial x_2}(x_1,x_2,x_3) \quad \frac{\partial\varphi}{\partial x_3}(x_1,x_2,x_3)\right) &= \\ = \left(\partial_1 g(f(x_1,x_2,x_3)) \quad \partial_2 g(f(x_1,x_2,x_3)) \quad \partial_3 g(f(x_1,x_2,x_3))\right) & \\ \times \begin{pmatrix} \cos(x_2)\sin(x_3) & -x_1\sin(x_2)\sin(x_3) & x_1\cos(x_2)\cos(x_3) \\ \sin(x_2)\sin(x_3) & x_1\cos(x_2)\sin(x_3) & x_1\sin(x_2)\cos(x_3) \\ \cos(x_3) & 0 & -x_1\sin(x_3) \end{pmatrix}&,\end{aligned}$$

les dérivées partielles de l'application f ayant été calculées au cours des exercices 10.4 et 10.7. Après avoir effectué le produit matriciel du membre de droite, on obtient par identification toutes les dérivées partielles de φ

$$\begin{aligned}\frac{\partial\varphi}{\partial x_1}(x_1,x_2,x_3) &= \cos(x_2)\sin(x_3)\ \partial_1 g(f(x_1,x_2,x_3)) \\ &\quad + \sin(x_2)\sin(x_3)\ \partial_2 g(f(x_1,x_2,x_3)) + \cos(x_3)\ \partial_3 g(f(x_1,x_2,x_3)) \\ \frac{\partial\varphi}{\partial x_2}(x_1,x_2,x_3) &= -x_1\sin(x_2)\sin(x_3)\ \partial_1 g(f(x_1,x_2,x_3)) \\ &\quad + x_1\cos(x_2)\sin(x_3)\ \partial_2 g(f(x_1,x_2,x_3)) \\ \frac{\partial\varphi}{\partial x_3}(x_1,x_2,x_3) &= x_1\cos(x_2)\cos(x_3)\ \partial_1 g(f(x_1,x_2,x_3)) \\ &\quad + x_1\sin(x_2)\cos(x_3)\ \partial_2 g(f(x_1,x_2,x_3)) - x_1\sin(x_3)\ \partial_3 g(f(x_1,x_2,x_3)).\end{aligned}$$

Solution de l'exercice 10.9

1 - On peut obtenir les dérivées partielles de f en $(x_1,x_2,x_3) \in \mathbb{R}^2 \times \mathbb{R}^*$ en dérivant l'expression x_1x_2/x_3^2 successivement par rapport à chacune des 3 variables en considérant à chaque fois les deux autres comme étant des constantes. On obtient

$$\frac{\partial f}{\partial x_1}(x_1,x_2,x_3) = \frac{x_2}{x_3^2}, \quad \frac{\partial f}{\partial x_2}(x_1,x_2,x_3) = \frac{x_1}{x_3^2}, \quad \frac{\partial f}{\partial x_3}(x_1,x_2,x_3) = -\frac{2x_1x_2}{x_3^3}.$$

On en déduit que $\dfrac{\partial f}{\partial x_1}(1,1,1) = 1$, $\dfrac{\partial f}{\partial x_2}(1,1,1) = 1$ et $\dfrac{\partial f}{\partial x_3}(1,1,1) = -2$.

2 - Les trois dérivées partielles de f sont des applications continues sur l'ouvert $U = \mathbb{R}^2 \times \mathbb{R}^*$ car il s'agit du quotient d'applications continues sur $\mathbb{R}^3$ dont le dénominateur ne s'annule pas sur U. D'après le théorème 10.2, on en déduit que f est de classe $\mathcal{C}^1$ sur U. Elle est donc différentiable sur U et en particulier en $(1,1,1)$. On déduit de la question précédente et de la proposition 10.9 que

$$\mathrm{d}f_{(1,1,1)} : (h_1,h_2,h_3) \mapsto h_1 + h_2 - 2h_3.$$

3 - La fonction f n'est pas continue à l'origine car en restriction à la courbe paramétrée de l'espace dont un paramétrage est $\varphi : t \in \mathbb{R} \mapsto (\sqrt{t}, \sqrt{t}, t)$ et pour laquelle $(0,0,0) = \varphi(0)$, la fonction f vaut $f(\varphi(t)) = 1/t$ et cette quantité n'a pas de limite. Comme f n'est pas continue à l'origine, d'après la proposition 10.2, f n'est pas différentiable à l'origine.

Solution de l'exercice 10.10

Notons U l'ouvert $\mathbb{R}^* \times R$.

1 - On obtient la première dérivée partielle de f en $(x_1, x_2) \in U$ en dérivant la quantité $x_1^2 \sin(x_2/x_1)$ par rapport à la variable x_1 en supposant x_2 fixée. On obtient

$$\frac{\partial f}{\partial x_1}(x_1, x_2) = 2x_1 \sin\left(\frac{x_2}{x_1}\right) - x_2 \cos\left(\frac{x_2}{x_1}\right).$$

Considérons à présent le cas où $(x_1, x_2) \notin U$, autrement dit un point de la forme $(0, x_2)$ avec $x_2 \in \mathbb{R}$. On a

$$\lim_{h \to 0} \frac{f(h, x_2) - f(0, x_2)}{h} = \lim_{h \to 0} h \sin\left(\frac{x_2}{h}\right) = 0,$$

ce qui prouve que la dérivée partielle de f selon le premier vecteur de la base canonique de $\mathbb{R}^2$ existe en $(0, x_2)$ et vaut 0. On en conclut que la première dérivée partielle d'ordre 1 de f est :

$$\frac{\partial f}{\partial x_1} : (x_1, x_2) \in \mathbb{R}^2 \longmapsto \begin{cases} 2x_1 \sin\left(\dfrac{x_2}{x_1}\right) - x_2 \cos\left(\dfrac{x_2}{x_1}\right) & \text{lorsque } x_1 \neq 0 \\ 0 & \text{lorsque } x_1 = 0 \end{cases}.$$

De la même manière, on montre que la seconde dérivée partielle de f est l'application

$$\frac{\partial f}{\partial x_2} : (x_1, x_2) \in \mathbb{R}^2 \longmapsto \begin{cases} x_1 \cos\left(\dfrac{x_2}{x_1}\right) & \text{lorsque } x_1 \neq 0 \\ 0 & \text{lorsque } x_1 = 0 \end{cases}.$$

2 - Évaluons la dérivée partielle seconde $\dfrac{\partial^2 f}{\partial x_2 \partial x_1}$ en $(0,0)$; on a $\dfrac{\partial^2 f}{\partial x_2 \partial x_1} = \dfrac{\partial}{\partial x_2}\left(\dfrac{\partial f}{\partial x_1}\right)$ et

$$\lim_{h \to 0} \frac{\frac{\partial f}{\partial x_1}(0, h) - \frac{\partial f}{\partial x_1}(0, 0)}{h} = \lim_{h \to 0} \frac{0 - 0}{h} = 0.$$

On en déduit que $\dfrac{\partial^2 f}{\partial x_2 \partial x_1}(0, 0) = 0$. De même, on a

$$\lim_{h \to 0} \frac{\frac{\partial f}{\partial x_2}(h, 0) - \frac{\partial f}{\partial x_2}(0, 0)}{h} = \lim_{h \to 0} \frac{h \cos(0/h) - 0}{h} = 1,$$

d'où $\dfrac{\partial^2 f}{\partial x_1 \partial x_2}(0,0) = 1$.

Solution de l'exercice 10.11

1 - L'application f n'est pas injective sur $\mathbb{R}^3$ car $f(1,0,0) = (0,0,1) = f(1,2\pi,0)$. L'application $f : \mathbb{R}^3 \to \mathbb{R}^3$ ne peut donc pas être un difféomorphisme.

2 - Notons $U = \mathbb{R}_+^* \times]0, 2\pi[\times]-\frac{\pi}{2}, \frac{\pi}{2}[$. Vérifions que l'application f est un $\mathcal{C}^1$-difféomorphisme de U sur $f(U)$.

– L'application f est une application de classe $\mathcal{C}^1$ sur $\mathbb{R}^3$ (c'est le produit d'applications de classe $\mathcal{C}^1$ sur $\mathbb{R}^3$), elle est donc aussi de classe $\mathcal{C}^1$ sur U.

– Le jacobien de f a été calculé lors de l'exercice 10.7. On a montré que $\mathcal{J}_f(x_1,x_2,x_3) = -x_1 \sin(x_3)$; il ne s'annule pas pour $(x_1,x_2,x_3) \in U$.

– L'application f est injective sur U. En effet, supposons que $f(\mathbf{x}) = f(\mathbf{y})$ où $\mathbf{x} = (x_1,x_2,x_3) \in U$ et $\mathbf{y} = (y_1,y_2,y_3) \in U$; on a

$$\begin{cases} x_1 \cos(x_2)\sin(x_3) = y_1 \cos(y_2)\sin(y_3), \\ x_1 \sin(x_2)\sin(x_3) = y_1 \sin(y_2)\sin(y_3), \\ x_1 \cos(x_3) = y_1 \cos(y_3). \end{cases} \tag{34}$$

En effectuant la somme du carré de chacune de ces trois égalités, on obtient l'égalité $x_1^2 = y_1^2$. Les deux réels x_1 et y_1 étant strictement positifs, on en déduit $x_1 = y_1$. La dernière équation du système (34) devient $\cos(x_3) = \cos(y_3)$ mais puisque x_3 et y_3 sont dans l'intervalle $]-\pi/2, \pi/2[$, on a nécessairement $x_3 = y_3$. En reportant les égalités $x_1 = y_1$ et $x_3 = y_3$ dans les deux premières équations du système (34), on obtient $\cos(x_2) = \cos(y_2)$ et $\sin(x_2) = \sin(y_2)$. Les réels x_2 et y_2 appartenant tous les deux à l'intervalle $]0, 2\pi[$, on a $x_2 = y_2$. Finalement on en conclut que $\mathbf{x} = \mathbf{y}$ et donc que la fonction f est injective.

Solution de l'exercice 10.12

L'application f de classe $\mathcal{C}^\infty$ sur $]-1,1[^2$ car c'est un quotient de deux applications de classe $\mathcal{C}^\infty$ sur $]-1,1[^2$ dont la fonction au dénominateur ne s'annule pas sur $]-1,1[^2$. Les dérivées partielles premières de f en $(x_1,x_2) \in]-1,1[^2$, obtenue en dérivant l'expression de $f(x_1,x_2)$ par rapport à l'une des variables, en supposant l'autre fixée, sont :

$$\frac{\partial f}{\partial x_1}(x_1,x_2) = \frac{3}{2} \frac{x_2\,(2\,x_1+2)}{\left((x_1+1)^2 + {x_2}^2\right)^{5/2}}$$

$$\frac{\partial f}{\partial x_1}(x_1,x_2) = -\frac{{x_1}^2 + 2\,x_1 + 1 - 2\,{x_2}^2}{({x_1}^2 + 2\,x_1 + 1 + {x_2}^2)^{5/2}}$$

On en déduit que la matrice jacobienne de f en $(0,0)$ est la matrice

$$\mathrm{J}_f(0,0) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}.$$

Les dérivées partielles secondes de f $(x_1, x_2) \in]-1,1[^2$ sont

$$\frac{\partial^2 f}{\partial x_1^2}(x_1,x_2) = -3\,\frac{x_2\,\left(4\,{x_1}^2+8\,x_1+4-{x_2}^2\right)}{\left({x_1}^2+2\,x_1+1+{x_2}^2\right)^{7/2}}$$

$$\frac{\partial^2 f}{\partial x_2^2}(x_1,x_2) = 3\,\frac{x_2\,\left(3\,{x_1}^2+6\,x_1+3-2\,{x_2}^2\right)}{\left({x_1}^2+2\,x_1+1+{x_2}^2\right)^{7/2}}$$

$$\frac{\partial^2 f}{\partial x_1 x_2}(x_1,x_2) = \frac{\partial^2 f}{\partial x_2 x_1}(x_1,x_2) = 3\,\frac{(x_1+1)\left({x_1}^2+2\,x_1+1-4\,{x_2}^2\right)}{\left({x_1}^2+2\,x_1+1+{x_2}^2\right)^{7/2}}$$

On en déduit que la matrice hessienne de f en $(0,0)$ est :

$$\mathrm{H}_f(0,0) = \begin{pmatrix} 0 & 3 \\ 3 & 0 \end{pmatrix}.$$

La formule de Taylor à l'ordre 2 au voisinage de $(0,0)$ pour f s'écrit

$$\begin{aligned} f(h_1,h_2) &= f(0,0) + \mathrm{J}_f(0,0)\ \begin{pmatrix} h_1 & h_2 \end{pmatrix} + \frac{1}{2}\ \begin{pmatrix} h_1 \\ h_2 \end{pmatrix}\ \mathrm{H}_f(0,0)\ \begin{pmatrix} h_1 & h_2 \end{pmatrix} \\ &\qquad + \mathcal{O}_{(0,0)}(h_1^2+h_2^2) \\ &= -h_2 + 3h_1h_2 + \mathcal{O}_{(0,0)}(h_1^2+h_2^2). \end{aligned}$$

Solution de l'exercice 10.13

La matrice jacobienne de f en $(x_1,x_2) \in \mathbb{R}^2$ vaut $\mathrm{J}_f(x_1,x_2) = (2x_1 \quad 3x_2^2)$ et s'annule uniquement pour $x_1 = x_2 = 0$. L'origine est le seul point critique pour f sur $\mathbb{R}^2$. En ce point, la hessienne vaut $\mathrm{H}_f(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$ et admet une valeur propre nulle : on ne peut pas conclure grâce au corollaire 10.3.

Néanmoins, on observe que pour tout $\delta \in \mathbb{R}$ on a $f(0,\delta) = \delta^3$ et que cette quantité est tantôt positive, tantôt négative selon le signe de δ : $(0,0)$ ne peut pas être un extremum.

Solution de l'exercice 10.14

1 - Une application $u \in \mathcal{C}^2(]0,+\infty[,\mathbb{R})$ est une solution de l'équation différentielle (31) si et seulement si $(ru'(r))' = 0$ pour tout $r \in]0,+\infty[$. Puisque la dérivée de la fonction $r \mapsto ru'(r)$ est nulle et que l'ensemble $]0,+\infty[$ est un intervalle, on obtient dans un premier temps que la fonction $r \mapsto ru'(r)$ est constante :

$$\exists A \in \mathbb{R} \forall r \in]0,+\infty[\quad ru'(r) = A.$$

On en déduit ensuite que

$$\exists (A,B) \in \mathbb{R}^2\ ;\ \forall r \in]0,+\infty[\quad u(r) = A\ln(r) + B.$$

2 - Soit h une fonction de la forme $h(x,y) = u(\sqrt{x^2+y^2})$. Exprimons le laplacien de la fonction h en terme de dérivées de la fonction u. En dérivant la relation $h(x,y) = u(\sqrt{x^2+y^2})$ par rapport à la variable x, on obtient pour tout $(x,y) \in \mathbb{R}^2_*$

$$\frac{\partial h}{\partial x}(x,y) = \frac{x}{\sqrt{x^2+y^2}}\, u'(\sqrt{x^2+y^2}).$$

En dérivant à nouveau cette relation par rapport à la variable x, on obtient

$$\frac{\partial^2 h}{\partial x^2}(x,y) = \frac{1}{\sqrt{x^2+y^2}}\, u'(\sqrt{x^2+y^2}) - \frac{x^2}{(x^2+y^2)^{3/2}}\, u'(\sqrt{x^2+y^2}) + \frac{x^2}{x^2+y^2}\, u''(\sqrt{x^2+y^2}).$$

De même, on évalue la dérivée seconde de h par rapport à la variable y. En ajoutant les deux résultats, on a

$$\Delta h(x,y) = \frac{1}{\sqrt{x^2+y^2}} u'(\sqrt{x^2+y^2}) + u''(\sqrt{x^2+y^2}).$$

Ainsi, l'application h est harmonique si et seulement si l'application u est solution sur $]0,+\infty[$ de l'équation différentielle $\frac{1}{r}u'(r) + u''(r) = 0$. D'après la question précédente, la solution u de cette équation différentielle s'écrit sous la forme $u(r) = A\ln(r) + B$, A et B étant deux constantes réelles quelconques.

3 - On remarque que $((1+t^2)v'(t))' = 2tv'(t) + (1+t^2)v''(t)$. On peut donc intégrer de la même manière que lors de la première question l'équation sur v. On obtient :

$$\exists (C,D) \in \mathbb{R}^2 \ ; \ v(t) = C\arctan(t) + D.$$

4 - Soit k une fonction du type $k(x,y) = v(y/x)$ où v est une application définie sur $\mathbb{R}^2 \setminus \{(0,y),\ y \in \mathbb{R}\}$. En dérivant la relation $k(x,y) = v(y/x)$ par rapport à la variable x et par rapport à la variable y on a

$$\frac{\partial k}{\partial x}(x,y) = -\frac{y}{x^2} v'\left(\frac{y}{x}\right) \quad \text{et} \quad \frac{\partial k}{\partial y}(x,y) = \frac{1}{x} v'\left(\frac{y}{x}\right).$$

En dérivant une seconde fois, on obtient

$$\frac{\partial^2 k}{\partial x^2}(x,y) = -\frac{2y}{x^3} v'\left(\frac{y}{x}\right) + \frac{y^2}{x^4} v''\left(\frac{y}{x}\right) \quad \text{et} \quad \frac{\partial^2 k}{\partial y^2}(x,y) = \frac{1}{x^2} v''\left(\frac{y}{x}\right).$$

Ainsi pour tout $(x,y) \in \mathbb{R}^2 \setminus \{(0,y)\ ;\ y \in \mathbb{R}\}$ on a la relation

$$\Delta k(x,y) = -\frac{1}{x^2}\left(2\frac{y}{x} v'\left(\frac{y}{x}\right) + \left(1 + \frac{y^2}{x^2}\right) v''\left(\frac{y}{x}\right)\right).$$

Par conséquent l'application k est harmonique si et seulement si l'application v est solution sur $\mathbb{R}$ de l'équation différentielle $2tv'(t) + (1+t^2)v''(t) = 0$. D'après la question précédente, la solution v de cette équation différentielle s'écrit sous la forme $v(t) = C\arctan(t) + D$ où C et D sont deux constantes réelles quelconques.

5 - L'application f_n est continue sur un ensemble fermé et borné (le disque D) on sait donc[90] qu'elle atteint ses bornes, et en particulier son maximum, sur D.

6 - La dérivée seconde de f_n par rapport à la première variable au point (a_n, b_n) est[91] la dérivée seconde de la fonction $x \mapsto f_n(x, b_n)$ au point a_n. Puisque cette application atteint son maximum en a_n, on sait[92] que la dérivée seconde y est négative[93]. On procède de la même manière pour la dérivée seconde par rapport à la variable y.

7 - Puisque l'application f est harmonique, le laplacien de f_n est égal à celui de la fonction $(x, y) \mapsto \frac{1}{n}(x^2 + y^2)$, c'est-à-dire qu'il vaut $\Delta f_n = \frac{4}{n}$. Au point (a_n, b_n) on a donc une contradiction puisque d'une part $\Delta f_n(a_n, b_n) = \frac{4}{n} > 0$ et d'autre part d'après la question précédente

$$\Delta f_n(a_n, b_n) = \frac{\partial^2 f_n}{\partial x^2}(a_n, b_n) + \frac{\partial^2 f_n}{\partial y^2}(a_n, b_n) \leqslant 0.$$

On en déduit que le point (a_n, b_n) n'est pas à l'intérieur du disque D mais sur sa frontière : le cercle C.

8 - D'après la caractérisation des compacts de $\mathbb{R}^n$, voir le théorème 7.7, p. 415, un ensemble fermé et borné est un ensemble compact. Or le cercle C est un ensemble borné (puisqu'il est contenu dans le disque D) et le cercle C est un ensemble fermé[94] et par conséquent le cercle C est un ensemble compact. D'après la définition d'un ensemble compact[95], on en déduit que l'on peut extraire des deux suites $(a_n)_{n\in\mathbb{N}^*}$ et $(b_n)_{n\in\mathbb{N}^*}$ deux sous-suites $(a_{n_j})_{j\in\mathbb{N}^*}$ et $(b_{n_j})_{j\in\mathbb{N}^*}$ qui convergent et telles que les limites respectives a et b sont aussi dans C. Soit $(x, y) \in D$; pour tout $j \in \mathbb{N}^*$ on a $f_{n_j}(a_{n_j}, b_{n_j}) \geqslant f_{n_j}(x, y)$. En passant à la limite lorsque j tend vers $+\infty$, on obtient $f(a, b) \geqslant f(x, y)$ ce qui prouve que (a, b) est un maximum pour l'application f.

---

[90] Voir le théorème 9.2, p. 564.

[91] Voir la définition 10.6, p. 630.

[92] Voir le *Cours de première année*, p. 736.

[93] Attention, ceci n'est valable uniquement si la fonction f_n est définie dans un voisinage autour de (a_n, b_n), ce qui est vrai si (a_n, b_n) est à l'intérieur de disque D.

[94] On peut démontrer que le cercle C est un ensemble fermé en remarquant que $C = D \cap (\mathbb{R}^2 \setminus \mathrm{Int}\,(D))$. En effet, le disque D est un ensemble fermé, le disque ouvert $\mathrm{Int}\,(D)$ est un ensemble ouvert donc son complémentaire $\mathbb{R}^2 \setminus \mathrm{Int}\,(D)$ est un ensemble fermé et l'intersection de deux ensembles fermés est un ensemble fermé, voir la proposition 7.9, p. 375.

[95] Voir la définition 7.23, p. 413.

9 - Considérons deux applications harmoniques dans $\mathbb{R}^2$ qui sont égales le long d'un cercle $\widetilde{C}$. On va montrer que ces deux applications sont égales dans tout le disque $\widetilde{D}$ de frontière $\widetilde{C}$. Notons f la différence de ces deux applications. L'application f est harmonique et nulle sur $\widetilde{C}$. D'après les questions précédentes[96], la maximum d'une fonction harmonique sur un disque est obtenu sur la frontière du disque. On en déduit $f \leqslant 0$ sur $\widetilde{D}$. En considérant l'application $-f$, qui est aussi harmonique et nulle sur $\widetilde{C}$, on en déduit que $-f \leqslant 0$ sur $\widetilde{D}$. Cela implique que $f = 0$ sur $\widetilde{D}$.

Solution de l'exercice 10.15

1 - Notons f l'application $X \in \mathcal{M}_2(\mathbb{R}) \mapsto X^2 \in \mathcal{M}_2(\mathbb{R})$. Afin de déterminer la différentielle de l'application f en $X \in \mathcal{M}_2(\mathbb{R})$ on va[97] d'abord former, pour $H \in \mathcal{M}_2(\mathbb{R})$, la différence entre les termes $f(X+H)$ et $f(X)$ puis identifier les « termes linéaires » en la variable H. En prenant garde au fait que la multiplication des matrices n'est pas une opération commutative, on obtient

$$f(X+H) - f(X) = (X+H)^2 - X^2 = XH + HX + H^2.$$

Il est aisé de vérifier que l'application $L_X : H \in \mathcal{M}_2(\mathbb{R}) \mapsto XH + HX$ est une application linéaire. Puisque l'ensemble $\mathcal{M}_2(\mathbb{R})$ est de dimension finie, cette application linéaire est une application continue : c'est naturellement le « bon candidat » pour être la différentielle de l'application f en X. Pour le prouver, il faut montrer que le terme H^2 obtenu en effectuant la différence $f(X+H)-f(X)$ est de la forme $\|H\|\varepsilon(H)$ et que l'application ε ainsi définie tend vers 0 en 0. Posons

$$\varepsilon : H \in \mathcal{M}_2(\mathbb{R}) \longmapsto \begin{cases} \dfrac{H^2}{\|H\|} & \text{si } H \neq 0, \\ 0 & \text{si } H = 0. \end{cases}$$

La norme $\|\cdot\|$ utilisée ici peut être n'importe quelle norme sur l'ensemble $\mathcal{M}_2(\mathbb{R})$ car on sait qu'en dimension finie toutes les normes sont équivalentes. On peut choisir une des normes de la proposition 7.5 p. 365. On notera que pour de telles normes on a la propriété suivante[98]

$$\forall (C, D) \in \mathcal{M}_2(\mathbb{R}) \quad \|CD\| \leqslant \|C\| \, \|D\|.$$

On en déduit que pour tout $H \in \mathcal{M}_2(\mathbb{R})$, on a $\|\varepsilon(H)\| \leqslant \|H\|$ ce qui prouve que $\varepsilon(H)$ tend vers 0 lorsque H tend vers 0. Ainsi, la relation

$$f(X+H) - f(X) = L_X(H) + \|H\|\varepsilon(H)$$

---

[96] On peut très aisément adapter les questions précédentes au cas d'un disque $\widetilde{D}$ quelconque du plan.

[97] Voir l'exemple de la page 611 dans lequel cette méthode est appliquée pour obtenir la différentielle d'une application.

[98] Voir l'exercice de la page 366 pour une démonstration de cette propriété.

montre que f est différentiable sur $\mathcal{M}_n(\mathbb{R})$ et que $\mathrm{d}f_X = L_X$. Au point $X = A_0$ on obtient

$$\mathrm{d}f_{A_0}(H) = A_0H + HA_0.$$

2 - Ainsi en notant[99] $H = \begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix}$ on a $\mathrm{d}f_{A_0}(H) = \begin{pmatrix} 2h_{11} & 3h_{12} \\ 3h_{21} & 4h_{22} \end{pmatrix}$. On peut en déduire l'expression de la matrice de $\mathrm{d}f_{A_0}$ dans la base canonique[100] de $\mathcal{M}_2(\mathbb{R})$:

$$\mathrm{Mat}(\mathrm{d}f_{A_0}, \mathrm{can}) = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}$$

Le déterminant de cette matrice est non nul. Puisque f est de classe $\mathcal{C}^\infty$, on peut appliquer le théorème d'inversion locale : il existe un ouvert $U \subset \mathcal{M}_2(\mathbb{R})$ contenant A_0 et un ouvert $V \subset \mathcal{M}_2(\mathbb{R})$ contenant B_0 tels que la fonction f soit un $\mathcal{C}^1$-difféomorphisme de U sur V. En particulier, la fonction f est une surjection sur V :

$$\forall B \in V \quad \exists X \in U X^2 = f(X) = B.$$

Par définition d'un ensemble ouvert[101] cela implique directement le résultat.

3 - Notons g l'application réciproque de f, définie grâce à la question précédente sur la boule ouverte de centre B_0 et de rayon α ; g est de classe $\mathcal{C}^\infty$ et la formule de Taylor à l'ordre 2 s'écrit, pour tout $B \in \mathcal{M}_2(\mathbb{R})$ telle que $\|B - B_0\| \leqslant \alpha$

$$X = g(B) = g(B_0) + \mathrm{d}g_{B_0}(H) + \frac{1}{2}\mathrm{d}^2 g_{B_0}(H, H) + \|H\|^2 \varepsilon(H)$$

où $H = B - B_0$ et où ε est une application de $\mathcal{M}_2(\mathbb{R})$ dans $\mathcal{M}_2(\mathbb{R})$ de limite nulle en 0. Évaluons les différents termes du membre de droite de cette relation.
– Puisque $f(A_0) = B_0$, on a $g(B_0) = A_0$. Le coefficient en position $(1,1)$ dans la matrice A_0 vaut 1.

– D'après la proposition 10.6 de la page 619 on a $\mathrm{d}(g \circ f)_{A_0} = \mathrm{d}g_{B_0} \circ \mathrm{d}f_{A_0}$. Puisque g est la réciproque de f, on a aussi $\mathrm{d}(g \circ f)_{A_0} = \mathrm{d}(\mathrm{id})_{A_0} = \mathrm{id}$ de sorte que $\mathrm{d}g_{B_0} = (\mathrm{d}f_{A_0})^{-1}$. D'après l'expression matricielle de la différentielle $\mathrm{d}f_{A_0}$, on en déduit l'expression matricielle de $(\mathrm{d}f_{A_0})^{-1}$:

$$(\mathrm{d}f_{A_0})^{-1}(H) = \begin{pmatrix} \frac{1}{2}h_{11} & \frac{1}{3}h_{12} \\ \frac{1}{3}h_{21} & \frac{1}{4}h_{22} \end{pmatrix}.$$

---

[99] Cette notation sera conservée au cours de cet exercice.

[100] La base canonique de $\mathcal{M}_2(\mathbb{R})$ est formée des matrices

$$E_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad E_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad E_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \quad \text{et} \quad E_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

[101] Voir la définition 7.7, p. 371. L'ensemble ouvert V contenant B_0, il contient aussi une boule ouverte de centre B_0 et de rayon $\alpha \in \mathbb{R}_+^*$.

Par conséquent que le coefficient en position $(1,1)$ dans la matrice $\mathrm{d}g_{\mathrm{B}_0}(\mathrm{H})$ vaut $\frac{1}{2}h_{11}$.

– De manière générale, la différentielle de l'application g en $\mathrm{B} \in \mathcal{M}_n(\mathbb{R})$ vaut $\mathrm{d}g_{\mathrm{B}} = (\mathrm{d}f_{g(\mathrm{B})})^{-1}$. On peut ainsi écrire $\mathrm{d}g = \Phi \circ \mathrm{d}f \circ g$ où Φ désigne l'application définie sur l'ensemble E^* des éléments inversibles de $E = L(\mathcal{M}_n(\mathbb{R}), \mathcal{M}_n(\mathbb{R}))$ par $\Phi : \Lambda \in E^* \mapsto \Lambda^{-1} \in E^*$. En utilisant la proposition 10.6 de la page 619, on calcule la différentielle de $\mathrm{d}g$, c'est-à-dire $\mathrm{d}^2 g$ comme la différentielle d'applications composées. On a $\mathrm{d}^2 g_{\mathrm{B}_0} = \mathrm{d}\Phi_{\mathrm{d}f_{g(\mathrm{B}_0)}} \circ \mathrm{d}^2 f_{g(\mathrm{B}_0)} \circ \mathrm{d}g_{\mathrm{B}_0}$, c'est-à-dire $\mathrm{d}^2 g_{\mathrm{B}_0} = \mathrm{d}\Phi_{\mathrm{d}f_{\mathrm{A}_0}} \circ \mathrm{d}^2 f_{\mathrm{A}_0} \circ \mathrm{d}g_{\mathrm{B}_0}$.

Remarquons tout d'abord[102] que l'application Φ est différentiable et que pour tout $\Lambda \in E^*$ on a

$$\mathrm{d}\Phi_\Lambda : \Psi \in E \longmapsto -\Lambda^{-1}\Psi\Lambda \in E.$$

Calculons maintenant la différentielle $\mathrm{d}^2 f_{\mathrm{A}_0}$. Cette différentielle seconde est la différentielle de l'application $F = \mathrm{d}f$ évaluée en A_0. Puisque l'application F a été déterminée lors de la première question, on évalue, pour $(\mathrm{H}, \mathrm{K}) \in \mathcal{M}_n(\mathbb{R})^2$, la quantité

$$\begin{aligned}\big(F(\mathrm{A}_0 + \mathrm{K}) - F(\mathrm{A}_0)\big)\mathrm{H} &= \big(\mathrm{d}f_{\mathrm{A}_0+\mathrm{K}} - \mathrm{d}f_{\mathrm{A}_0}\big)\mathrm{H} \\ &= (\mathrm{A}_0 + \mathrm{K})\mathrm{H} + \mathrm{H}(\mathrm{A}_0 + \mathrm{K}) - \mathrm{A}_0\mathrm{H} - \mathrm{H}\mathrm{A}_0 \\ &= \mathrm{H}\mathrm{K} + \mathrm{K}\mathrm{H}.\end{aligned}$$

On en déduit que, l'application f est deux fois différentiable et $\mathrm{d}^2 f_{\mathrm{A}_0}$ est l'application

$$\mathrm{d}^2 f_{\mathrm{A}_0} : (\mathrm{H}, \mathrm{K}) \in \mathcal{M}_n(\mathbb{R}) \times \mathcal{M}_n(\mathbb{R}) \longmapsto \mathrm{H}\mathrm{K} + \mathrm{K}\mathrm{H}.$$

Ainsi, pour tout $(\mathrm{H}, \mathrm{K}) \in \mathcal{M}_n(\mathbb{R}) \times \mathcal{M}_n(\mathbb{R})$, on a

$$\begin{aligned}\mathrm{d}^2 g_{\mathrm{B}_0}(\mathrm{H}, \mathrm{K}) &= -(\mathrm{d}f_{\mathrm{A}_0})^{-1}\left(\mathrm{d}^2 f_{\mathrm{A}_0} \circ \mathrm{d}g_{\mathrm{B}_0}(\mathrm{H})\right)\mathrm{d}f_{\mathrm{A}_0}(\mathrm{K}) \\ &= -(\mathrm{d}f_{\mathrm{A}_0})^{-1}\left(\mathrm{d}^2 f_{\mathrm{A}_0}(\mathrm{d}g_{\mathrm{B}_0}(\mathrm{H}), \mathrm{d}f_{\mathrm{A}_0}(\mathrm{K}))\right) \\ &= -(\mathrm{d}f_{\mathrm{A}_0})^{-1}\left(\mathrm{d}g_{\mathrm{B}_0}(\mathrm{H})\,\mathrm{d}f_{\mathrm{A}_0}(\mathrm{K}) + \mathrm{d}f_{\mathrm{A}_0}(\mathrm{K})\,\mathrm{d}g_{\mathrm{B}_0}(\mathrm{H})\right).\end{aligned}$$

En particulier, pour tout $\mathrm{H} \in \mathcal{M}_n(\mathbb{R})$, on obtient

$$\mathrm{d}^2 g_{\mathrm{B}_0}(\mathrm{H}, \mathrm{H}) = -2(\mathrm{d}f_{\mathrm{A}_0})^{-1}\left(\mathrm{d}g_{\mathrm{B}_0}(\mathrm{H})\,\mathrm{d}f_{\mathrm{A}_0}(\mathrm{H})\right).$$

Ainsi en utilisant les relations déterminées auparavant pour les matrices $\mathrm{d}f_{\mathrm{A}_0}(\mathrm{H})$ et $\mathrm{d}g_{\mathrm{B}_0}(\mathrm{H})$, on en déduit que le coefficient en position $(1,1)$ dans la matrice $\mathrm{d}^2 g_{\mathrm{B}_0}(\mathrm{H}, \mathrm{H})$ vaut $-h_{11}^2 - h_{12}h_{21}$.

On en conclut que le développement demandé est

$$x_{11} = 1 + \frac{1}{2}h_{11} - \frac{1}{2}\left(h_{11}^2 + h_{12}h_{21}\right) + \|\mathrm{H}\|^2 \varepsilon_{11}(\mathrm{H}),$$

où $\varepsilon_{11}(\mathrm{H})$ correspond à la composante en position $(1,1)$ de la matrice $\varepsilon(\mathrm{H})$. L'application ε_{11} ainsi définie satisfait $\lim_{\mathrm{H}\to 0} \varepsilon_{11}(\mathrm{H}) = 0$.

---

[102] Pour un calcul analogue, on pourra consulter l'exemple de la p. 613.

CHAPITRE 11

Courbes et surfaces

Très souvent, la géométrie est apparentée aux dessins qu'elles peut engendrer[1] : tracés de droites, de courbes ou de surfaces. Ces notions apparaissent dans des domaines variés : on parle de ski « parabolique », de molécule d'ADN « hélicoïdale » ou de trajectoire « elliptique » des planètes. Ce chapitre a pour but d'introduire les notions de courbe et de surface dans l'espace et de répondre aux questions suivantes. Comment peut-on mathématiquement représenter une courbe ou une surface ? Existe-t-il plusieurs façons de le faire ? Laquelle est la meilleure ?

La première partie de ce chapitre est consacrée à préciser mathématiquement les différentes représentations de courbes et de surfaces (représentations paramétrée, cartésienne et implicite) et à en étudier les propriétés. La seconde partie du chapitre a pour objectif d'établir et d'explorer les liens existant entre ces différentes descriptions d'une courbe ou d'une surface, en ayant recours au théorème d'inversion locale et au théorème des fonctions implicites.

11.1 Trois définitions de courbes et de surfaces

On a coutume de concevoir les courbes dans le plan ou dans l'espace et les surfaces dans l'espace mais du point de vue mathématique, ces notions de courbes et de surfaces existent et se définissent de façon identique dans $\mathbb{R}^n$ pour tout $n \in \mathbb{N}^*$. Nous allons dans cette partie présenter trois façons de définir une courbe ou une surface : de manière paramétrée, sous forme d'équation cartésienne et de manière implicite.

11.1.1 Le règne des paramètres

La façon la plus intuitive de voir une courbe est de l'imaginer comme un « segment tordu » et c'est effectivement une manière de définir mathématiquement une courbe, voir la figure 1.

[1] On peut citer à ce sujet le site `mathcurve.com` de R. Ferréol ainsi que le numéro 45 de la revue du Palais de la Découverte de décembre 1995 qui constituent de véritables albums de courbes et surfaces mathématiques.



Fig. 1 Une courbe paramétrée.

DÉFINITION 11.1 (Arc paramétré)
Pour $n \in \mathbb{N}^*$ et $k \in \mathbb{N} \cup \{\infty\}$, on appelle arc paramétré *de classe $\mathcal{C}^k$ de $\mathbb{R}^n$ toute application φ définie sur un intervalle I de $\mathbb{R}$ et à valeurs dans $\mathbb{R}^n$ qui est de classe $\mathcal{C}^k$ sur I. On note $\varphi \in \mathcal{C}^k(I, \mathbb{R}^n)$.*

Remarque L'hypothèse que φ est de classe $\mathcal{C}^k$ sur un intervalle I de $\mathbb{R}$ peut sembler restrictive. Si l'application φ est de classe $\mathcal{C}^k$ sur un ensemble qui est une réunion finie d'intervalles, on préfère considérer la restriction de φ à chacun de ces intervalles séparément. Par ailleurs, l'étude des arcs paramétrés qui sont continus sans être de classe $\mathcal{C}^1$ conduit à des difficultés mathématiques importantes (Peano[2] a par exemple montré qu'il existe un arc paramétré continue de $\mathbb{R}^2$, défini sur l'intervalle $[0,1]$, dont l'image est le pavé $[0,1] \times [0,1]$). On se limitera donc principalement dans ce chapitre à l'étude des arcs paramétrés qui sont de classe $\mathcal{C}^k$ avec $k \in \mathbb{N}^* \cup \{\infty\}$. ◇

Considérons un arc paramétré $\varphi : t \in I \mapsto (\varphi_1(t), \ldots, \varphi_n(t))$. Si on note $M(t)$ le point de $\mathbb{R}^n$ dont les coordonnées dans la base canonique de $\mathbb{R}^n$ sont $(\varphi_1(t), \varphi_2(t), \ldots, \varphi_n(t))$ alors lorsque le paramètre t parcourt I, le point $M(t)$ décrit un sous-ensemble C de $\mathbb{R}^n$, appelé *courbe paramétrée* de (ou associée à) φ. On parle aussi de la *trajectoire* de φ ou du *support* de φ. Une courbe paramétrée est donc un sous-ensemble de $\mathbb{R}^n$ défini par

$$C = \{x \in \mathbb{R}^n \ ; \ \exists t \in I \quad x = \varphi(t)\}$$

où $\varphi \in \mathcal{C}^0(I, \mathbb{R}^n)$ est un arc paramétré. On dit que l'arc paramétré φ est un *paramétrage de la courbe paramétrée* C et que la courbe paramétrée C admet pour *représentation paramétrique* l'arc paramétré φ.

Il ne faut pas confondre une courbe paramétrée et un arc paramétré. Selon les définitions précédentes, un arc paramétré $\varphi \in \mathcal{C}^0(I, \mathbb{R}^n)$ est une application alors qu'une courbe paramétrée est un sous-ensemble de $\mathbb{R}^n$ correspondant à l'image $\varphi(I) \subset \mathbb{R}^n$ d'un arc paramétré.

[2] PEANO, Giuseppe (1858, Cuneo - 1932, Turin). Mathématicien et philosophe italien qui s'intéressa principalement aux fondements axiomatiques des mathématiques. La courbe de Peano, d'une conception très astucieuse, est une courbe du plan, au sens où elle est l'image du segment $[0,1]$ par une application continue, qui cependant remplit un carré. Il s'agit d'une courbe fractale : bien que formée d'une simple « ligne », elle est de « dimension 2 ».

Par abus de langage, on parlera de courbe paramétrée de classe $\mathcal{C}^k$ pour désigner une courbe paramétrée admettant pour représentation paramétrique un arc paramétré de classe $\mathcal{C}^k$.

DÉFINITION 11.2 *On dit qu'un arc paramétré $\varphi \in \mathcal{C}^0(I, \mathbb{R}^n)$ est* simple *si l'application φ est injective.*

Exemples

1. L'arc paramétré $\varphi : [0, 2\pi[\mapsto (\cos(t), \sin(t)) \in \mathbb{R}^2$ est un arc simple. L'arc paramétré $\psi : [0, 4\pi] \mapsto (\cos(t), \sin(t)) \in \mathbb{R}^2$ n'est pas un arc simple car $\psi(\pi) = \psi(3\pi)$. Ces deux arcs paramétrés ont pour image la même courbe paramétrée : le cercle de centre $(0, 0)$ et de rayon 1 dans le plan.

2. Une courbe de Lissajous[(3)] est l'image de l'arc paramétré

$$\varphi : t \in \mathbb{R} \longmapsto \big(a \sin(t),\ b \sin(\eta t + \phi)\big)$$

où ϕ et η sont deux paramètres satisfaisant aux conditions $0 \leqslant \phi \leqslant \pi/2$ et $\eta \geqslant 1$. Ces courbes peuvent avoir des allures très variées, allant d'une ellipse (lorsque $\eta = 1$) à une courbe dense dans le rectangle $[-a, a] \times [-b, b]$ (lorsque η est nombre irrationnel) en passant par des motifs comme ceux de la figure 2.

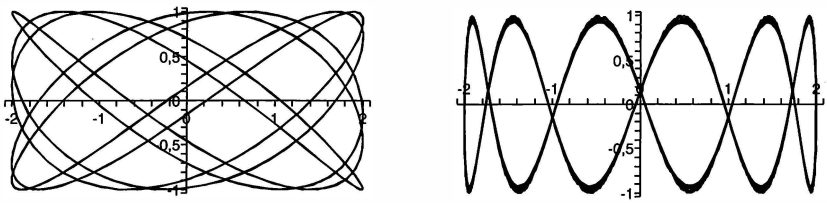


Fig. 2 Courbes de Lissajous correspondant aux valeurs $\eta = 7/5$ (à gauche) et $\eta = 6$ (à droite) avec $a = 2$, $b = 1$, $\phi = \pi/20$ pour $t \in [-100, 100]$.

On dit qu'une courbe paramétrée est simple lorsqu'elle admet pour représentation paramétrique un arc paramétré simple.

(3) Cette famille de courbes correspond aux trajectoires d'un point du plan dont les coordonnées x et y ont un mouvement sinusoïdal. Cette famille fut étudiée par Nathaniel Bowditch (mathématicien américain, 1773-1838) en 1815, puis plus en détail par Jules Lissajous en 1857.

LISSAJOUS, Jules (1822, Versailles - 1880, Plombières-les-Dijon).



Lissajous a été l'un des premiers physiciens français qui se soit consacré à l'étude des phénomènes acoustiques. Professeur de physique au Lycée Saint-Louis à Paris, il a développé une méthode qui permet d'étudier les sons à l'aide de l'œil, sans avoir recours à l'oreille. Elle se concrétise par le tracé de courbes dont les formes, obtenues par la réflexion successive sur des miroirs attachés à deux diapasons placés à angle droit, sont liées aux caractéristiques des sons. A la fin de sa carrière, Lissajous fut recteur de l'académie de Chambéry puis de celle de Besançon.

Remarque On rencontre fréquemment des arcs paramétrés $\varphi \in \mathcal{C}^0(\mathbb{R}, \mathbb{R}^n)$ où les applications coordonnées $\varphi_1, \ldots, \varphi_n$ sont des applications de $\mathbb{R}$ dans $\mathbb{R}$ périodiques. Lorsque les applications $\varphi_1, \ldots, \varphi_n$ possèdent un multiple de période commun, l'arc paramétré est lui-même périodique et on peut se contenter de l'étudier sur un intervalle de période. Par exemple, l'arc paramétré $\varphi : t \in \mathbb{R} \mapsto (\sin(\frac{1}{3}t),\ \cos(\frac{1}{2}t))$ est périodique de période 12π car φ_1 est périodique de période 6π et φ_2 est périodique de période 4π. Par contre, l'arc paramétré $\varphi : t \in \mathbb{R} \mapsto (\sin(\sqrt{2}t),\ \cos(t))$ n'est pas périodique bien que les applications coordonnées le soient (φ_1 est périodique de période $\sqrt{2}\pi$ et φ_2 est périodique de période 2π) ; cela est lié au fait que $\sqrt{2}$ est irrationnel et ne possède donc pas de multiple commun avec 2. ◇

Une façon de définir la notion de surface consiste à généraliser la définition 11.1 : une courbe paramétrée étant l'image d'un segment par une application à valeurs dans $\mathbb{R}^n$, une surface paramétrée peut être définie comme l'image d'un sous-ensemble de $\mathbb{R}^2$ par une application à valeurs dans $\mathbb{R}^n$.

DÉFINITION 11.3 (Nappe paramétrée)
Pour $n \in \mathbb{N}$, $n \geqslant 2$ et $k \in \mathbb{N} \cup \{\infty\}$, on appelle nappe paramétrée *de classe $\mathcal{C}^k$ de $\mathbb{R}^n$ toute application φ définie sur un sous-ensemble U de $\mathbb{R}^2$ et à valeurs dans $\mathbb{R}^n$ qui est de classe $\mathcal{C}^k$ sur U. On note $\varphi \in \mathcal{C}^k(U, \mathbb{R}^n)$.*

Considérons une nappe paramétrée $\varphi : (u,v) \in U \mapsto (\varphi_1(u,v), \ldots, \varphi_n(u,v))$ où $U \subset \mathbb{R}^2$. Si on note $M(u,v)$ le point de $\mathbb{R}^n$ dont les coordonnées dans la base canonique de $\mathbb{R}^n$ sont $(\varphi_1(u,v), \ldots, \varphi_n(u,v))$ alors lorsque les paramètres (u,v) parcourent U, le point $M(u,v)$ décrit un sous-ensemble S de $\mathbb{R}^n$, appelé la *surface paramétrée* de (ou associée à) φ. Une surface paramétrée est donc un sous-ensemble de $\mathbb{R}^n$ défini par

$$S = \{x \in \mathbb{R}^n \ ;\ \exists (u,v) \in U \quad x = \varphi(u,v)\}$$

où φ est une application de U dans $\mathbb{R}^n$. On dit que la nappe paramétrée φ est un *paramétrage de la surface paramétrée* S et que la surface paramétrée S admet pour *représentation paramétrique* la nappe paramétrée φ.

De la même manière qu'il ne faut pas confondre *courbe paramétrée* et *arc paramétré*, il ne faut pas confondre *surface paramétrée* et *nappe paramétré* : une nappe paramétrée $\varphi \in \mathcal{C}^0(U, \mathbb{R}^n)$ est une application alors qu'une surface paramétrée est un sous-ensemble de $\mathbb{R}^n$ qui correspond à l'image $\varphi(U) \subset \mathbb{R}^n$ d'une nappe paramétrée.

Par abus de langage, on parlera de surface paramétrée de classe $\mathcal{C}^k$ pour désigner une surface paramétrée admettant pour représentation paramétrique une nappe paramétrée de classe $\mathcal{C}^k$.

Remarque En général, l'ensemble des paramètres U est un *domaine* de $\mathbb{R}^2$, c'est-à-dire un ouvert connexe de $\mathbb{R}^2$. Lorsque l'ensemble des paramètres comporte plusieurs composantes connexes $U = U_1 \cup U_2$ avec $U_1 \cap U_2 = \varnothing$, on préférera parler de deux surfaces, l'une pour laquelle les paramètres sont dans U_1, l'autre pour laquelle les paramètres sont dans U_2. ◇

On dit qu'une nappe paramétrée $\varphi \in \mathcal{C}^k(U, \mathbb{R}^n)$ est simple si l'application φ est injective.

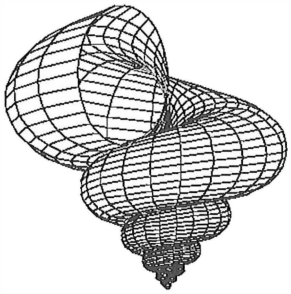
Exemples

1. Étant donné a, b, m et k quatre nombres réels, considérons la surface paramétrée par

$$\varphi : (u, v) \in \mathbb{R}^2 \longmapsto \begin{pmatrix} (a + b\cos(v))e^{mu}\cos(u) \\ (a + b\cos(v))e^{mu}\sin(u) \\ (ka + b\sin(v))e^{mu} \end{pmatrix}.$$

L'application φ est une nappe paramétrée de classe $\mathcal{C}^\infty$. En raison de leur forme, ces surfaces sont appelées *coquillages*. Les outils informatiques, comme MAPLE, se révèlent très utiles pour visualiser de telles surfaces. La commande `plot3d` du logiciel MAPLE permet de représenter graphiquement des surfaces paramétrées. La syntaxe est la suivante :
`plot3d([phi1(u,v),phi2(u,v),phi3(u,v)],u=umin..umax,v=vmin..vmax);`
où `phi1(u,v),phi2(u,v),phi3(u,v)` correspondent aux trois fonctions coordonnées de la nappe paramétrée φ et où $U = [\texttt{umin}, \texttt{umax}] \times [\texttt{vmin}, \texttt{vmax}]$.

```
> a:=1: m:=1/10: k:=3: b=:8/10:
> plot3d([(a+b*cos(v))*exp(m*u)*cos(u),(a+b*cos(v))*exp(m*u)*sin(u),
(k*a+b*sin(v))*exp(m*u)], u=-50..0,v=0..2*Pi,numpoints=10000):
```



2. La nappe paramétrée $\varphi : (u, v) \in \mathbb{R}^2 \longmapsto ((u+v)^2,\ uv,\ u^2+v^2) \in \mathbb{R}^3$ est de classe $\mathcal{C}^\infty$. Montrons que la surface paramétrée S associée à cette nappe paramétrée est incluse dans le plan $\mathcal{P}$ d'équation $x - 2y - z = 0$. Pour cela, considérons un point (x, y, z) de S et montrons que ce point appartient au plan $\mathcal{P}$ en établissant que ses coordonnées satisfont $x - 2y - z = 0$. Puisque $(x, y, z) \in S$, il existe des paramètres $(u, v) \in \mathbb{R}^2$ tels que

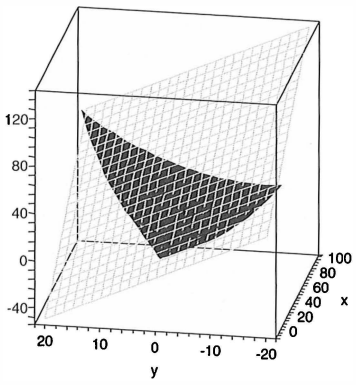
$$x = (u+v)^2, \quad y = u\,v, \quad z = u^2 + v^2.$$

On a $x - 2y - z = (u+v)^2 - 2u\,v - u^2 - v^2 = 0$ ce qui prouve que $S \subset \mathcal{P}$. Remarquons que l'on n'a pas égalité entre les deux surfaces. Par exemple le point de coordonnées $(-1, 0, -1)$ est un point du plan $\mathcal{P}$ alors que ce n'est pas un point de S car il n'existe pas de couple $(u, v) \in \mathbb{R}^2$ solution du système d'équations

$$(u+v)^2 = -1, \quad u\,v = 0 \quad \text{et} \quad u^2 + v^2 = -1.$$

On peut visualiser la situation en ayant recours à MAPLE. On utilise ici le fait que le plan d'équation $x - 2y - z = 0$ coïncide avec la nappe paramétrée $\psi : (x, y) \in \mathbb{R}^2 \mapsto (x,\ y,\ x - 2y)$.

```
> with(plots):
> fig1:=plot3d([(u+v)^2,u*v,u^2+v^2],u=-7..7,v=-3..3):
> fig2:=plot3d([x,y,x-2*y],x=-10..100,y=-20..20):
> display({fig1,fig2},axes=boxed,orientation=[-170,70]);
```



11.1.2 Ensembles implicites

Dans les chapitres précédents, nous avons rencontré des surfaces qui n'étaient pas décrites de façon paramétrée. En particulier, l'étude des quadriques (voir p. 515) a permis d'obtenir une famille de surfaces définies par l'annulation d'une forme quadratique. Plus généralement, on peut définir une courbe de $\mathbb{R}^2$ (resp. une surface de $\mathbb{R}^3$) comme l'ensemble des points $(x, y) \in \mathbb{R}^2$ (resp. des points $(x, y, z) \in \mathbb{R}^3$) liés par une condition du type $F(x, y) = 0$ (resp. du type $F(x, y, z) = 0$). On parle de courbes (resp. de surfaces) définies par une équation implicite.

DÉFINITION 11.4 (Courbe et surface implicites)

✗ *On appelle* courbe implicite *de classe* $\mathcal{C}^k$, $k \in \mathbb{N} \cup \{\infty\}$, *un sous-ensemble de* $\mathbb{R}^2$ *de la forme*

$$\{(x, y) \in U \ ; \ F(x, y) = 0\}$$

où F *est une application définie sur un ouvert* U *de* $\mathbb{R}^2$, *à valeurs dans* $\mathbb{R}$ *qui est de classe* $\mathcal{C}^k$ *sur* U.

✗ *On appelle* surface implicite *de classe* $\mathcal{C}^k$, $k \in \mathbb{N} \cup \{\infty\}$, *un sous-ensemble de* $\mathbb{R}^3$ *de la forme*

$$\{(x, y, z) \in U \ ; \ F(x, y, z) = 0\}$$

où F *est une application définie sur un ouvert* U *de* $\mathbb{R}^3$, *à valeurs dans* $\mathbb{R}$ *qui est de classe* $\mathcal{C}^k$ *sur* U.

Exemples

1. Dans le plan, l'équation $x^2 + y^2 = 1$ définit le cercle de centre $(0, 0)$ et de rayon 1 comme une courbe implicite de la forme $\{(x, y) \in \mathbb{R}^2 \ ; \ F(x, y) = 0\}$ où $F : (x, y) \in \mathbb{R}^2 \mapsto x^2 + y^2 - 1$. On sera vigilant au fait que cette même équation $x^2 + y^2 = 1$ définit dans l'espace un cylindre de hauteur infinie, la fonction décrivant la surface de façon implicite étant $F : (x, y, z) \in \mathbb{R}^3 \mapsto x^2 + y^2 - 1$.

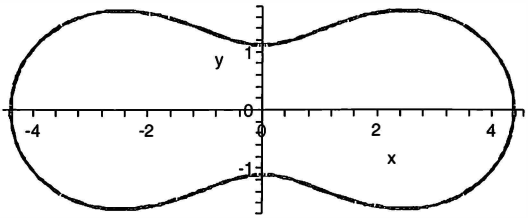
2. Dans le plan $\mathbb{R}^2$, l'équation

$$(x^2 + y^2)^2 + 2a^2(y^2 - x^2) + a^4 - b^4 = 0$$

où a et b sont deux réels fixés définit des courbes implicites appelées *Ovales de Cassini*[4]. Pour représenter une telle courbe, on peut avoir recours à l'instruction `implicitplot` de la librairie `plots` de MAPLE.

[4] CASSINI, Jean-Dominique (1625, Nice - 1712, Paris). Astronome, il fut le premier directeur de l'Observatoire de Paris. Il découvre la grande tache rouge de Jupiter en 1665, et détermine la même année la vitesse de rotation de Jupiter, Mars et Vénus. Il découvre également quatre satellites de Saturne, ainsi que la division des anneaux de Saturne en 1675. En 1673, il fait la première mesure précise de la distance de la terre au soleil. Cassini étudia les « Ovales de Cassini » en 1680 au cours d'études sur le mouvement de la terre par rapport au soleil.

```
> with(plots):
> a:=3: b:=3.2:
> implicitplot((x^2+y^2)^2+2*a^2*(y^2-x^2)+a^4-b^4=0,
  x=-6..6, y=-6..6);
```

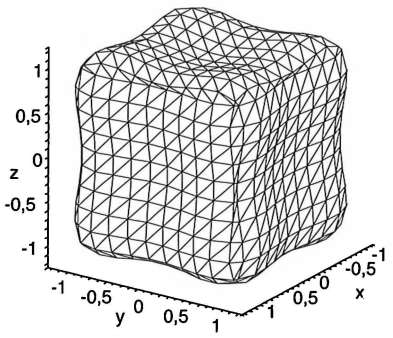


3. Dans $\mathbb{R}^3$, l'équation $x^4+y^4+z^4 = x^2+y^2+z^2$ définit une surface implicite de la forme $\{(x,y,z) \in \mathbb{R}^3 \ ; \ F(x,y,z) = 0\}$ où

$$F : (x,y,z) \in \mathbb{R}^3 \mapsto x^4 - x^2 + y^4 - y^2 + z^4 - z^2.$$

Pour représenter cette surface implicite, on peut avoir recours à l'instruction `implicitplot3d` de la librairie `plots` de MAPLE.

```
> with(plots):
> implicitplot3d(x^4+y^4+z^4 = x^2+y^2+z^2, x=-1.2..1.2,
  y=-1.2..1.2, z=-1.2..1.2);
```



Remarque On n'a pas défini la notion de courbe implicite dans $\mathbb{R}^3$. La raison en est la difficulté intrinsèque à formaliser mathématiquement une telle définition. On pourrait penser à définir une courbe implicite dans l'espace en considérant une fonction $F : (x,y,z) \in \mathbb{R}^3 \mapsto (F_1(x,y,z), F_2(x,y,z)) \in \mathbb{R}^2$ et, à l'instar de la définition 11.4, à voir une courbe implicite dans l'espace comme l'ensemble des points (x,y,z) de $\mathbb{R}^3$ annulant les deux fonctions réelles F_1 et F_2. Une telle courbe serait l'intersection des deux surfaces implicites définies par $F_1(x,y,z) = 0$ et $F_2(x,y,z) = 0$.

Un des inconvénients majeurs de cette approche est que l'intersection de deux surfaces n'est pas toujours une courbe au sens usuel du terme. On peut s'en convaincre aisément en considérant l'intersection de deux plans : celle-ci peut-être vide si les deux plans sont parallèles non confondus, le plan lui-même si les deux plans sont confondus ou une droite dans le cas général. Pour lever cette difficulté dans une tentative de définir les courbes implicites dans $\mathbb{R}^3$, il est nécessaire d'ajouter une condition d'indépendance entre les fonctions F_1 et F_2. La notion mathématique sous-jacente à ce type de condition est celle de « sous-variété ». Cette notion sort du cadre de ce cours et nous nous restreindrons[5] aux seuls cas de la définition 11.4 : les courbes implicites dans le plan et les surfaces implicites dans l'espace. ◇

11.1.3 Courbes et surfaces cartésiennes

Un cas particulier de surface implicite (on pourrait faire la même remarque pour les courbes implicites) est celui où l'une des coordonnées dans l'équation implicite $F(x, y, z) = 0$ peut s'exprimer en fonction des deux autres coordonnées. Ainsi, si on a une équivalence du type

$$F(x, y, z) = 0 \quad \Longleftrightarrow \quad z = f(x, y),$$

alors la surface implicite $\{(x, y, z) \in U \ ; \ F(x, y, z) = 0\}$ coïncide avec le graphe[6] de l'application f.

DÉFINITION 11.5 (Courbe et surface cartésiennes)

✗ *On appelle* courbe cartésienne *de classe $\mathcal{C}^k$, $k \in \mathbb{N} \cup \{\infty\}$, le graphe d'une application définie sur un ensemble I de $\mathbb{R}$ à valeurs dans $\mathbb{R}$ qui est de classe $\mathcal{C}^k$ sur I.*

✗ *On appelle* surface cartésienne *de classe $\mathcal{C}^k$, $k \in \mathbb{N} \cup \{\infty\}$, le graphe d'une application définie sur un ensemble U de $\mathbb{R}^2$ à valeurs dans $\mathbb{R}$ qui est de classe $\mathcal{C}^k$ sur U.*

Remarques

1. On parle de la courbe (resp. surface) cartésienne d'équation $y = f(x)$ (resp. $z = f(x, y)$) pour désigner la courbe (resp. surface) cartésienne qui est le graphe de la fonction f de $\mathbb{R}$ dans $\mathbb{R}$ (resp. de $\mathbb{R}^2$ dans $\mathbb{R}$).

---

[5] Quelques aspects de la notion de sous-variété seront néanmoins abordés lors de la dernière partie de ce chapitre.

[6] On ne confondra pas le graphe d'une fonction f de $\mathbb{R}^2$ dans $\mathbb{R}$ qui est défini comme l'ensemble $\{(x, y, z) \in \mathbb{R}^3 \ ; \ z = f(x, y)\}$ avec la représentation graphique de la dite fonction.

2. Il résulte de la définition 11.5, qu'une courbe cartésienne est un sous-ensemble de $\mathbb{R}^2$ et qu'une surface cartésienne est un sous-ensemble de $\mathbb{R}^3$. Ainsi que nous l'avons indiqué au sujet des courbes et surfaces implicites, bien qu'une généralisation de ces notions à des espaces de dimension supérieure existe, dans ce cours nous ne nous intéresserons pas à la notion de courbes cartésiennes dans $\mathbb{R}^n$ avec $n \geqslant 3$ ni à la notion de surface dans $\mathbb{R}^m$ avec $m \geqslant 4$ compte tenu des prérequis nécessaires à une telle étude.

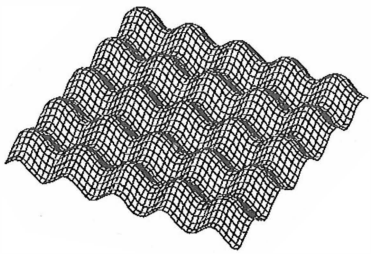
3. Puisqu'une courbe cartésienne correspond au graphe d'une fonction f de $\mathbb{R}$ dans $\mathbb{R}$, l'étude d'une telle courbe se réduit à l'étude de la fonction f. Nous n'insisterons donc pas sur l'étude des courbes cartésiennes dans ce chapitre. $\diamond$

Exemples

1. L'application $f : x \in \mathbb{R} \mapsto x^2 \in \mathbb{R}$ définit une courbe cartésienne d'équation $y = x^2$. Cette courbe est une parabole.

2. Une « boîte à œufs » est une surface cartésienne dont l'équation est de la forme $z = a(\sin(x/b) + \sin(y/b))$, a et b étant deux réels non nuls. Cette surface est obtenue en translatant une sinusoïde le long d'une autre sinusoïde isométrique[7], perpendiculairement à celle-ci. L'instruction MAPLE `plot3d` permet de visualiser des surfaces cartésiennes.

```
> a:=1: b:=1:
> plot3d(a*(sin(x/b)+sin(y/b)),x=0..30,y=0..30);
```



3. Nous avons vu au chap. 8 que la surface cartésienne d'équation $z = x^2 - y^2$ est un paraboloïde hyperbolique[8].

11.1.4 Premiers liens entre les différentes représentations

Un des objectifs de ce chapitre est de montrer que des liens très étroits existent entre les différentes représentations de courbes et de surfaces que nous avons présentées (paramétrées, implicites et cartésiennes).

---

[7] C'est-à-dire qui est l'image de la première sinusoïde par une isométrie, voir la définition 9.9 p. 560.

[8] Voir en particulier la partie consacrée à la classification des quadriques, p. 515.

Un premier lien s'obtient en remarquant qu'une surface cartésienne définie comme le graphe d'une application $f : (x, y) \in U \mapsto f(x, y) \in \mathbb{R}$, où U est un sous-ensemble de $\mathbb{R}^2$, peut également être vue comme une surface paramétrée représentée dont un paramétrage est $\varphi : (u, v) \in U \mapsto (u, v, f(u, v)) \in \mathbb{R}^3$. Le même constat est valable pour les courbes du plan : toute courbe cartésienne définie comme le graphe d'une application $f : x \in I \mapsto f(x) \in \mathbb{R}$, où I est un intervalle de $\mathbb{R}$, peut être vue comme une courbe paramétrée dont un paramétrage est $\varphi : t \in I \mapsto (t, f(t)) \in \mathbb{R}^2$. Dans un tel cas, on parle de *paramétrage cartésien* de la surface ou de la courbe.

Un autre lien est que toute courbe (resp. surface) cartésienne peut être vue comme une courbe (resp. surface) implicite. En effet, une courbe cartésienne définie comme le graphe d'une application $f : x \in I \mapsto f(x) \in \mathbb{R}$, où I est un intervalle de $\mathbb{R}$, correspond à la courbe implicite définie par $F(x, y) = 0$ où $F : (x, y) \in I \times f(I) \mapsto y - f(x) \in \mathbb{R}$. De même, une surface cartésienne définie comme le graphe d'une application $f : (x, y) \in U \mapsto f(x, y) \in \mathbb{R}$, où U est un sous-ensemble de $\mathbb{R}^2$, correspond à la surface implicite définie par $F(x, y, z) = 0$ où $F : (x, y, z) \in U \times f(U) \mapsto z - f(x, y) \in \mathbb{R}$.

On peut visualiser ces liens entre les différentes représentations de surfaces (ou de courbes) par le diagramme suivant[9] :

Représentation paramétrique ⟵ Représentation cartésienne ⟶ Représentation implicite

En ce qui concerne plus spécifiquement les courbes, la situation que l'on sait le mieux étudier est celle des courbes cartésiennes. En effet, l'analyse des fonctions d'une variable réelle[10] permet d'obtenir de nombreuses informations aussi bien locales (concernant les tangentes, la concavité ou les asymptotes) que globales (variations, maxima et minima) sur la courbe. Pour pouvoir étudier facilement des courbes paramétrées ou implicites, il est donc intéressant de pouvoir les décrire comme des courbes cartésiennes.

Nous verrons ultérieurement que sous certaines conditions les courbes et les nappes paramétrées admettent aussi des représentations cartésiennes. L'outil pour établir ce lien sera le théorème d'inversion locale. De même, nous introduirons le théorème des fonctions implicites dont l'interprétation géométrique permet de décrire une courbe ou une surface implicite comme une courbe ou une surface cartésienne. Schématiquement, nous mettrons en évidence les liens suivants :

[9] Une flèche $A \to B$ indique que si une surface est décrite par A alors on peut aussi la décrire par B

[10] Voir le *Cours de première année* chap. 16.

$$\text{Représentation paramétrique} \xrightarrow{\text{(a)}} \text{Représentation cartésienne} \xleftarrow{\text{(b)}} \text{Représentation implicite}$$

(a) Sous les hypothèses du théorème d'inversion locale (voir la section 11.5).
(b) Sous les hypothèses du théorème des fonctions implicites (voir la section 11.6).

11.1.5 Quelques courbes et surfaces de référence

La droite dans le plan

Dans le plan, un paramétrage de la droite passant par le point de coordonnées (b, d) et dirigée par le vecteur (a, c) est

$$\varphi : t \in \mathbb{R} \longmapsto \big(at + b,\ ct + d\big)\,.$$

Une équation implicite de la droite de vecteur directeur $(-\beta, \alpha)$ passant par le point (x_0, y_0) est

$$\alpha x + \beta y + \gamma = 0,$$

où $\gamma = -\alpha x_0 + \beta y_0$. On pourra remarquer que si une même droite est décrite à la fois de façon paramétrée et de façon implicite alors les réels a, b, c, d d'une part et α, β, γ d'autre part sont liés par les relations $\alpha = -c$, $\beta = a$ et $\gamma = ad - bc$.

L'équation cartésienne d'une droite est de l'une des formes suivantes

$$y = \lambda x + \mu \quad \text{ou} \quad x = \delta$$

où λ, μ et δ désignent trois réels. La seconde forme correspond à une droite parallèle à l'axe des ordonnées.

Le cercle dans le plan

Dans le plan, un paramétrage du cercle de centre $(a, b) \in \mathbb{R}^2$ et de rayon $R \in \mathbb{R}_+^*$ est

$$\varphi : t \in [0, 2\pi] \longmapsto \big(R\cos(t) + a,\ R\sin(t) + b\big)\,.$$

En raison de la périodicité des applications cosinus et sinus d'autres paramétrages du cercle sont obtenus en considérant à la place de $[0, 2\pi]$ un intervalle quelconque de longueur supérieur à 2π. Un autre paramétrage est obtenu en échangeant les rôles de sinus et cosinus car $\cos(\pi/2 - t) = \sin(t)$ et $\sin(\pi/2 - t) = \cos(t)$.

Une équation implicite du cercle de centre $(a, b) \in \mathbb{R}^2$ et de rayon $R \in \mathbb{R}_+^*$ dans le plan est

$$(x - a)^2 + (y - b)^2 = R^2.$$

On ne peut pas décrire le cercle par une équation cartésienne. Toutefois, chaque demi-cercle « supérieur » et « inférieur » admet une représentation cartésienne respectivement d'équation

$$y = b + \sqrt{R^2 - (x - a)^2} \quad \text{et} \quad y = b - \sqrt{R^2 - (x - a)^2}.$$

L'ellipse dans le plan

Étant donné deux réels strictement positifs R_1 et R_2, la courbe paramétrée dont un paramétrage est

$$\varphi : t \in [0, 2\pi] \longmapsto \big(a + R_1 \cos(t),\ b + R_2 \sin(t)\big)$$

est une ellipse dans le plan dont les axes sont les axes des cordonnées et qui est centrée en (a, b). Lorsque $R_1 = R_2$, on retrouve le paramétrage d'un cercle dont le centre est (a, b).

Une équation implicite de cette ellipse est

$$\left(\frac{x-a}{R_1}\right)^2 + \left(\frac{y-b}{R_2}\right)^2 = 1.$$

Comme pour le cercle, on ne peut pas donner de représentation cartésienne globale de l'ellipse.

L'hyperbole dans le plan

Étant donné deux réels strictement positifs R_1 et R_2, l'arc paramétré

$$\varphi : t \in \mathbb{R} \longmapsto \big(R_1 \operatorname{ch}(t),\ R_2 \operatorname{sh}(t)\big)$$

a pour image une hyperbole[11] dans le plan. L'hyperbole est appelée *hyperbole équilatère* lorsque $R_1 = R_2$.

Une équation implicite de cette hyperbole est

$$\left(\frac{x}{R_1}\right)^2 - \left(\frac{y}{R_2}\right)^2 = 1.$$

Signalons que[12] dans un repère $(O; \mathbf{u}, \mathbf{v})$ dont les axes sont les asymptotes de cette hyperbole, l'équation implicite de l'hyperbole s'écrit

$$XY = \frac{R_1^2 + R_2^2}{2},$$

où X et Y désignent les coordonnées de ce nouveau repère.

Dans le repère $(O; \mathbf{u}, \mathbf{v})$ une représentation cartésienne de l'hyperbole est

$$Y = \frac{R_1^2 + R_2^2}{2X}.$$

Comme pour l'ellipse, il n'est pas possible de donner une représentation cartésienne dans le repère initial.

---

[11] D'où la terminologie de *fonctions hyperboliques* données à ces deux fonctions.

[12] Voir à ce sujet l'étude des coniques effectuée p. 515.

Le plan dans l'espace

Dans l'espace, un paramétrage du plan contenant le point de coordonnées (c_1, c_2, c_3) et de vecteurs directeurs (a_1, a_2, a_3) et (b_1, b_2, b_3) est

$$\varphi : (u, v) \in \mathbb{R}^2 \longmapsto \big(a_1u + b_1v + c_1,\ a_2u + b_2v + c_2,\ a_3u + b_3v + c_3\big)\,.$$

Une équation implicite du plan dans $\mathbb{R}^3$ de vecteur normal (α, β, γ) et contenant le point (x_0, y_0, z_0) est

$$\alpha x + \beta y + \gamma z + \delta = 0$$

où $\delta = -(\alpha x_0 + \beta y_0 + \gamma z_0)$.

Selon les valeurs des coefficients α, β ou γ, on pourra obtenir une équation cartésienne de ce plan sous l'une des formes suivantes

$$z = \lambda x + \mu y + \nu \quad \text{ou} \quad y = \lambda' x + \nu' \quad \text{ou} \quad x = \delta$$

où λ, μ, ν, λ', ν' et δ désignent des réels.

La droite dans l'espace

Dans l'espace, un paramétrage de la droite passant par le point de coordonnées (c_1, c_2, c_3) et de vecteur directeur (a_1, a_2, a_3) est

$$\varphi : t \in \mathbb{R} \longmapsto \big(a_1t + c_1,\ a_2t + c_2,\ a_3t + c_3\big)\,.$$

Nous n'avons pas introduit les notions permettant de définir la représentation cartésienne et la représentation implicite d'une droite dans l'espace. Insistons sur le fait que dans $\mathbb{R}^3$ une équation de la forme $\alpha x+\beta y+\gamma z+\delta = 0$ correspond à une représentation d'un plan et non d'une droite. Une droite dans l'espace peut toutefois être décrite comme l'intersection de deux plans.

La sphère dans l'espace

Un paramétrage de la sphère de centre $(a, b, c) \in \mathbb{R}^3$ et de rayon $R \in \mathbb{R}_+^*$ dans l'espace est

$$\varphi : (u, v) \in [0, 2\pi] \times [-\tfrac{1}{2}\pi, \tfrac{1}{2}\pi] \longmapsto \begin{pmatrix} a + R\cos(u)\sin(v) \\ b + R\sin(u)\sin(v) \\ c + R\cos(v) \end{pmatrix}.$$

En raison de la périodicité des applications cosinus et sinus, d'autres paramétrages de la sphère sont obtenus en considérant des ensembles de définition de la forme $[\alpha, \alpha + 2\pi] \times [\beta, \beta + \pi]$ ou $[\alpha, \alpha + \pi] \times [\beta, \beta + 2\pi]$ avec $(\alpha, \beta) \in \mathbb{R}^2$.

Une équation implicite de la sphère de centre $(a, b, c) \in \mathbb{R}^3$ et de rayon $R \in \mathbb{R}_+^*$ dans l'espace est

$$(x - a)^2 + (y - b)^2 + (z - c)^2 = R^2.$$

Comme c'est le cas pour le cercle dans le plan, une sphère dans l'espace n'admet pas de représentation cartésienne. Toutefois, la demi-sphère de centre $(0, 0, 0)$ et de rayon $R \in \mathbb{R}_+^*$ située dans le demi-espace $\{(x, y, z) \in \mathbb{R}^3 \ ; \ z \geqslant 0\}$ admet une représentation cartésienne d'équation $z = \sqrt{R^2 - x^2 - y^2}$ et celle située dans le demi-espace $\{(x, y, z) \in \mathbb{R}^3 \ ; \ z \leqslant 0\}$ admet une représentation cartésienne d'équation $z = -\sqrt{R^2 - x^2 - y^2}$.

L'ellipsoïde dans l'espace

Un paramétrage de l'ellipsoïde de centre (a, b, c) et de longueur d'axes R_1, R_2 et R_3 (où $R_1, R_2, R_3 \in \mathbb{R}_+^*$) est

$$\varphi : (u, v) \in [0, 2\pi] \times [-\tfrac{1}{2}\pi, \tfrac{1}{2}\pi] \longmapsto \begin{pmatrix} a + R_1 \cos(u) \sin(v) \\ b + R_2 \sin(u) \sin(v) \\ c + R_3 \cos(v) \end{pmatrix}.$$

Une équation implicite de cet ellipsoïde est

$$\left(\frac{x-a}{R_1}\right)^2 + \left(\frac{y-b}{R_2}\right)^2 + \left(\frac{z-c}{R_3}\right)^2 = 1.$$

Comme pour la sphère, il n'existe pas d'équation cartésienne de l'ellipsoïde.

Le cylindre de révolution dans l'espace

Un paramétrage du cylindre de révolution[13] d'axe $(O; \mathbf{k})$ et de rayon $R \in \mathbb{R}_+^*$ dans l'espace est

$$\varphi : (u, v) \in [0, 2\pi] \times \mathbb{R} \longmapsto \big(R\cos(u),\ R\sin(u),\ v\big).$$

Une équation implicite de ce cylindre dans l'espace est[14]

$$x^2 + y^2 = R^2.$$

11.2 Surfaces cartésiennes dans l'espace

Les courbes cartésiennes étant définies comme le graphe d'une fonction f de $\mathbb{R}$ dans $\mathbb{R}$, l'étude d'une telle courbe se réduit à l'étude de la fonction f. Nous n'insistons pas sur l'étude des courbes cartésiennes dans ce chapitre et renvoyons au *Cours de première année*.

---

[13] Plus généralement, un paramétrage du cylindre d'axe $(O; \mathbf{k})$ porté par la courbe Γ du plan $z = 0$ paramétrée par $\gamma : t \in I \mapsto (\gamma_1(t), \gamma_2(t))$ est $\varphi : (u, v) \in I \times \mathbb{R} \mapsto (\gamma_1(u), \gamma_2(u), v)$.

[14] À ne pas confondre avec l'équation implicite d'un cercle dans le plan. Il est donc important de préciser la dimension de l'espace vectoriel dans lequel on se place. Plus généralement, un cylindre d'axe $(O; \mathbf{k})$ porté par la courbe Γ du plan $z = 0$, d'équation implicite $f(x, y) = 0$, est une surface de $\mathbb{R}^3$ qui admet pour représentation implicite $f(x, y) = 0$.

11.2.1 Courbes de niveau et lignes coordonnées

Considérons une surface cartésienne S de classe $\mathcal{C}^2$ qui est le graphe d'une application f définie sur un ouvert U de $\mathbb{R}^2$ et à valeurs dans $\mathbb{R}$. Géométriquement, la surface S est obtenue en associant à tout point $m \in U$ de coordonnées (x, y) le point $M \in \mathbb{R}^3$ de coordonnées (x, y, z) où $z = f(x, y)$. Pour représenter S il est commode de « couper » la surface par des plans « horizontaux » ou « verticaux ». On définit ainsi les *courbes de niveau* et les *lignes coordonnées* de la surface S.

Courbes de niveau

En « coupant » la surface S par des plans « horizontaux », autrement dit en cherchant l'intersection de la surface S avec des plans d'équation $z = z_0$, on obtient des ensembles[15] appelés *courbes de niveau* de f. D'un point de vue pratique, pour obtenir une visualisation de la surface par courbes de niveau, on projette ces « courbes » dans le plan $z = 0$. Il est alors nécessaire de fournir, avec ce projeté, l'information de la cote z_0 de chacune des courbes projetées. Cette façon de procéder est celle qui est employée sur les cartes topographiques, comme celles de l'Institut Géographique National[16]. Ainsi, la surface peut être « reconstruite » en trois dimensions à partir du projeté des courbes de niveau et de la cote qui est mentionnée sur chaque courbe. Les logiciels de visualisation graphique offrent une alternative consistant à utiliser un code de couleur pour donner l'information sur la cote.

Considérons par exemple la surface cartésienne S qui est le graphe de l'application $f : (x, y) \in \mathbb{R}^2 \mapsto x^2 - y^2$. Cette surface est connue sous le nom de *paraboloïde hyperbolique*, voir l'étude des quadriques effectuée p. 517. La courbe de niveau de S correspondant à l'intersection de S avec le plan d'équation $z = z_0$ a pour équation

$$x^2 - y^2 = z_0.$$

Pour $z_0 \neq 0$, il s'agit d'une hyperbole dans le plan d'équation $z = z_0$ dont les foyers sont $(\sqrt{2z_0}, 0, z_0)$ et $(-\sqrt{2z_0}, 0, z_0)$ si $z_0 > 0$ et dont les foyers sont $(0, \sqrt{2|z_0|}, z_0)$ et $(0, -\sqrt{2|z_0|}, z_0)$ si $z_0 < 0$. Les courbes de niveau dans le plan d'équation $z = 0$ sont deux droites d'équation $y = x$ et $y = -x$. Le fait que les courbes de niveau de S soient des hyperboles explique la présence du terme *hyperbolique* dans la dénomination de la surface.

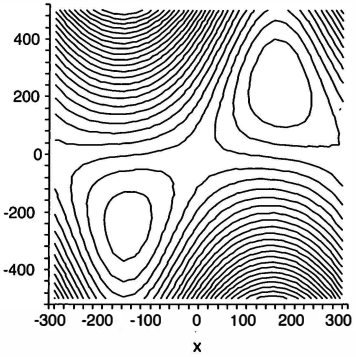
---

[15] Rappelons que l'intersection de deux surfaces n'est pas toujours une courbe au sens « usuel », voir à ce sujet la remarque de la p. 720. Dans la situation envisagée ici, l'intersection est définie de façon implicite : l'intersection de la surface $z = f(x, y)$ et du plan $z = z_0$ est l'ensemble $\{(x, y) \in U \ ; \ f(x, y) - z_0 = 0\}$.

[16] Des exemples sont disponibles sur : `www.ign.fr`

Il n'est pas toujours aussi aisé de déterminer les courbes de niveau d'une surface cartésienne et le recours à des logiciels pour visualiser ces courbes de niveau est très utile. La commande `contourplot` de la librairie `plots` du logiciel MAPLE permet de visualiser les courbes de niveau d'une surface cartésienne. Le nombre de courbes de niveau est déterminé par l'option `contours`. Par exemple, la représentation des courbes de niveau de la surface cartésienne qui est le graphe de la fonction $f : (x, y) \mapsto 2500 + \frac{1}{20}\, y \sin(\frac{x}{100}) - (\frac{y}{100})^2 + \frac{x}{100}$ s'obtient de la manière suivante.

```
> with(plots):
> f:= (x,y)->2500+y*sin(x/100)/20-(y/100)^2+x/100:
> contourplot(f(x,y),x=-300..300, y=-500..500, contours=25);
```



Lignes coordonnées

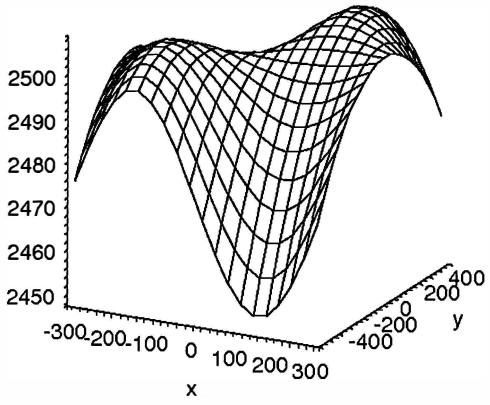
Une autre technique pour visualiser une surface cartésienne S consiste à tracer les *lignes coordonnées* qui sont les courbes obtenues par intersection de la surface avec les plans d'équation $x = x_0$ ou d'équation $y = y_0$. Il existe donc deux lignes coordonnées Γ_{x_0} et Γ_{y_0} passant par un point de coordonnées $(x_0, y_0, f(x_0, y_0))$ de la surface S.

- La première ligne coordonnée Γ_{x_0} est, dans le plan d'équation $x = x_0$, la courbe cartésienne d'équation $z = f(x_0, y)$. Cette courbe est le graphe de la fonction $y \in \mathbb{R} \mapsto f(x_0, y)$ qui est la seconde fonction partielle de f en (x_0, y_0).
- La seconde ligne coordonnée Γ_{y_0} est, dans le plan d'équation $y = y_0$, la courbe cartésienne d'équation $z = f(x, y_0)$. Il s'agit du graphe de la fonction $y \in \mathbb{R} \mapsto f(x, y_0)$ qui est la première fonction partielle de f en (x_0, y_0).

Reprenons l'exemple du paraboloïde hyperbolique qui est le graphe de l'application $f : (x, y) \in \mathbb{R}^2 \mapsto x^2 - y^2$. La première ligne coordonnée Γ_{x_0} qui passe par le point $(x_0, y_0, x_0^2 - y_0^2)$ a pour équation $z = x_0^2 - y^2$. Il s'agit d'une parabole de centre $(x_0, 0, x_0^2)$ et de foyer $(x_0, 0, x_0^2 - \frac{1}{4})$ dans le plan $x = x_0$. La seconde ligne coordonnée Γ_{y_0} a pour équation $z = x^2 - y_0^2$. Il s'agit d'une parabole de centre $(0, y_0, -y_0^2)$ et de foyer $(0, y_0, -y_0^2 + \frac{1}{4})$ dans le plan $y = y_0$. Le fait que les lignes coordonnées de cette surface soient des paraboles explique la présence du terme *paraboloïde* dans la dénomination de la surface.

Cette façon de visualiser les surfaces par lignes coordonnées est également largement utilisée par les logiciels graphiques. Par exemple, la commande `plot3d` de MAPLE permet de représenter une surface sous cette forme. La surface cartésienne qui est le graphe de $f : (x, y) \mapsto 2500 + \frac{1}{20}\, y \sin(\frac{x}{100}) - (\frac{y}{100})^2 + \frac{x}{100}$ se visualise à l'aide de lignes coordonnées de la manière suivante sous MAPLE.

```
> plot3d(f(x,y),x=-300..300,y=-500..500,grid=[20,20]);
```



11.2.2 Plan tangent à une surface cartésienne

Rappelons[17] que si f est une fonction de $\mathbb{R}$ dans $\mathbb{R}$ dérivable en $x_0 \in \mathbb{R}$ alors la tangente au graphe de f au point $(x_0, f(x_0))$ est la droite d'équation

$$y = f(x_0) + (x - x_0)\, f'(x_0)$$

dont un vecteur directeur est $\mathbf{u} = (1, f'(x_0))$. Nous allons voir qu'il est possible d'étendre cette notion aux fonctions de $\mathbb{R}^2$ dans $\mathbb{R}$.

Une application f d'un sous-ensemble U de $\mathbb{R}^2$ à valeurs dans $\mathbb{R}$ qui est de classe $\mathcal{C}^1$ admet un graphe S qui est une surface cartésienne de classe $\mathcal{C}^1$. On définit le *plan tangent* au point $(x_0, y_0, f(x_0, y_0))$ de la surface S comme étant le plan contenant les droites tangentes aux deux lignes coordonnées Γ_{x_0} et Γ_{y_0} passant par ce point.

La première ligne coordonnée Γ_{x_0} est le graphe de $g : y \in \mathbb{R} \mapsto f(x_0, y)$ et la pente de la tangente à cette courbe en y_0 est $g'(y_0) = \partial_y f(x_0, y)$. Cela signifie que le vecteur

$$\mathbf{u} = \begin{pmatrix} 0 \\ 1 \\ \partial_y f(x_0, y_0) \end{pmatrix}$$

[17] Voir le *Cours de première année* p. 714.

est un vecteur directeur de la tangente à la ligne coordonnée Γ_{x_0}. De la même façon, on peut vérifier que le vecteur

$$\mathbf{v} = \begin{pmatrix} 1 \\ 0 \\ \partial_x f(x_0, y_0) \end{pmatrix}$$

est un vecteur tangent à la ligne coordonnée Γ_{y_0} d'équation $z = f(x, y_0)$. Les deux vecteurs $\mathbf{u}$ et $\mathbf{v}$ étant linéairement indépendants, ils engendrent un plan dans $\mathbb{R}^3$. Ce plan est appelé *plan tangent* à la surface S au point $(x_0, y_0, f(x_0, y_0))$.

DÉFINITION 11.6 (Plan tangent)
Soient f une application de classe $\mathcal{C}^1$ définie sur un ouvert U de $\mathbb{R}^2$, à valeurs dans $\mathbb{R}$ et (x_0, y_0) un élément de U. Le plan tangent *à la surface cartésienne S d'équation $z = f(x, y)$ au point $M_0 = (x_0, y_0, f(x_0, y_0))$ est le plan passant par M_0 et dirigé par les vecteurs*[18]

$$\mathbf{u} = \begin{pmatrix} 1 \\ 0 \\ \partial_x f(x_0, y_0) \end{pmatrix} \quad \textit{et} \quad \mathbf{v} = \begin{pmatrix} 0 \\ 1 \\ \partial_y f(x_0, y_0) \end{pmatrix}.$$

Un vecteur normal $\mathbf{N}$ au plan tangent en $M_0 = (x_0, y_0, f(x_0, y_0))$ à la surface cartésienne S d'équation $z = f(x, y)$ est obtenu en effectuant le produit vectoriel entre les deux vecteurs $\mathbf{u}$ et $\mathbf{v}$; on a

$$\mathbf{N} = \begin{pmatrix} 1 \\ 0 \\ \partial_x f(x_0, y_0) \end{pmatrix} \wedge \begin{pmatrix} 0 \\ 1 \\ \partial_y f(x_0, y_0) \end{pmatrix} = \begin{pmatrix} -\partial_x f(x_0, y_0) \\ -\partial_y f(x_0, y_0) \\ 1 \end{pmatrix}.$$

On en déduit[19] que l'équation du plan tangent est :

$$z = f(x_0, y_0) + (x - x_0)\frac{\partial f}{\partial x}(x_0, y_0) + (y - y_0)\frac{\partial f}{\partial y}(x_0, y_0). \qquad (1)$$

Exemple Les surfaces cartésiennes d'équation $z = f_1(x, y)$ et $z = f_2(x, y)$ où

$$f_1 : (x, y) \in \mathbb{R}^2 \longmapsto e^{x+y} \qquad \text{et} \qquad f_2 : (x, y) \in \mathbb{R}^2 \longmapsto 2 - \frac{1}{(1+x)(1+y)}$$

passent toutes les deux par le point $(0, 0, 1)$ car $f_1(0, 0) = f_2(0, 0) = 1$. Déterminons leur plan tangent au point $(0, 0, 1)$. On a

$$\frac{\partial f_1}{\partial x}(x, y) = \frac{\partial f_1}{\partial y}(x, y) = e^{x+y} \qquad \text{donc} \qquad \frac{\partial f_1}{\partial x}(0, 0) = \frac{\partial f_1}{\partial y}(0, 0) = 1.$$

[18] On remarquera que pour que le plan tangent à la surface cartésienne S au point M_0 soit défini, il suffit que f soit différentiable en (x_0, y_0).

[19] Voir p. 726 l'équation d'un plan dans l'espace.

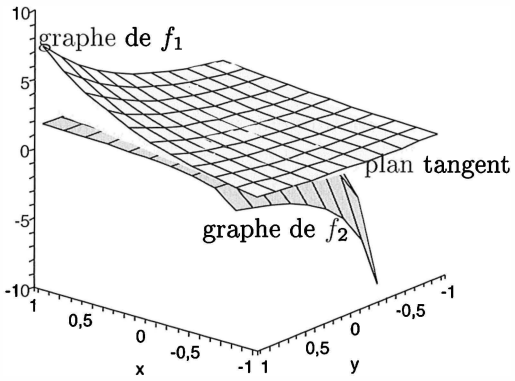
L'équation du plan tangent en $(0, 0, 1)$ à la surface cartésienne S_1 qui est le graphe de f_1 est donc $z = 1 + x + y$. Par ailleurs, on a

$$\frac{\partial f_2}{\partial x}(x, y) = \frac{1}{(1+x)^2(1+y)} \qquad \text{et} \qquad \frac{\partial f_2}{\partial y}(x, y) = \frac{1}{(1+y)^2(1+x)}$$

donc $\partial_x f_2(0, 0) = \partial_y f_2(0, 0) = 1$ et l'équation du plan tangent en $(0, 0, 1)$ à la surface cartésienne S_2 qui est le graphe de f_2 est $z = 1 + x + y$. Les deux surfaces cartésiennes S_1 et S_2 ont le même plan tangent en $(0, 0, 1)$. On dit que les deux surfaces sont des *surfaces tangentes.*

On peut visualiser les deux surfaces cartésiennes S_1 et S_2 ainsi que leur plan tangent en $(0, 0, 1)$ en ayant recours à MAPLE.

```
> f1:= (x,y)-> exp(x+y):
> f2:=(x,y)-> 2-1/((1+x)*(1+y)):
> plot3d({f1(x,y),f2(x,y),1+x+y},x=-1..1,y=-1..1,view=-10..10);
```



11.2.3 Position d'une surface par rapport à son plan tangent

La droite tangente au graphe d'une fonction f de $\mathbb{R}$ dans $\mathbb{R}$ en un point $(x_0, f(x_0))$ a pour équation $y = f(x_0) + (x - x_0)f'(x_0)$. Rappelons[20] que la position de la courbe par rapport à sa tangente est obtenue en déterminant le signe de $\Delta = f(x) - (f(x_0) + (x - x_0)f'(x_0))$. De la même manière, la position d'une surface cartésienne, qui est le graphe d'une fonction f de $\mathbb{R}^2$ dans $\mathbb{R}$, par rapport à son plan tangent en un point $(x_0, y_0, f(x_0, y_0))$ est obtenue en étudiant le signe de de la différence

$$\Delta = f(x, y) - \Big(f(x_0, y_0) + (x - x_0)\,\partial_x f(x_0, y_0) + (y - y_0)\,\partial_y f(x_0, y_0)\Big).$$

[20] Cette étude a été effectuée dans le *Cours de première année* p. 755 en ayant recours à la formule de Taylor-Lagrange.

Si f est de classe $\mathcal{C}^2$, en utilisant la formule de Taylor[21] à l'ordre 2 au voisinage de (x_0, y_0), on obtient :

$$\begin{aligned}\Delta &= \tfrac{1}{2}\big((x-x_0)^2\,\partial^2_{x^2} f(x_0,y_0) + 2(x-x_0)(y-y_0)\,\partial^2_{xy} f(x_0,y_0) \\ &\qquad + (y-y_0)^2\,\partial^2_{y^2} f(x_0,y_0)\big) + \mathcal{O}_{(x_0,y_0)}(\|(x-x_0,y-y_0)\|^2) \\ &= \frac{1}{2}\,{}^t\begin{pmatrix} x-x_0 \\ y-y_0 \end{pmatrix} \cdot \mathrm{H}_f(x_0,y_0) \cdot \begin{pmatrix} x-x_0 \\ y-y_0 \end{pmatrix} + \mathcal{O}_{(x_0,y_0)}(\|(x-x_0,y-y_0)\|^2).\end{aligned}$$

Le signe de Δ au voisinage de (x_0, y_0) est[22] donné par le signe de la quantité

$${}^t\begin{pmatrix} x-x_0 \\ y-y_0 \end{pmatrix} \cdot \mathrm{H}_f(x_0,y_0) \cdot \begin{pmatrix} x-x_0 \\ y-y_0 \end{pmatrix}.$$

Il dépend des valeurs propres de la matrice hessienne $\mathrm{H}_f(x_0, y_0) \in \mathcal{M}_2(\mathbb{R})$. Rappelons[23] que sous l'hypothèse où f est de classe $\mathcal{C}^2$, le théorème de Schwarz indique que la matrice $\mathrm{H}_f(x_0, y_0)$ est symétrique et que par conséquent ses deux valeurs propres sont réelles. Selon le signe de celles-ci, on peut déterminer[22] localement la position relative de la surface par rapport à son plan tangent en $(x_0, y_0, f(x_0, y_0))$. Les différentes situations envisageables au voisinage de ce point sont précisées dans la proposition suivante.

Proposition 11.1 (Position du plan tangent)
Soit f une application de classe $\mathcal{C}^2$ définie sur un ouvert U de $\mathbb{R}^2$ à valeurs dans $\mathbb{R}$ et soient λ_1 et λ_2 (avec $\lambda_1 \leqslant \lambda_2$) les deux valeurs propres réelles de la matrice hessienne de f en $(x_0, y_0) \in U$.

La position de la surface cartésienne d'équation $z = f(x, y)$ par rapport à son plan tangent en $M_0 = (x_0, y_0, f(x_0, y_0))$ est donnée, au voisinage de M_0, de la manière suivante :

1. *si $0 < \lambda_1 \leqslant \lambda_2$ alors le plan tangent en M_0 est en dessous de la surface ;*
2. *si $\lambda_1 \leqslant \lambda_2 < 0$ alors le plan tangent en M_0 est au-dessus de la surface ;*
3. *si $\lambda_1 < 0 < \lambda_2$ alors le plan tangent en M_0 traverse la surface ;*
4. *si $\lambda_1 = 0$ ou $\lambda_2 = 0$ alors on ne peut pas conclure.*

Exemples

1. Considérons le paraboloïde elliptique d'équation $z = x^2 + y^2$. L'application $f : (x, y) \in \mathbb{R}^2 \mapsto x^2 + y^2$ est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}^2$ et a pour dérivées partielles premières en $(x, y) \in \mathbb{R}^2$

$$\frac{\partial f}{\partial x}(x,y) = 2x \quad \text{et} \quad \frac{\partial f}{\partial y}(x,y) = 2y.$$

[21] Voir p. 684.

[22] Une étude analogue a été menée lors de l'étude des extrema de fonctions de plusieurs variables, en particulier lors de la preuve du théorème 10.10 p. 693.

[23] Voir le théorème 10.3 p. 650.

En $(0,0)$, ces deux dérivées sont nulles et l'équation du plan tangent est $z = 0$. Les dérivées partielles secondes de f sont

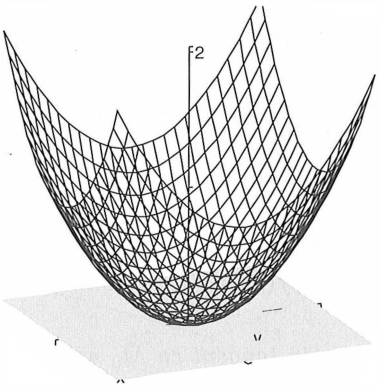
$$\frac{\partial^2 f}{\partial x^2}(x,y) = \frac{\partial^2 f}{\partial y^2}(x,y) = 2 \quad \text{et} \quad \frac{\partial^2 f}{\partial x \partial y}(x,y) = \frac{\partial^2 f}{\partial y \partial x}(x,y) = 0$$

ce qui implique que la matrice hessienne de f en $(0,0)$ est :

$$\mathrm{H}_f(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

Elle admet une valeur propre double qui est 2. D'après la proposition 11.1, on en déduit qu'au voisinage de l'origine le paraboloïde elliptique est au-dessus du plan $z = 0$. La proposition 11.1 ne permet de préciser la position du paraboloïde par rapport au plan tangent d'équation $z = 0$ qu'au voisinage de l'origine mais comme que pour tout $(x,y) \in \mathbb{R}^2$ on a $f(x,y) = x^2 + y^2 \geqslant 0$, on peut affirmer que le paraboloïde est partout au-dessus du plan d'équation $z = 0$.
On peut visualiser avec MAPLE le paraboloïde et son plan tangent à l'origine.

```
> plot3d({x^2+y^2,0},x=-1..1,y=-1..1);
```



2. Considérons le paraboloïde hyperbolique d'équation $z = x^2 - y^2$. L'application $f : (x,y) \in \mathbb{R}^2 \mapsto x^2 - y^2$ est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}^2$ et a pour dérivées partielles premières en $(x,y) \in \mathbb{R}^2$

$$\frac{\partial f}{\partial x}(x,y) = 2x \quad \text{et} \quad \frac{\partial f}{\partial y}(x,y) = -2y.$$

En $(0,0)$, ces deux dérivées sont nulles et l'équation du plan tangent est $z = 0$. Les dérivées partielles secondes de f sont

$$\frac{\partial^2 f}{\partial x^2}(x,y) = 2, \quad \frac{\partial^2 f}{\partial y^2}(x,y) = -2 \quad \text{et} \quad \frac{\partial^2 f}{\partial x \partial y}(x,y) = \frac{\partial^2 f}{\partial y \partial x}(x,y) = 0$$

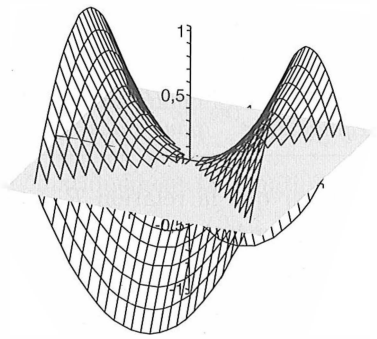
ce qui implique que la matrice hessienne de f en $(0,0)$ est :

$$\mathrm{H}_f(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}.$$

Elle possède deux valeurs propres distinctes qui sont 2 et -2. D'après la proposition 11.1, on en déduit que le plan $z = 0$ traverse la surface au point $(0, 0)$.

On peut visualiser avec MAPLE la surface et son plan tangent à l'origine.

```
> plot3d({x^2-y^2,0},x=-1..1,y=-1..1);
```



11.3 Étude des arcs paramétrés

Dans cette partie I et J désignent deux intervalles dans $\mathbb{R}$. Nous définissons en toute généralité la notion d'arc paramétré dans $\mathbb{R}^n$; les deux cas particuliers importants sont le cas $n = 2$ correspondant aux arcs paramétrés dans le plan et le cas $n = 3$ correspondant aux arcs paramétrés dans l'espace.

11.3.1 Classification des arcs paramétrés

Plusieurs arcs paramétrés[(24)] peuvent avoir la même image et décrire la même courbe paramétrée. Par exemple, le demi-cercle de centre $(0, 0)$ et de rayon 1 situé dans le demi-plan supérieur peut être décrit par les deux arcs paramétrés $\varphi : t \in [0, \pi] \mapsto (\cos(t), \sin(t))$ et $\psi : t \in [-1, 1] \mapsto (t, \sqrt{1 - t^2})$. Il est donc naturel de vouloir essayer de regrouper *certains* arcs paramétrés décrivant une même courbe et possédant des propriétés communes. Ce regroupement s'effectue par le biais d'une *relation d'équivalence* sur l'ensemble des arcs paramétrés.

DÉFINITION 11.7 (Arcs paramétrés $\mathcal{C}^k$-équivalents)
Soit $k \in \mathbb{N}^* \cup \{\infty\}$. On dit que l'arc paramétré $\psi \in \mathcal{C}^0(J, \mathbb{R}^n)$ est $\mathcal{C}^k$-équivalent *à l'arc paramétré $\varphi \in \mathcal{C}^0(I, \mathbb{R}^n)$ s'il existe un $\mathcal{C}^k$-difféomorphisme[(25)] θ de I sur J tel que l'on ait $\varphi = \psi \circ \theta$.*

(24) Voir la définition 11.1 p. 714.

(25) Rappelons, voir la définition 10.15 p. 664 qu'un $\mathcal{C}^k$-difféomorphisme est une application de classe $\mathcal{C}^k$, bijective dont la réciproque est elle aussi de classe $\mathcal{C}^k$. S'agissant ici d'un difféomorphisme entre intervalles de $\mathbb{R}$, on notera que la dérivé de θ ne s'annule pas sur I.

Si l'arc paramétré $\psi \in \mathcal{C}^0(J,\mathbb{R}^n)$ est $\mathcal{C}^k$-équivalent à l'arc paramétré $\varphi \in \mathcal{C}^0(I,\mathbb{R}^n)$, on dit que ψ se déduit de φ par le *changement de paramétrage* θ.

PROPOSITION 11.2 (Arc géométrique)
Pour tout $k \in \mathbb{N}^* \cup \{\infty\}$, la relation $\mathcal{R}$ définie par « les arcs paramétrés $\psi \in \mathcal{C}^0(J,\mathbb{R}^n)$ et $\varphi \in \mathcal{C}^0(I,\mathbb{R}^n)$ sont $\mathcal{C}^k$-équivalents » est une relation d'équivalence[26] *sur l'ensemble des arcs paramétrés. On appelle* arc géométrique *de classe $\mathcal{C}^k$ toute classe d'équivalence*[26] *pour cette relation.*

Démonstration Pour montrer que la relation $\mathcal{R}$ est une relation d'équivalence sur l'ensemble des arcs paramétrés, il faut montrer[26] qu'elle est réflexive, symétrique et transitive.

$\trianglerighteq$ La relation $\mathcal{R}$ est réflexive puisque tout arc paramétré $\varphi \in \mathcal{C}^0(I,\mathbb{R}^n)$ est équivalent à lui-même. On a en effet $\varphi = \varphi \circ \theta$ où θ est l'identité et l'identité est un $\mathcal{C}^k$-difféomorphisme pour tout $k \in \mathbb{N}^* \cup \{\infty\}$.

$\trianglerighteq$ Montrons que la relation $\mathcal{R}$ est symétrique. Si $\varphi \in \mathcal{C}^0(I,\mathbb{R}^n)$ est un arc paramétré $\mathcal{C}^k$-équivalent à l'arc paramétré $\psi \in \mathcal{C}^0(J,\mathbb{R}^n)$ alors il existe un $\mathcal{C}^k$-difféomorphisme θ de I sur J tel que $\varphi = \psi \circ \theta$. L'application θ étant bijective, on a $\psi = \varphi \circ \theta^{-1}$ où θ^{-1} désigne la bijection réciproque de θ et par définition d'un difféomorphisme[25], l'application θ^{-1} est aussi un $\mathcal{C}^k$-difféomorphisme. La relation $\psi = \varphi \circ \theta^{-1}$ permet donc de conclure que l'arc paramétré $\psi \in \mathcal{C}^0(J,\mathbb{R}^n)$ est $\mathcal{C}^k$-équivalent à l'arc paramétré $\varphi \in \mathcal{C}^0(I,\mathbb{R}^n)$.

$\trianglerighteq$ Montrons que la relation $\mathcal{R}$ est transitive. Considérons trois arcs paramétrés $\varphi \in \mathcal{C}^0(I,\mathbb{R}^n)$, $\psi \in \mathcal{C}^0(J,\mathbb{R}^n)$ et $\zeta \in \mathcal{C}^0(K,\mathbb{R}^n)$ où I, J et K désignent trois intervalles de $\mathbb{R}$ et supposons d'une part que ψ est $\mathcal{C}^k$-équivalent à φ et d'autre part que ζ est $\mathcal{C}^k$-équivalent à ψ. Nous devons montrer que ζ est $\mathcal{C}^k$-équivalent à φ. D'après les hypothèses, il existe un $\mathcal{C}^k$-difféomorphisme θ_1 de I sur J tel que $\varphi = \psi \circ \theta_1$ et il existe un $\mathcal{C}^k$-difféomorphisme θ_2 de J sur K tel que $\psi = \zeta \circ \theta_2$. On en déduit que $\varphi = \zeta \circ \theta_2 \circ \theta_1$. L'application $\theta = \theta_2 \circ \theta_1$ qui est la composée de deux bijections de classe $\mathcal{C}^k$ est une bijection de classe $\mathcal{C}^k$ sur I. Sa bijection réciproque $\theta^{-1} = \theta_1^{-1} \circ \theta_2^{-1}$ qui est la composée de deux bijections de classe $\mathcal{C}^k$ est une bijection de classe $\mathcal{C}^k$ sur K. L'application θ est donc un $\mathcal{C}^k$-difféomorphisme de I sur K et la relation $\varphi = \zeta \circ \theta$ indique que l'arc paramétré ζ est $\mathcal{C}^k$-équivalent à l'arc paramétré φ. □

On ne confondra pas *arc paramétré* et *arc géométrique*. Un arc paramétré est une application d'un intervalle de $\mathbb{R}$ dans $\mathbb{R}^n$ alors qu'un arc géométrique, défini comme une classe d'équivalence pour la relation $\mathcal{R}$, est un ensemble d'applications ; c'est l'ensemble de tous les arcs paramétrés $\mathcal{C}^k$-équivalents à un arc paramétré donné (que l'on appelle le *représentant* de la classe).

[26] Voir les définitions 2.28 et 2.29 p. 55 du *Cours de première année*.

Remarques

1. Comme une application de classe $\mathcal{C}^k$ est également de classe $\mathcal{C}^\ell$ pour tout entier ℓ tel que $1 \leqslant \ell \leqslant k$, si deux arcs paramétrés sont $\mathcal{C}^k$-équivalents alors ils sont également $\mathcal{C}^\ell$-équivalents pour tout entier ℓ tel que $1 \leqslant \ell \leqslant k$.

2. Deux arcs paramétrés $\varphi \in \mathcal{C}^0(I, \mathbb{R}^n)$ et $\psi \in \mathcal{C}^0(J, \mathbb{R}^n)$ qui sont $\mathcal{C}^k$-équivalents ont pour image la même courbe paramétrée C puisque $\varphi(I) = \psi(\theta(I)) = \psi(J)$. Ces deux arcs paramétrés appartiennent au même arc géométrique γ_φ ; la courbe paramétrée C qui est leur image commune est appelée *support de l'arc géométrique* γ_φ. ◇

On peut représenter schématiquement la situation de la manière suivante :

Arcs paramétrés $\mathcal{C}^k$-équivalents	Arc géométrique = classe d'équivalence	$\mathbb{R}^n$
$\varphi : I \mapsto \mathbb{R}^n$	$\gamma_\phi = \{\psi_k : J_k \mapsto \mathbb{R}^n \ ; \ \psi_k \mathcal{R} \varphi\}$	Courbe paramétrée associée à φ, ψ_1, ψ_2
$\psi_1 : J_1 \mapsto \mathbb{R}^n$		
$\psi_2 : J_2 \mapsto \mathbb{R}^n$		Support de l'arc géométrique

Deux arcs paramétrés peuvent avoir pour image la même courbe paramétrée sans être $\mathcal{C}^k$-équivalents pour aucun $k \in \mathbb{N}^* \cup \{\infty\}$ comme le montrent les exemples suivants.

Exemples

1. Les deux arcs paramétrés

$$\varphi : t \in [0, \pi] \longmapsto (\cos(t), \sin(t)) \quad \text{et} \quad \psi : t \in [-1, 1] \longmapsto (t, \sqrt{1 - t^2})$$

vérifient $\varphi = \psi \circ \theta$ où $\theta : t \in [0, \pi] \mapsto \cos(t) \in [-1, 1]$. L'application θ est une bijection de classe $\mathcal{C}^\infty$ dont la réciproque arccos n'est pas dérivable en -1 et 1. Il ne s'agit donc pas d'un $\mathcal{C}^1$-difféomorphisme et par conséquent les arcs paramétrés φ et ψ ne sont pas $\mathcal{C}^1$-équivalents bien qu'ils aient la même image : le demi-cercle de centre $(0, 0)$ et de rayon 1 situé dans le demi-plan supérieur. Toutefois, l'application arccos étant de classe $\mathcal{C}^\infty$ sur $]-1, 1[$, les deux arcs paramétrés

$$\varphi : t \in]0, \pi[\longmapsto (\cos(t), \sin(t)) \quad \text{et} \quad \psi : t \in]-1, 1[\longmapsto (t, \sqrt{1 - t^2})$$

sont $\mathcal{C}^\infty$-équivalents.

2. Les deux arcs paramétrés

$$\varphi : t \in [0,1] \longmapsto (t,t) \quad \text{et} \quad \psi : t \in [0,1] \longmapsto (1-t, 1-t)$$

sont $\mathcal{C}^\infty$-équivalents. En effet, on a $\varphi = \psi \circ \theta$ où $\theta : t \in [0,1] \mapsto 1-t \in [0,1]$ est un $\mathcal{C}^\infty$-difféomorphisme car[27] c'est une application

- de classe $\mathcal{C}^\infty$ sur $[0,1]$,
- injective puisque strictement décroissante sur $[0,1]$,
- surjective puisque $\theta([0,1]) = [0,1]$,
- dont la dérivée θ' ne s'annule pas sur $[0,1]$.

Ils ont pour image la courbe paramétrée $\{(x,y) \in \mathbb{R}^2 \ ; \ x \in [0,1] \ \text{ et } \ x = y\}$, autrement dit le segment de droite joignant le point de coordonnées $(1,1)$ à l'origine $(0,0)$.

3. Les arcs paramétrés

$$\varphi : t \in [-1,1] \longmapsto (1-|t|, 1-|t|) \quad \text{et} \quad \psi : t \in [0,1] \mapsto (t,t)$$

ont pour image la courbe paramétrée $\{(x,y) \in \mathbb{R}^2 \ ; \ x \in [0,1] \ \text{ et } \ x = y\}$, autrement dit le segment de droite joignant le point de coordonnées $(1,1)$ à l'origine. Cependant, ces deux arcs paramétrés ne sont $\mathcal{C}^k$-équivalents pour aucune valeur de $k \in \mathbb{N}^* \cup \{\infty\}$. En effet, pour que ces arcs paramétrés soient $\mathcal{C}^k$-équivalents, il doit exister un $\mathcal{C}^k$-difféomorphisme θ de $[-1,1]$ dans $[0,1]$ tel que $\varphi = \psi \circ \theta$, autrement dit tel que

$$\forall t \in [-1,1] \qquad 1-|t| = \theta(t).$$

La seule application θ telle que $\varphi = \psi \circ \theta$ est donc $\theta : t \in [-1,1] \mapsto 1-|t|$; elle n'est pas bijective (remarquer par exemple que $\theta(1) = \theta(-1) = 0$) donc ce n'est pas un difféomorphisme.

4. Montrons que l'arc paramétré $\varphi : t \in [0,1] \mapsto \left((t+t^{\frac{3}{2}})^2, 0\right)$ qui est de classe $\mathcal{C}^2$ et l'arc paramétré $\psi : t \in [0,2] \mapsto (t^2,0)$ qui est de classe $\mathcal{C}^\infty$ sont $\mathcal{C}^1$-équivalents mais ne sont pas $\mathcal{C}^2$-équivalents. Remarquons tout d'abord que ψ est de classe $\mathcal{C}^2$ sur $[0,1]$ mais n'est pas de classe $\mathcal{C}^3$; en effet, la première fonction coordonnées s'écrit $\psi_1 : t \mapsto t^3 + 2t^{\frac{5}{2}} + t^2$ et la fonction puissance $5/2$ est deux fois dérivable sur $[0,1]$ et admet pour dérivée seconde l'application $t \in [0,1] \mapsto \frac{15}{4}\sqrt{t}$ qui est continue sur $[0,1]$ mais n'est pas dérivable en 0. Établir que les arc paramétrés φ et ψ sont $\mathcal{C}^1$-équivalents revient à établir l'existence d'un $\mathcal{C}^1$-difféomorphisme θ de $[0,1]$ dans $[0,2]$ tel que $\varphi = \psi \circ \theta$, autrement dit tel que

$$\forall t \in [0,1] \qquad \left((t+t^{\frac{3}{2}})^2, 0\right) = (\theta^2(t), 0).$$

Le fait que θ doit être continue implique que θ est l'une des deux applications

$$t \in [0,1] \longmapsto t + t^{\frac{3}{2}} \quad \text{ou} \quad t \in [0,1] \longmapsto -(t+t^{\frac{3}{2}}).$$

Seule la première de ces deux applications est à valeurs dans $[0,2]$ ce qui implique que $\theta : t \in [0,1] \mapsto t + t^{\frac{3}{2}} \in [0,2]$. Cette application :

---

[27] Voir le théorème d'inversion globale p. 665.

- est de classe $\mathcal{C}^1$ sur $[0,1]$ mais n'est pas de classe $\mathcal{C}^2$ sur $[0,1]$ car elle n'est pas deux fois dérivable en 0,
- est injective puisqu'elle est strictement croissante sur $[0,1]$,
- est surjective puisque $\theta([0,1]) = [0,2]$,
- admet une dérivée $\theta' : t \in [0,1] \mapsto 1 + \frac{3}{2}\sqrt{t}$ qui ne s'annule pas sur $[0,1]$.

On en conclut[27] que l'application θ est $\mathcal{C}^1$-difféomorphisme mais n'est pas un $\mathcal{C}^2$-difféomorphisme ; les deux arcs paramétrés φ et ψ sont donc $\mathcal{C}^1$-équivalents mais ne sont pas $\mathcal{C}^2$-équivalents.

Exercice 11.1
1 - Montrer que les applications $\varphi_1 : t \in [0,1[\mapsto (\cos(2\pi t^2), \sin(2\pi t^2))$ et $\varphi_2 : t \in [0,2\pi[\mapsto (\cos(t), \sin(t))$ définissent des arcs paramétrés simples ayant la même image.
2 - Ces arcs paramétrés sont-ils $\mathcal{C}^1$-équivalents ?
3 - Montrer que l'arc paramétré $\varphi_3 : t \in]0,1] \mapsto (\cos(2\pi t^2), \sin(2\pi t^2))$ est $\mathcal{C}^\infty$-équivalent à l'arc paramétré $\varphi_4 : t \in]0,2\pi] \mapsto (\cos(t), \sin(t))$.

Orientation des arcs paramétrés

Soient $\varphi \in \mathcal{C}^0(I, \mathbb{R}^n)$ et $\psi \in \mathcal{C}^0(J, \mathbb{R}^n)$ deux arcs paramétrés $\mathcal{C}^k$-équivalents. D'après la définition 11.7, il existe un $\mathcal{C}^k$-difféomorphisme θ de I sur J tel que $\varphi = \psi \circ \theta$. Tous les difféomorphismes d'un intervalle de $\mathbb{R}$ vers un intervalle de $\mathbb{R}$ étant des applications strictement monotones[28], on peut les classer en deux catégories : ceux qui sont strictement croissants et ceux qui sont strictement décroissants. Cela nous amène à introduire la définition suivante.

Définition 11.8 *Soient $\varphi \in \mathcal{C}^0(I, \mathbb{R}^n)$ et $\psi \in \mathcal{C}^0(J, \mathbb{R}^n)$ deux arcs paramétrés $\mathcal{C}^k$-équivalents ($k \in \mathbb{N}^* \cup \{\infty\}$). On dit que les deux arcs paramétrés φ et ψ sont $\mathcal{C}^k$-équivalents s'il existe un $\mathcal{C}^k$-difféomorphisme θ de I sur J,* strictement croissant *tel que $\varphi = \psi \circ \theta$.*

Considérons un arc paramétré $\varphi \in \mathcal{C}^0(I, \mathbb{R}^n)$ et notons γ_φ l'arc géométrique correspondant à la classe d'équivalence $c\ell(\varphi)$ de φ pour la relation $\mathcal{R}$:« être $\mathcal{C}^k$-équivalent », voir la proposition 11.2. L'arc géométrique γ_φ est l'ensemble des arcs paramétrés qui sont $\mathcal{C}^k$-équivalents à l'arc paramétré φ et la définition 11.8 introduit une relation d'équivalence sur l'ensemble $c\ell(\varphi)$ qui permet de regrouper entre eux les arcs paramétrés de γ_φ qui se déduisent l'un de l'autre par un changement de paramétrage strictement croissant.

[28] En effet, un difféomorphisme θ de I vers $J = \theta(I)$ étant continu, si il n'était pas strictement monotone, il existerait des éléments distincts x_1 et x_2 dans I tel que $\theta(x_1) = \theta(x_2)$ ce qui contredirait le fait qu'un difféomorphisme est injectif.

Exercice 11.2 Montrer que la relation $\mathcal{R}_{\mathcal{O}}$ définie par « les arcs paramétrés $\varphi \in \mathcal{C}^0(J, \mathbb{R}^n)$ et $\psi \in \mathcal{C}^0(I, \mathbb{R}^n)$ sont $\mathcal{O}^k$-équivalents » est une relation d'équivalence sur l'ensemble des arcs paramétrés[29].

LEMME 11.1 *Tout arc géométrique de classe $\mathcal{C}^k$ est la réunion d'au plus deux classes d'équivalence pour la relation $\mathcal{R}_{\mathcal{O}}$: « être $\mathcal{O}^k$-équivalent ».*

Démonstration Raisonnons par l'absurde ; supposons qu'il y ait au moins trois classes d'équivalence, pour la relation $\mathcal{R}_{\mathcal{O}}$ dans un arc géométrique γ. Notons $\varphi_1 \in \mathcal{C}^0(I_1, \mathbb{R}^n)$, $\varphi_2 \in \mathcal{C}^0(I_2, \mathbb{R}^n)$ et $\varphi_3 \in \mathcal{C}^0(I_3, \mathbb{R}^n)$ trois arcs paramétrés appartenant à γ et représentant ces trois classes pour la relation $\mathcal{R}_{\mathcal{O}}$. Les trois arcs φ_1, φ_2 et φ_3 sont par hypothèse $\mathcal{C}^k$-équivalents deux à deux (car ils appartiennent à γ) mais ne sont pas $\mathcal{O}^k$-équivalents (car ils n'appartiennent pas à la même classe d'équivalence pour $\mathcal{R}_{\mathcal{O}}$). Il existe par conséquent un $\mathcal{C}^k$-difféomorphisme θ_{12} de I_2 vers I_1 qui n'est pas strictement croissant tel que $\varphi_2 = \varphi_1 \circ \theta_{12}$. Puisqu'un difféomorphisme d'un intervalle de $\mathbb{R}$ vers un intervalle de $\mathbb{R}$ qui n'est pas strictement croissant est strictement décroissant[28], l'application θ_{12} est strictement décroissante. De même, il existe un $\mathcal{C}^k$-difféomorphisme θ_{23} de I_3 vers I_2, strictement décroissant, tel que $\varphi_3 = \varphi_2 \circ \theta_{23}$. On déduit de ces considérations que $\varphi_3 = \varphi_1 \circ \theta_{12} \circ \theta_{23}$, autrement dit que $\varphi_3 = \varphi_1 \circ \theta$ où $\theta = \theta_{12} \circ \theta_{23}$ est un $\mathcal{C}^k$-difféomorphisme de I_3 dans I_1. L'application θ est strictement croissante car c'est la composée de deux applications strictement décroissantes. On peut donc affirmer d'après la relation $\varphi_3 = \varphi_1 \circ \theta$ que les arcs φ_1 et φ_3 sont $\mathcal{O}^k$-équivalents. Cela contredit l'hypothèse initiale. Il ne peut donc exister plus de deux classes d'équivalence pour la relation $\mathcal{R}_{\mathcal{O}}$ dans un arc géométrique γ. □

Il découle du lemme 11.1, qu'un arc géométrique admet ou bien une ou bien deux classes d'équivalence.

DÉFINITION 11.9 (Arc orientable)
On dit qu'un arc géométrique est orientable *lorsqu'il est la réunion d'exactement deux classes d'équivalence pour la relation « être $\mathcal{O}^k$-équivalent ».* Orienter *un arc géométrique, c'est choisir l'une de ces deux classes.*

Exemples

1. Considérons dans le plan, la droite D passant par le point A de coordonnées $(1, 1)$ et dirigée par le vecteur de composantes $(1, -1)$, voir la figure 3. Un paramétrage de la droite D est, voir p. 724,

$$\varphi : t \in \mathbb{R} \longmapsto \big(1 + t,\ 1 - t\big).$$

[29] On pourra s'inspirer de la démonstration de la proposition 11.2.

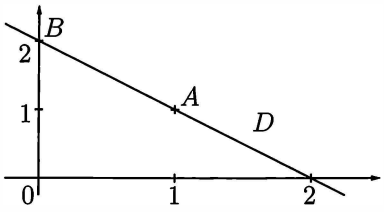


Fig. 3 Situation considérée.

Un autre vecteur directeur de la droite D est le vecteur de composantes $(-2, 2)$ et la droite passe par le point B de coordonnées $(2, 0)$. Cette seconde description conduit à proposer pour autre paramétrage de la droite D, voir p. 724,

$$\psi : t \in \mathbb{R} \longmapsto \big(2 - 2t,\ 2t\big)\,.$$

Les deux arcs paramétrés φ et ψ sont $\mathcal{C}^\infty$-équivalents puisque $\varphi = \psi \circ \theta$ où $\theta : t \in \mathbb{R} \mapsto \frac{1}{2}(1 - t) \in \mathbb{R}$ est un $\mathcal{C}^\infty$-difféomorphisme. Puisque l'application θ est strictement décroissante sur $\mathbb{R}$, les deux arcs paramétrés φ et ψ ne sont pas $\mathcal{O}^1$-équivalents. Il n'appartiennent donc pas à la même classe d'équivalence pour la relation $\mathcal{R}_\mathcal{O}$ « être $\mathcal{O}^k$-équivalent ». On peut montrer[30] qu'il n'existe pas de difféomorphisme $\widetilde{\theta}$ croissant tel que $\varphi = \psi \circ \widetilde{\theta}$. Il existe donc deux classes d'équivalence pour la relation $\mathcal{R}_\mathcal{O}$. On en déduit que l'arc géométrique dont le support est la droite D est orientable et que les deux arcs paramétrés φ et ψ fournissent les deux orientations de l'arc géométrique.

2. Les applications suivantes constituent deux paramétrages du cercle centré à l'origine et de rayon 1, voir p. 724,

$$\varphi : t \in [0, 2\pi] \longmapsto \big(\cos(t),\ \sin(t)\big)\,, \qquad \psi : t \in [0, 2\pi] \longmapsto \big(\cos(t),\ -\sin(t)\big)\,.$$

Les deux arcs paramétrés φ et ψ sont $\mathcal{C}^\infty$-équivalents puisque $\varphi = \psi \circ \theta$ où on établit aisément que $\theta : t \in [0, 2\pi] \mapsto 2\pi - t \in [0, 2\pi]$ est un $\mathcal{C}^\infty$-difféomorphisme. Puisque l'application θ est strictement décroissante sur $\mathbb{R}$, les deux arcs paramétrés φ et ψ ne sont pas $\mathcal{O}^1$-équivalents. Il n'appartiennent donc pas à la même classe d'équivalence pour la relation $\mathcal{R}_\mathcal{O}$ « être $\mathcal{O}^k$-équivalent ». On peut montrer[30] qu'il n'existe pas de difféomorphisme $\widetilde{\theta}$ croissant tel que $\varphi = \psi \circ \widetilde{\theta}$. Il existe donc deux classes d'équivalence pour la relation $\mathcal{R}_\mathcal{O}$. On en déduit que l'arc géométrique dont le support est le cercle centré à l'origine et de rayon 1 est orientable et les deux arcs paramétrés φ et ψ fournissent les deux orientations possibles de l'arc géométrique. L'orientation déterminée par φ est habituellement appelée *orientation dans le sens sens trigonométrique* et l'orientation déterminée par ψ *orientation dans le sens sens antitrigonométrique.*

[30] Voir la preuve de la proposition 11.3 pour un aperçu de la méthode.

Remarques

1. La notion d'orientation pour les arcs géométriques induit une notion d'orientation pour les arcs paramétrés : un arc paramétré φ est orientable si l'arc géométrique γ_φ correspondant à la classe d'équivalence de φ pour la relation « être $\mathcal{C}^k$-équivalent » est orientable (autrement dit si γ_φ possède exactement deux classes d'équivalence pour la relation « être $\mathcal{O}^k$-équivalent »).

2. Deux arcs paramétrés qui sont $\mathcal{C}^k$-équivalents pour $k \in \mathbb{N}^* \cup \{\infty\}$ étant également $\mathcal{C}^1$-équivalents, on peut contenter de la seule relation « être $\mathcal{O}^1$-équivalent » pour étudier la possibilité d'orienter un arc (géométrique ou paramétré). ◇

Les résultats que nous allons établir dans la suite du chapitre vont montrer que les arcs paramétrés qui ne sont pas orientables sont relativement peu courants.

PROPOSITION 11.3 *Tout arc paramétré simple*[31] *est orientable.*

Démonstration Considérons un arc paramétré simple $\varphi \in \mathcal{C}^0(I, \mathbb{R}^n)$ et notons γ_φ l'arc géométrique correspondant à l'ensemble des arcs paramétrés $\mathcal{C}^1$-équivalents à l'arc φ. Afin de montrer que l'arc paramétré φ est orientable, il faut montrer qu'il existe dans l'ensemble γ_φ exactement deux classes d'équivalence pour la relation $\mathcal{R}_\mathcal{O}$: « être $\mathcal{O}^1$-équivalent ».

Considérons l'arc paramétré $\psi : t \in J \mapsto \varphi(-t)$ où $J = \{-t \in \mathbb{R} \ ; \ t \in I\}$. Cet arc paramétré est $\mathcal{C}^1$-équivalent à l'arc φ car on a $\varphi = \psi \circ \theta$ où $\theta : t \in I \mapsto -t \in J$ est clairement un $\mathcal{C}^1$-difféomorphisme. On a donc $\psi \in \gamma_\varphi$.

Pour montrer que l'arc paramétré φ est orientable, nous allons montrer que dans γ_φ, l'arc paramétré ψ n'appartient pas à la même classe d'équivalence que φ pour la relation $\mathcal{R}_\mathcal{O}$. Cela établira que γ_φ possède deux classes d'équivalence pour la relation $\mathcal{R}_\mathcal{O}$ et par conséquent qu'il est orientable. Raisonnons par l'absurde et supposons que les arcs φ et ψ sont $\mathcal{O}^1$-équivalents : il existe un difféomorphisme $\widetilde{\theta}$ de J sur I strictement croissant tel que $\psi = \varphi \circ \widetilde{\theta}$. Comme on a aussi $\varphi = \psi \circ \theta$, on en déduit que $\varphi = \varphi \circ \widetilde{\theta} \circ \theta$, c'est-à-dire que

$$\forall t \in I \quad \varphi(t) = \varphi(\widetilde{\theta}(\theta(t))).$$

Puisque l'application φ est injective, cette assertion implique que[32] pour tout $t \in I$ on a $t = \widetilde{\theta}(\theta(t))$, c'est-à-dire que $\widetilde{\theta} \circ \theta = \mathrm{id}$. Comme l'application $\theta : t \in I \mapsto -t \in J$ est strictement croissante et que l'application $\widetilde{\theta}$ est strictement décroissante, la composée $\widetilde{\theta} \circ \theta$ est strictement décroissante. Cela est absurde car $\widetilde{\theta} \circ \theta = \mathrm{id}$ et l'application $\mathrm{id} : t \in I \mapsto t \in I$ est strictement croissante. □

---

[31] Rappelons, voir la définition 11.2, qu'un arc paramétré φ est dit simple si l'application φ est injective.

[32] Rappelons, voir la définition 2.9 p. 41 du *Cours de première année*, que φ est injective signifie que : $\forall (x, x') \in I^2 \quad (\varphi(x) = \varphi(x') \ \Rightarrow \ x = x')$.

11.3.2 Propriétés des arcs paramétrés

DÉFINITION 11.10 (Point simple, point multiple)
Soit $\varphi \in \mathcal{C}^0(I, \mathbb{R}^n)$ un arc paramétré et C la courbe paramétrée qui lui est associé.

✗ *On dit que le point M_0 de C est un* point simple *de l'arc paramétré φ si il existe un unique réel $t_0 \in I$ tel que $M_0 = \varphi(t_0)$.*

✗ *Si M_0 n'est pas un point simple de l'arc paramétré φ on dit que c'est un* point multiple.

✗ *On dit qu'un point multiple M_0 est un* point double *de l'arc paramétré φ s'il existe exactement deux réels distincts t_1 et t_2 dans I tels que $M_0 = \varphi(t_1) = \varphi(t_2)$.*

Remarques

1. Un arc paramétré simple, voir la définition 11.2, est un arc paramétré dont tous les points sont simples.

2. Si un arc paramétré $\varphi \in \mathcal{C}^0(I, \mathbb{R}^n)$ possède un nombre fini de points multiples alors il existe un ensemble fini de réels deux à deux distincts $\{t_0, \ldots, t_N\} \in I^N$ tel que l'application $\varphi : I \setminus \{t_0, \ldots, t_N\} \to \mathbb{R}^n$ soit injective, autrement dit définisse un arc paramétré simple. Un arc paramétré possédant un nombre fini de points multiples peut donc être considéré comme la réunion d'un nombre fini d'arcs paramétrés simples.

3. Deux arcs paramétrés qui sont $\mathcal{C}^k$-équivalents ont les mêmes points simples (et les mêmes points multiples). Cela implique en particulier que si on dispose de deux arcs paramétrés qui n'ont pas les mêmes points simples ou les mêmes points multiples, alors quel que soit $k \in \mathbb{N}^* \cup \{\infty\}$, ces deux arcs paramétrés ne sont pas $\mathcal{C}^k$-équivalents. C'est le cas par exemple des arcs paramétrés $\varphi : t \in [0, 2\pi[\mapsto (\cos(t), \sin(t))$ et $\psi : t \in [0, 2\pi] \mapsto (\cos(t), \sin(t))$ car le point, de coordonnées $(1, 0)$ est un point simple de φ mais un point double de ψ (on a $\psi(0) = \psi(2\pi) = (1, 0)$). Ces deux arcs paramétrés admettent la même image qui est le cercle centré à l'origine et de rayon 1 mais ne sont pas $\mathcal{C}^1$-équivalents.

4. En complément à la proposition 11.3, signalons que l'on peut montrer[33] qu'un arc paramétré qui possède au moins deux points simples est orientable. ◇

Exercice 11.3 Soient $\varphi \in \mathcal{C}^0(I, \mathbb{R}^n)$ et $\psi \in \mathcal{C}^0(J, \mathbb{R}^n)$ deux arcs paramétrés $\mathcal{C}^k$-équivalents et C la courbe paramétrée qui leur est associée.
1 - Montrer que si $M_0 \in C$ est un point simple de l'arc paramétré φ alors M_0 est aussi un point simple de l'arc paramétré ψ.
2 - En déduire que si un arc paramétré est simple alors tous les arcs paramétrés qui lui sont $\mathcal{C}^k$-équivalents sont simples.

[33] Voir par exemple J. Lelong-Ferrand et J.M. Arnaudiès, *Cours de mathématiques*, Dunod, Paris, 1977, t. 3, p. 259.

Exemples

1. L'origine est un point double de l'arc $\varphi : t \in [0, 4\pi[\mapsto (2\sin(t), \cos(\frac{3}{2}t))$ dont l'image est une courbe de Lissajous dans $\mathbb{R}^2$. En effet, le système

$$\begin{cases} \sin(t) = 0 \\ \cos(\frac{3}{2}t) = 0 \end{cases}$$

admet pour solutions dans $[0, 4\pi]$ les réels π et 3π. On pourra noter que l'application $t \in \mathbb{R} \mapsto (2\sin(t), \cos(\frac{3}{2}t))$ est périodique, de période 4π, et on pourrait établir en résolvant le système

$$\begin{cases} \sin(t) = \sin(s) \\ \cos(\frac{3}{2}t) = \cos(\frac{3}{2}s) \end{cases}$$

pour $(s, t) \in [0, 4\pi[^2$ que l'arc paramétré φ possède sept points doubles.

2. L'arc paramétré $\varphi : t \in \mathbb{R} \mapsto (\sin(t), \sin(\sqrt{2}t))$ représente lui aussi une courbe de Lissajous dans $\mathbb{R}^2$. L'arc paramétré φ n'est pas périodique (cela est lié au fait que $\sqrt{2}$ n'est pas un nombre rationnel). Montrons que l'origine est un point simple de cet arc paramétré en résolvant le système

$$\begin{cases} \sin(t) = 0 \\ \sin(\sqrt{2}t) = 0 \end{cases}$$

La première équation implique que $t = k\pi$ avec $k \in \mathbb{Z}$ et la seconde que $t = k'\pi/\sqrt{2}$ avec $k' \in \mathbb{Z}$. On doit donc avoir $k = k'/\sqrt{2}$ pour que t soit solution des deux équations simultanément. Comme $\sqrt{2}$ est un nombre irrationnel, $k'/\sqrt{2}$ est lui aussi un nombre irrationnel si $k' \neq 0$ et ne peut donc être égal à k qui est un entier relatif. La seule solution du système est $t = 0$ et correspond à $k = k' = 0$.

Par contre, cette courbe paramétrée admet des points multiples. On peut par exemple vérifier que $\varphi(\pi/(2\sqrt{2}) - \pi) = \varphi(\pi/(2\sqrt{2}) + \pi)$.

3. Les courbes de Lissajous dans $\mathbb{R}^3$ sont les trajectoires d'un point dans l'espace dont les composantes ont un mouvement sinusoïdal. Un paramétrage d'une courbe de Lissajous est

$$\varphi : t \in \mathbb{R} \longmapsto \begin{pmatrix} a\sin(t) \\ b\sin(\eta t + \phi) \\ c\sin(\mu t + \psi) \end{pmatrix}, \tag{2}$$

où a, b, c, η, μ, ϕ et ψ sont des réels. Lorsque la courbe ne s'« auto-intersecte » pas (c'est-à-dire lorsqu'elle n'a pas de point multiple sur un intervalle de période), on dit qu'elle forme un « nœud » dans l'espace, appelé *nœud de Lissajous*. Les arcs paramétrés suivants sont des nœuds de Lissajous[34] :

$$t \in \mathbb{R} \longmapsto \begin{pmatrix} \cos(3t) \\ \sin(2t) \\ 2\sin(7t) \end{pmatrix}, \; t \in \mathbb{R} \longmapsto \begin{pmatrix} \cos(3t) \\ \sin(4t) \\ 2\sin(17t) \end{pmatrix} \text{ et } \; t \in \mathbb{R} \longmapsto \begin{pmatrix} \cos(3t) \\ \sin(5t) \\ 2\sin(22t) \end{pmatrix}.$$

[34] On vérifiera qu'ils sont C^∞-équivalents à un arc paramétré de la forme (2).

Ils sont périodiques de période 2π. On pourra vérifier qu'en restriction à l'intervalle $[0, 2\pi[$, ces arcs paramétrés sont injectifs et par conséquent qu'ils ne s'auto-intersectent pas.

La commande `spacecurve` du logiciel MAPLE permet de tracer des courbes paramétrées de $\mathbb{R}^3$. Ainsi la première des trois courbes définies ci-dessus se visualise en effectuant les commandes suivantes :

```
> with(plots):
> spacecurve([cos(3*t),sin(2*t),2*sin(7*t)], t=0..2*Pi);
```

Les trois nœuds de Lissajous sont représentés sur la figure 4.

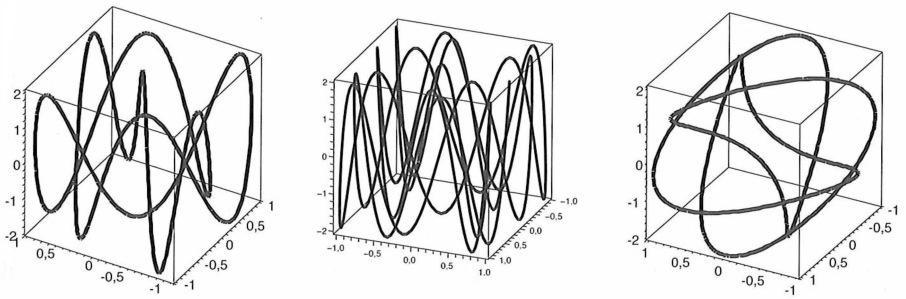


Fig. 4 Trois nœuds de Lissajous.

Régularité des arcs paramétrés

DÉFINITION 11.11 (Point régulier, point singulier)
Soit $\varphi \in \mathcal{C}^1(I, \mathbb{R}^n)$ un arc paramétré et C la courbe paramétrée qui lui est associée.

✗ *On dit qu'un point M_0 de C tel que $M_0 = \varphi(t_0)$ avec $t_0 \in I$ est un point* régulier *de C pour le paramétrage φ si $\varphi'(t_0) \neq \mathbf{0}$.*

✗ *Un point qui n'est pas régulier est appelé* point singulier[35].

✗ *Dans le cas où $\varphi \in \mathcal{C}^2(I, \mathbb{R}^n)$, on dit que le point M_0 de C est* birégulier *pour le paramétrage φ si les deux vecteurs $\varphi'(t_0)$ et $\varphi''(t_0)$ sont linéairement indépendants.*

Les notions de point régulier, de point birégulier et de point singulier pour une courbe paramétrée ne dépendent pas du paramétrage considéré. On dit que ces notions sont « intrinsèques à la courbe ». Plus précisément, on a la proposition suivante.

[35] Parfois, le terme de point *stationnaire* est employé à la place de point singulier. Ce vocabulaire est utilisé lorsque le paramètre t représente la variable temps. Physiquement, un point stationnaire est un point dont la vitesse (dérivée de la position $\varphi(t)$) est nulle.

Proposition 11.4 *Soient φ et ψ deux arcs paramétrés de classe $\mathcal{C}^1$ (resp. de classe $\mathcal{C}^2$) qui sont $\mathcal{C}^1$-équivalents (resp. $\mathcal{C}^2$-équivalents) et C la courbe paramétrée qui leur est associée.*

Un point $M_0 \in C$ est régulier (resp. birégulier) pour le paramétrage φ si et seulement si il est régulier (resp. birégulier) pour le paramétrage ψ.

Démonstration $\trianglerighteq$ Considérons deux arcs paramétrés $\varphi \in \mathcal{C}^1(I, \mathbb{R}^n)$ et $\psi \in \mathcal{C}^1(J, \mathbb{R}^n)$ qui sont $\mathcal{C}^1$-équivalents. D'après la définition 11.7, il existe un $\mathcal{C}^1$-difféomorphisme θ de I sur J tel que $\varphi = \psi \circ \theta$. Tout point M_0 de la courbe paramétrée C associée à φ et à ψ s'écrit $M_0 = \varphi(t_0)$ avec $t_0 \in I$ et $M_0 = \psi(s_0)$ avec $s_0 = \theta(t_0) \in J$.

En dérivant la relation $\varphi = \psi \circ \theta$, on obtient en $t_0 \in I$

$$\varphi'(t_0) = \psi'(\theta(t_0))\,\theta'(t_0). \tag{3}$$

On a $\psi'(\theta(t_0)) = \psi'(s_0)$ et puisque θ est un difféomorphisme, sa dérivée ne s'annule pas ; la relation (3) fournit donc l'équivalence

$$\varphi'(t_0) = \mathbf{0} \iff \psi'(s_0) = \mathbf{0}.$$

On en déduit que le point M_0 est régulier pour le paramétrage φ si et seulement si il est régulier pour le paramétrage ψ.

$\trianglerighteq$ Considérons à présent deux arcs paramétrés $\varphi \in \mathcal{C}^2(I, \mathbb{R}^n)$ et $\psi \in \mathcal{C}^2(J, \mathbb{R}^n)$, qui sont $\mathcal{C}^2$-équivalents. Il existe donc un $\mathcal{C}^2$-difféomorphisme θ de I sur J tel que $\varphi = \psi \circ \theta$ et tout point M_0 de la courbe paramétrée C associée à φ et à ψ s'écrit $M_0 = \varphi(t_0)$ avec $t_0 \in I$ et $M_0 = \psi(s_0)$ avec $s_0 = \theta(t_0) \in J$. Montrons que si le point M_0 est birégulier pour l'arc paramétré ψ alors il est également birégulier pour l'arc paramétré φ. Pour se faire, utilisons un raisonnement par contraposée[36] et montrons que si le point M_0 n'est pas birégulier pour φ alors il n'est pas non plus birégulier pour ψ.

Supposer que le point M_0 n'est pas birégulier pour φ revient à supposer que les vecteurs $\varphi'(t_0)$ et $\varphi''(t_0)$ sont liés, c'est-à-dire qu'il existe deux réels α et β non tous les deux nuls tels que

$$\alpha\,\varphi'(t_0) + \beta\,\varphi''(t_0) = \mathbf{0}. \tag{4}$$

D'une part, d'après la relation (3) on a $\varphi'(t_0) = \psi'(s_0)\,\theta'(t_0)$ et d'autre part en dérivant deux fois la relation $\varphi = \psi \circ \theta$, on obtient en $t_0 \in I$

$$\varphi''(t_0) = \psi''(\theta(t_0))\,(\theta'(t_0))^2 + \psi'(\theta(t_0))\,\theta''(t_0)$$

autrement dit : $\varphi''(t_0) = \psi''(s_0)\,(\theta'(t_0))^2 + \psi'(s_0)\,\theta''(t_0)$. En substituant ces expressions de $\varphi'(t_0)$ et de $\varphi''(t_0)$ en fonction de $\psi'(s_0)$ et $\psi''(s_0)$ dans la relation (4), on obtient :

$$\big(\alpha\,\theta'(t_0) + \beta\,\theta''(t_0)\big)\,\psi'(s_0) + \beta\,(\theta'(t_0))^2\,\psi''(s_0) = \mathbf{0}. \tag{5}$$

[36] Voir le *Cours de première année* p. 17.

Notons $\lambda = \alpha\theta'(t_0) + \beta\theta''(t_0)$ et $\mu = \beta(\theta'(t_0))^2$. Puisque θ est un difféomorphisme, sa dérivée ne s'annule pas et par conséquent l'un au moins des deux réels α ou β est non nul. La relation (5) permet donc d'affirmer que les deux vecteurs $\psi'(s_0)$ et $\psi''(s_0)$ sont liés, autrement dit que le point M_0 n'est pas birégulier pour ψ.

On a ainsi montré qu'un point birégulier pour un arc paramétré est également un point birégulier pour tous les arcs paramétrés qui lui sont $\mathcal{C}^\infty$-équivalents, ce qui établit l'équivalence énoncée dans la proposition. □

Remarques

1. Puisque d'après la proposition 11.4 la notion de point régulier (resp. de point birégulier) est invariante par changement de paramétrage $\mathcal{C}^1$-équivalent (resp. $\mathcal{C}^2$-équivalent), on dit qu'une courbe paramétrée est régulière (resp. birégulière) si et seulement si il existe un paramétrage φ de C tel que tous les points de C sont réguliers (resp. birréguliers) pour φ.

2. Il résulte de la définition 11.11 qu'un point $M_0 = \varphi(t_0)$ d'une courbe paramétrée C de classe $\mathcal{C}^2$ n'est pas birégulier pour un paramétrage φ de C si les deux vecteurs $\varphi'(t_0)$ et $\varphi''(t_0)$ sont liés. Dans le cas d'une courbe paramétrée dans le plan $\mathbb{R}^2$, les deux vecteurs $\varphi'(t_0)$ et $\varphi''(t_0)$ de $\mathbb{R}^2$ sont liés si et seulement si[37] le déterminant $\det(\varphi'(t_0), \varphi''(t_0))$ est nul. On peut donc déterminer les points birréguliers pour un paramétrage φ d'une courbe paramétrée de classe $\mathcal{C}^2$ en déterminant les valeurs du paramètre t pour lequel $\det(\varphi'(t), \varphi''(t)) \neq 0$. ◇

Les deux arcs paramétrés $\psi : t \in [-1, 1] \mapsto (\cos(2\pi t^2), \sin(2\pi t^2))$ et $\varphi : t \in [0, 2\pi] \mapsto (\cos(t), \sin(t))$ sont des paramétrages du cercle de centre $(0, 0)$ et de rayon 1. Le point $(1, 0)$ est régulier pour φ mais singulier pour ψ. Une telle singularité est qualifiée de cinématique car elle est liée à la façon dont la courbe est paramétrée et non à l'allure de la courbe elle-même. On a établi au cours de l'exercice 11.1 que ces deux arcs paramétrés n'étaient pas C^1-équivalents.
Par ailleurs, un point d'une courbe paramétrée C peut être à la fois régulier et singulier pour un même paramétrage φ. En effet, si l'arc paramétré φ n'est pas injectif alors il possède des points multiples et un point multiple $M_0 = \varphi(t_1) = \varphi(t_2)$ avec $t_1 \neq t_2$ peut être régulier pour la valeur t_1 et singulier pour la valeur t_2 ; une telle situation se produit dans l'exemple 2 ci-après.

Exemples

1. Déterminons les points singuliers de la *quartique piriforme* qui est la courbe paramétrée du plan $\mathbb{R}^2$ dont un paramétrage est

$$\varphi : t \in [0, \pi] \longmapsto (\cos^2(t),\ \cos^3(t)\sin(t)) \in \mathbb{R}^2.$$

[37] Voir le *Cours de première année* p. 465.

Pour tout $t \in [0,\pi]$, on a

$$\begin{aligned}\big(\cos^3(t)\sin(t)\big)' &= -3\sin^2(t)\cos^2(t) + \cos^4(t) = \cos^2(t)\big(-3\sin^2(t)+\cos^2(t)\big)\\ &= \cos^2(t)\big(-3+4\cos^2(t)\big).\end{aligned}$$

La dérivée de φ est donc l'application

$$\varphi' : t \in [0,\pi] \longmapsto \big(-\sin(2t),\ \cos^2(t)(3+4\cos^2(t))\big);$$

elle ne s'annule que pour $t = \pi/2$. La quartique piriforme possède donc un unique point singulier qui est le point $M_0 = \varphi(\pi/2) = (0,0)$. Tous les autres points de cette courbe sont réguliers.

Déterminons à présent les points birréguliers de la quartique piriforme. Pour tout $t \in [0,\pi]$, on a

$$\big(\cos^3(t)\sin(t)\big)'' = 6\cos(t)\sin(t) - 16\cos^3(t)\sin(t) = \sin(2t)\left(3-8\cos^2(t)\right)$$

et la dérivée seconde de φ est

$$\varphi'' : t \in [0,\pi] \longmapsto \big(-2\cos(2t),\ \sin(2t)\,(3-8\cos^2(t))\big).$$

On en déduit que le déterminant $\det(\varphi'(t),\varphi''(t))$ vaut

$$\varphi_1'(t)\varphi_2''(t) - \varphi_2'(t)\varphi_1''(t) = -2\cos^2(t)\left(8\cos^4(t) - 12\cos^2(t) + 3\right).$$

Ce déterminant est nul si $t = \pi/2$ ou si $\cos^2(t)$ est racine du polynôme $8X^2 - 12X + 3$. Ce polynôme admet deux racines réelles qui sont $\lambda_1 = \frac{3-\sqrt{3}}{4}$ et $\lambda_2 = \frac{3+\sqrt{3}}{4}$. Seule λ_1 appartient à $[0,1]$ (on a $\lambda_2 > 1$). Les valeurs de $t \in [0,\pi]$ pour lesquelles $\cos^2(t) = \lambda_1$ (i.e. $\cos(t) = \sqrt{\lambda_1}$ ou $\cos(t) = -\sqrt{\lambda_1}$) sont

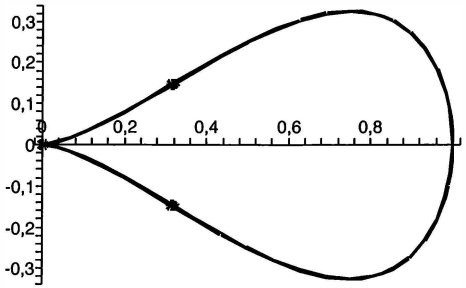
$$t = \arccos\left(\sqrt{\lambda_1}\right) \quad \text{et} \quad t = \arccos\left(-\sqrt{\lambda_1}\right) = \pi - \arccos\left(\sqrt{\lambda_1}\right).$$

Tous les points de la quartique piriforme sont donc birréguliers excepté les trois points $M_0 = \varphi(t)$ où

$$t \in \left\{\frac{\pi}{2}, \pi - \arccos\left(\frac{\sqrt{3-\sqrt{3}}}{2}\right), \arccos\left(\frac{\sqrt{3-\sqrt{3}}}{2}\right)\right\}.$$

On peut utiliser MAPLE pour représenter la quartique piriforme et les trois points qui ne sont pas birréguliers.

```
> phi:=t->[cos(t)^2,cos(t)^3*sin(t)]:
> lambda1:=(3-sqrt(3))/4:
> J:={Pi/2,arccos(sqrt(lambda1)),Pi-arccos(sqrt(lambda1))}:
> des1:=plot([op(phi(t)),t=0..Pi],scaling=constrained):
> des2:=plot([seq([op(phi(s))],s=J)],style=point,symbol=circle):
> display(des1,des2);
```



2. Considérons la courbe paramétrée du plan dont un paramétrage est

$$\varphi : t \in \mathbb{R} \longmapsto \left(t^3(t-2),\quad t^2(t-2)(t-1)\right).$$

L'origine est un point double de la courbe car $(0,0) = \varphi(0) = \varphi(2)$. On a pour tout $t \in \mathbb{R}$,

$$\varphi'(t) = \left(2t^2(2t-1),\quad t(4t^2-9t+4)\right).$$

Puisque $\varphi'(0) = \mathbf{0}$, l'origine est un point singulier de la courbe paramétrée pour le paramétrage φ et puisque $\varphi'(2) = (24,4) \neq \mathbf{0}$ l'origine est également un point régulier pour le paramétrage φ.

PROPOSITION 11.5 *Tout arc paramétré régulier est orientable.*

Démonstration Le raisonnement est très semblable à celui ayant permis de démontrer la proposition 11.3. Considérons un arc régulier $\varphi \in \mathcal{C}^1(I,\mathbb{R}^n)$ et notons γ_φ l'ensemble des arcs paramétrés $\mathcal{C}^1$-équivalents à l'arc φ. Remarquons que l'application $\psi : s \in J \mapsto \varphi(-s)$ où $J = \{-t \in \mathbb{R}\ ;\ t \in I\}$ est un arc paramétré $\mathcal{C}^1$-équivalent à l'arc φ puisque $\varphi = \psi \circ \theta$ avec $\theta : t \in I \mapsto -t \in J$ et l'application θ est très clairement un $\mathcal{C}^1$-difféomorphisme.

D'après la définition 11.9, montrer que l'arc paramétré φ est orientable revient à montrer que dans l'ensemble γ_φ il existe exactement[38] deux classes d'équivalence pour la relation $\mathcal{R}_\mathcal{O}$: « être $\mathcal{O}^1$-équivalent ». Pour cela, nous allons montrer que les deux arcs paramétrés φ et ψ, qui appartiennent tous deux à γ_φ, ne sont pas $\mathcal{O}^1$-équivalents ; on pourra conclure qu'ils appartiennent à deux classes d'équivalence distinctes pour la relation $\mathcal{R}_\mathcal{O}$ dans γ_φ, ce qui établira l'existence de deux classes d'équivalence.

Raisonnons par l'absurde et supposons que les arcs paramétrés φ et ψ sont $\mathcal{O}^1$-équivalents ; d'après la définition 11.8, il existe un difféomorphisme $\widetilde{\theta}$ de J sur I, strictement croissant, tel que $\psi = \varphi \circ \widetilde{\theta}$. Comme par ailleurs on a $\varphi = \psi \circ \theta$, on en déduit que $\varphi = \varphi \circ \widetilde{\theta} \circ \theta$, c'est-à-dire que

$$\forall t \in I \quad \varphi(t) = \varphi(\widetilde{\theta}(\theta(t))). \tag{6}$$

[38] Rappelons que d'après le lemme 11.1 nous savons qu'il existe au plus deux classes d'équivalence.

Notons $f = \widetilde{\theta} \circ \theta$; comme θ est un $\mathcal{C}^1$-difféomorphisme de I sur J et que $\widetilde{\theta}$ est un $\mathcal{C}^1$-difféomorphisme de J sur I, l'application f est une application de classe $\mathcal{C}^1$ de I dans I. D'une part, comme f est une application continue de l'intervalle I dans lui-même, elle admet[39] un point fixe $c \in I$. D'autre part, l'application f est dérivable sur I et la relation (6), qui se réécrit $\varphi(t) = \varphi(f(t))$ pour tout $t \in I$, fournit, en dérivant les deux membres de l'égalité, la relation

$$\forall t \in I \qquad \varphi'(t) = \varphi'(f(t))\, f'(t).$$

En $t = c$, on obtient $\varphi'(c) = \varphi'(c)\, f'(c)$. Puisque l'arc paramétré φ est régulier, on a $\varphi'(c) \neq \mathbf{0}$ et par conséquent $f'(c) = 1$. Cela engendre une contradiction. En effet, comme $\theta : t \in I \mapsto -t \in J$ est une application strictement décroissante et que l'application $\widetilde{\theta}$ est strictement croissante, l'application f est strictement décroissante sur I, donc à dérivée négative ou nulle en c.

On en conclut que les arcs paramétrés φ et ψ ne sont pas $\mathcal{C}^1$-équivalents et par conséquent que γ_ϕ est orientable. □

Remarque Les propositions 11.3 et 11.5 indiquent que tous les arcs paramétrés simples ou réguliers sont orientables. Autrement dit, la plupart des arcs paramétrés « usuels » sont orientables. Néanmoins, l'arc paramétré par

$$t \in [-1, 1] \longmapsto \big(\cos(2\pi t^2),\quad \sin(2\pi t^2)\big)$$

n'est pas un arc simple (puisque $\varphi(-1) = \varphi(1)$), ni un arc régulier (puisque $\varphi'(0) = \mathbf{0}$). La courbe paramétrée associée à cet arc paramétré est un cercle parcouru une fois dans le sens trigonométrique puis une fois dans le sens inverse. ◇

11.3.3 Étude des arcs paramétrés dans le plan

Intéressons-nous à présent plus spécifiquement à l'étude des arcs paramétrés dans le plan. Les paragraphes précédents fournissent un certain nombre de notions permettant d'étudier les arcs paramétrés que nous allons compléter en introduisant des notions qui sont, elles, spécifiques au cas des arcs paramétrés dans le plan.

En cinématique, les courbes paramétrées dans le plan représentent la trajectoire d'un mobile ; le paramètre t désigne alors le temps et le vecteur $\varphi(t)$ la position du mobile dans un repère donné.

Symétries d'une courbe paramétrée dans le plan

L'étude d'une courbe paramétrée dans le plan peut être notablement simplifiée si celle-ci possède des symétries, c'est-à-dire s'il existe des isométries du plan qui laissent la courbe globalement invariante.

[39] Voir l'exercice 13 p. 600 du *Cours de première année*.

Soit $\varphi \in \mathcal{C}^0(I, \mathbb{R}^2)$ un arc paramétré et C la courbe paramétrée du plan qui lui est associée. On cherche à déterminer un intervalle J inclus dans I (de longueur la plus petite possible) tel que l'on puisse déduire les propriétés de la courbe paramétrée $C = \varphi(I)$ et sa construction de celle de $C' = \varphi(J)$.

Supposons qu'il existe une application θ de $\mathbb{R}$ dans $\mathbb{R}$, une application s de $\mathbb{R}^2$ dans $\mathbb{R}^2$ et un intervalle J inclus dans I tel que

- $I = J \cup \theta(J)$,
- $\forall t \in I \quad \varphi(\theta(t)) = s(\varphi(t))$;

on a dans ce cas

$$\varphi(I) = \varphi(J) \cup s(\varphi(J)),$$

autrement dit, la courbe paramétrée $C = \varphi(I)$ est la réunion de la courbe paramétrée $C' = \varphi(J)$ et d'une courbe paramétrée C'' qui se déduit de C' à l'aide de l'application s.

On rencontre souvent pour la fonction θ l'une des formes suivantes :

- $I = [-a, a]$ (où $a \in \mathbb{R}_+^*$) et $\theta : t \in I \mapsto -t \in I$ et dans ce cas on peut prendre $J = [0, a]$;
- $I = [a, b]$ (où $(a, b) \in \mathbb{R}^2, a < b$) et $\theta : t \in I \mapsto a + b - t \in I$ et dans ce cas on peut prendre $J = [a, \frac{1}{2}(a + b)]$;
- $I = \mathbb{R}_+^*$ et $\theta : t \in I \mapsto 1/t \in I$ et dans ce cas on peut prendre $J =]0, 1]$.

La fonction s est quant à elle fréquemment de l'une des formes suivantes :

- $s : (x, y) \in \mathbb{R}^2 \mapsto (x, y) \in \mathbb{R}^2$ et dans ce cas $C'' = C'$;
- $s : (x, y) \in \mathbb{R}^2 \mapsto (-x, y) \in \mathbb{R}^2$ et dans ce cas la courbe C'' est l'image de la courbe C' par la symétrie orthogonale par rapport à l'axe des ordonnées ;
- $s : (x, y) \in \mathbb{R}^2 \mapsto (x, -y) \in \mathbb{R}^2$ et dans ce cas la courbe C'' est l'image de la courbe C' par la symétrie orthogonale par rapport à l'axe des abscisses ;
- $s : (x, y) \in \mathbb{R}^2 \mapsto (-x, -y) \in \mathbb{R}^2$ et dans ce cas la courbe C'' est l'image de la courbe C' par la symétrie centrale de centre l'origine 0 du repère ;
- $s : (x, y) \in \mathbb{R}^2 \mapsto (y, x) \in \mathbb{R}^2$ et dans ce cas la courbe C'' est l'image de la courbe C' par la symétrie orthogonale par rapport la droite $y = x$ (première bissectrice) ;
- $s : (x, y) \in \mathbb{R}^2 \mapsto (-y, -x) \in \mathbb{R}^2$ et dans ce cas la courbe C'' est l'image de la courbe C' par la symétrie orthogonale par rapport la droite $y = -x$ (deuxième bissectrice) ;
- $s : (x, y) \in \mathbb{R}^2 \mapsto (-y, x) \in \mathbb{R}^2$ et dans ce cas la courbe C'' est l'image de la courbe C' par la rotation de centre l'origine O du repère et d'angle[(40)] $\frac{\pi}{2}$;
- $s : (x, y) \in \mathbb{R}^2 \mapsto (y, -x) \in \mathbb{R}^2$ et dans ce cas la courbe C'' est l'image de la courbe C' par la rotation de centre l'origine O du repère et d'angle[(40)] $-\frac{\pi}{2}$;

(40) Mesuré dans le sens direct ou trigonométrique (sens opposé au sens de parcours des aiguilles d'une montre).

- $s : (x, y) \in \mathbb{R}^2 \mapsto (x + u_1, y + u_2) \in \mathbb{R}^2$ avec $(u_1, u_2) \in \mathbb{R}^2$ et dans ce cas la courbe C'' est l'image de la courbe C' par la translation de vecteur $\mathbf{u} = (u_1, u_2)$.

Il arrive assez souvent que la courbe paramétrée étudiée possède plusieurs symétries et que l'on soit amené à utiliser plusieurs fois consécutivement cette stratégie de réduction de l'intervalle d'étude.

Exemples

1. Considérons l'arc paramétré $\varphi : t \in \mathbb{R} \mapsto (\cos(t), \sin(\frac{1}{3}t))$. On peut commencer par remarquer que la première fonction coordonnée φ_1 est une application 2π-périodique et que la seconde fonction coordonnée φ_2 est une application 6π-périodique de sorte que

$$\forall t \in \mathbb{R} \quad \varphi(t + 6\pi) = \varphi(t).$$

On s'intéresse donc à l'arc paramétré φ en restriction à un intervalle de période I de longueur 6π. Notons $x(t) = \cos(t)$, $y(t) = \sin(\frac{1}{3}t)$ et C la courbe paramétrée qui est l'image de I par φ. En prenant $I = [-3\pi, 3\pi]$, on a

$$\forall t \in I \qquad \big(-t \in I \quad \text{et} \quad x(-t) = x(t) \quad \text{et} \quad y(-t) = -y(t) \big)$$

ce qui signifie que la courbe paramétrée C admet l'axe des abscisses comme axe de symétrie. On a ici $\theta : t \in I \mapsto -t \in I$ et $s : (x, y) \in \mathbb{R}^2 \mapsto (x, -y) \in \mathbb{R}^2$ et en restreignant l'étude de l'arc paramétré φ à l'intervalle $J_1 = [0, 3\pi]$, on pourra construire C à partir de $C_1 = \varphi(J_1)$ en effectuant une symétrie orthogonale par rapport à l'axe des abscisses.

Considérons à présent l'application $\theta : t \in [0, 3\pi] \mapsto 3\pi - t \in [0, 3\pi]$; on a pour tout $t \in [0, 3\pi]$,

$$\varphi(\theta(t)) = \begin{pmatrix} x(3\pi - t) \\ y(3\pi - t) \end{pmatrix} = \begin{pmatrix} \cos(\pi - t) \\ \sin(\pi - \frac{t}{3}) \end{pmatrix} = \begin{pmatrix} -\cos(t) \\ \sin(\frac{t}{3}) \end{pmatrix} = s(\varphi(t))$$

où $s : (x, y) \in \mathbb{R}^2 \mapsto (-x, y) \in \mathbb{R}^2$. De plus, en posant $J_2 = [0, \frac{3}{2}\pi]$, on a $J_1 = [0, \frac{3}{2}\pi] \cup [\frac{3}{2}\pi, 3\pi] = J_2 \cup \theta(J_2)$, donc la courbe $C_1 = \varphi(J_1)$ peut être construite à partir de la courbe $C_2 = \varphi(J_2)$ par la symétrie orthogonale par rapport à l'axe des ordonnées.

En résumé, on peut se contenter d'étudier l'arc paramétré φ sur $[0, \frac{3}{2}\pi]$; la courbe paramétrée correspondant à $C = \varphi(I) = \varphi(\mathbb{R})$ s'obtient à partir de la courbe $C_2 = \varphi([0, \frac{3}{2}\pi])$ en effectuant successivement une symétrie orthogonale par rapport à l'axe des ordonnées puis une symétrie orthogonale par rapport à l'axe des abscisses comme cela est illustré sur la figure 5.

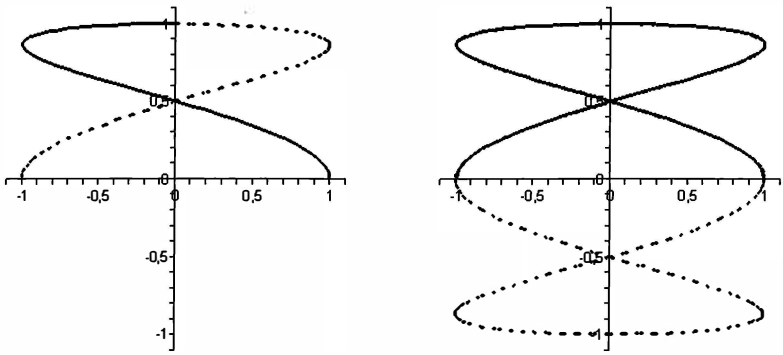


Fig. 5 Construction de la courbe paramétrée $\varphi(\mathbb{R})$ à partir de $\varphi([0, \frac{3}{2}\pi])$.

2. Intéressons-nous à l'arc paramétré $\varphi : t \in [-\pi, 3\pi] \mapsto (\frac{1}{4}t + \sin(t),\ \sin(t))$. Notons $x(t) = \frac{1}{4}t + \sin(t)$, $y(t) = \sin(t)$ et C la courbe paramétrée qui est l'image de $I = [-\pi, 3\pi]$ par φ.

Considérons l'application $\theta : t \in [-\pi, \pi] \mapsto t + 2\pi \in [\pi, 3\pi]$; en posant $J_1 = [-\pi, \pi]$, on a $I = [-\pi, \pi] \cup [\pi, 3\pi] = J_1 \cup \theta(J_1)$ et pour tout $t \in [-\pi, \pi]$,

$$\varphi(\theta(t)) = \begin{pmatrix} x(t+2\pi) \\ y(t+2\pi) \end{pmatrix} = \begin{pmatrix} \frac{\pi}{2} + \frac{1}{4}t + \sin(t) \\ \sin(t) \end{pmatrix} = \begin{pmatrix} \frac{\pi}{2} + x(t) \\ y(t) \end{pmatrix} = s(\varphi(t))$$

où $s : (x, y) \in \mathbb{R}^2 \mapsto (\frac{1}{2}\pi + x, y) \in \mathbb{R}^2$. On en déduit que la courbe paramétrée $C = \varphi(I)$ peut être déduite de la courbe $C_1 = \varphi(J_1)$ par la translation de vecteur $\mathbf{u} = (\frac{1}{2}\pi, 0)$.

Considérons à présent l'application $\theta : t \in [-\pi, \pi] \mapsto -t \in [-\pi, \pi]$; en posant $J_2 = [0, \pi]$, on a $J_1 = [-\pi, 0] \cup [0, \pi] = J_2 \cup \theta(J_2)$ et pour tout $t \in [0, \pi]$,

$$\varphi(\theta(t)) = \begin{pmatrix} x(-t) \\ y(-t) \end{pmatrix} = \begin{pmatrix} -\frac{1}{4}t - \sin(t) \\ -\sin(t) \end{pmatrix} = \begin{pmatrix} -x(t) \\ -y(t) \end{pmatrix} = s(\varphi(t))$$

où $s : (x, y) \in \mathbb{R}^2 \mapsto (-x, -y) \in \mathbb{R}^2$. On en déduit que la courbe paramétrée $C_1 = \varphi(J_1)$ peut être déduite de la courbe $C_2 = \varphi(J_2)$ par la symétrie axiale dont le centre est l'origine du repère.

En résumé, on peut se contenter d'étudier l'arc paramétré φ sur $[0, \pi]$; la courbe paramétrée correspondant à $\varphi([-\pi, 3\pi])$ s'obtient à partir de la courbe $\varphi([0, \pi])$ en effectuant successivement une symétrie axiale dont le centre est l'origine du repère puis translation de vecteur $\mathbf{u} = (\frac{1}{2}\pi, 0)$ comme cela est illustré sur la figure 6.

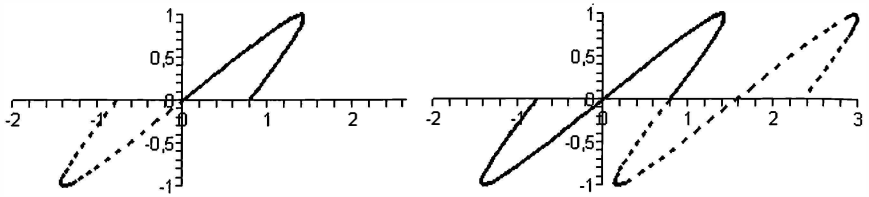


Fig. 6 Construction de la courbe paramétrée $\varphi([-\pi, 3\pi])$ à partir de $\varphi([0, \pi])$.

Tangente à une courbe paramétrée dans le plan

Soit $\varphi \in \mathcal{C}^\infty(I, \mathbb{R}^2)$ un arc paramétré[41] et $t_0 \in I$ une valeur du paramètre tel que l'une au moins des dérivées successives de φ en t_0 ne soit pas nulle. On note alors p l'entier défini par[42] $p = \min\left\{j \in \mathbb{N}^* \ ; \ \varphi^{(j)}(t_0) \neq \mathbf{0}\right\}$.

DÉFINITION 11.12 (Tangente à une courbe paramétrée)
Soient $\varphi \in \mathcal{C}^\infty(I, \mathbb{R}^2)$ un arc paramétré, C la courbe paramétrée qui lui est associée et $M_0 = \varphi(t_0)$, où $t_0 \in I$, un point de C.
On appelle tangente *à la courbe paramétrée C au point M_0, la droite passant par M_0 et dirigée par le vecteur $\varphi^{(p)}(t_0)$ où $p = \min\left\{j \in \mathbb{N}^* \ ; \ \varphi^{(j)}(t_0) \neq \mathbf{0}\right\}$.*

La droite tangente à la courbe paramétrée C au point $M_0 = \varphi(t_0)$ peut être vue[43] comme la droite qui correspond à la limite des droites passant par les points M_0 et $M = \varphi(t)$ lorsque t tend vers t_0, voir la figure 7. Un vecteur directeur unitaire à cette droite tangente est alors la limite lorsque t tend vers t_0 du vecteur directeur unitaire[44] $\mathbf{u}(t) = \overrightarrow{M_0M}/\|\overrightarrow{M_0M}\|$ de la droite (M_0M).

Établissons cette affirmation. La formule de Taylor pour l'application φ à l'ordre p au voisinage de $t_0 \in I$ s'écrit[45] :

$$\varphi(t) = \sum_{k=0}^{p} \frac{(t-t_0)^k}{k!}\,\varphi^{(k)}(t_0) + (t-t_0)^p\,\varepsilon(t)$$

---

[41] On suppose que l'application φ est de classe $\mathcal{C}^\infty$ afin de pouvoir la dériver autant que nécessaire par la suite ; toutefois, l'étude qui va être menée ici peut s'effectuer dès que l'on est en mesure de définir les entiers p et q introduits ultérieurement.

[42] On suppose par la suite qu'un tel entier existe toujours. On remarquera tout de même que ceci n'est pas toujours le cas. Par exemple en considérant la courbe paramétrée par $\varphi : t \in \mathbb{R} \mapsto (\mathrm{e}^{-t^2}, \mathrm{e}^{-t^2})$ et la valeur $t_0 = 0$ on a pour tout $j \in \mathbb{N}^*$ l'égalité $\varphi^{(j)}(t_0) = 0$.

[43] Ou même définie.

[44] Il convient de considérer un vecteur unitaire à ces droites car le vecteur directeur de la droite (M_0M) a une norme qui tend vers 0 quand t tend vers t_0.

[45] L'application φ est une application de I à valeurs dans $\mathbb{R}^2$. Les deux applications composantes φ_1 et φ_2 sont deux applications de $I \subset \mathbb{R}$ dans $\mathbb{R}$ pour lesquels on peut utiliser la formule de Taylor établie dans le *Cours de première année*, p. 778.

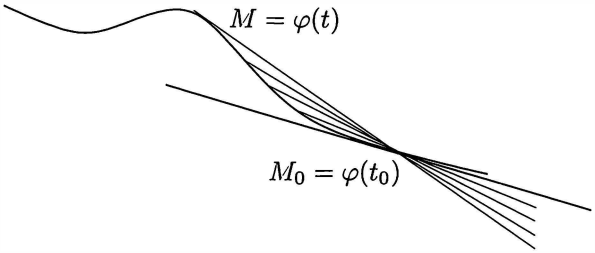


Fig. 7 Interprétation graphique de la notion de droite tangente.

où ε est une fonction de $\mathbb{R}$ dans $\mathbb{R}^2$ de limite nulle en t_0. Puisque toutes les dérivées d'ordre inférieur à p sont nulles en t_0, on a

$$\varphi(t) = \varphi(t_0) + \frac{(t-t_0)^p}{p!}\varphi^{(p)}(t_0) + (t-t_0)^p\,\varepsilon(t).$$

On en déduit que

$$\mathbf{u}(t) = \frac{\overrightarrow{M_0M}}{\|\overrightarrow{M_0M}\|} = \frac{\varphi(t)-\varphi(t_0)}{\|\varphi(t)-\varphi(t_0)\|} = \frac{(t-t_0)^p}{|(t-t_0)^p|}\,\frac{\varphi^{(p)}(t_0)+p!\,\varepsilon(t)}{\|\varphi^{(p)}(t_0)+p!\,\varepsilon(t)\|}.$$

Si p est pair, on a $\lim\limits_{t\to t_0}\mathbf{u}(t) = \dfrac{\varphi^{(p)}(t_0)}{\|\varphi^{(p)}(t_0))\|}$ et si p est impair, on a

$$\lim_{\substack{t\to t_0\\ t>t_0}}\mathbf{u}(t) = \frac{\varphi^{(p)}(t_0)}{\|\varphi^{(p)}(t_0))\|} \quad\text{et}\quad \lim_{\substack{t\to t_0\\ t<t_0}}\mathbf{u}(t) = -\frac{\varphi^{(p)}(t_0)}{\|\varphi^{(p)}(t_0))\|}.$$

Dans tous les cas, on peut en conclure que $\varphi^{(p)}(t_0)$, que nous avons défini comme étant vecteur directeur de la tangente en M_0 à la courbe paramétrée C, est un vecteur directeur de la droite qui correspond à la limite des droites passant par les points M_0 et $M = \varphi(t)$ lorsque t tend vers t_0.

La tangente à une courbe paramétrée est une notion intrinsèque : elle ne dépend pas du paramétrage de la courbe choisi comme l'indique la proposition suivante.

PROPOSITION 11.6 *Soient C une courbe paramétrée dans le plan et M_0 un point de C. La tangente à la courbe paramétrée C au point M_0 est indépendante du choix du paramétrage $\mathcal{C}^1$-équivalent de C considéré.*

Démonstration Considérons deux paramétrages $\varphi \in \mathcal{C}^\infty(I, \mathbb{R}^2)$ et $\psi \in \mathcal{C}^\infty(J, \mathbb{R}^2)$ de la courbe paramétrée C qui sont $\mathcal{C}^1$-équivalents. Il existe un $\mathcal{C}^1$-difféomorphisme θ de I sur J tel que $\varphi = \psi \circ \theta$ et un point M_0 de la courbe paramétrée C est décrit par $M_0 = \varphi(t_0)$ avec $t_0 \in I$ et par $M_0 = \psi(s_0)$ avec $s_0 = \theta(t_0) \in J$. On montre en utilisant un raisonnement par récurrence qu'en dérivant r fois (pour $r \in \mathbb{N}^*$) la relation $\varphi = \psi \circ \theta$, on obtient[46]

$$\varphi^{(r)}(t) = \psi^{(r)}(\theta(t))\,(\theta'(t))^r + \sum_{i=1}^{r-1} \psi^{(i)}(\theta(t))\, f_{i,r}(t),$$

où $f_{i,r}$ désigne une application s'exprimant uniquement à l'aide de l'application θ et de ses dérivées. En particulier puisque $\theta(t_0) = s_0$, on a

$$\varphi^{(r)}(t_0) = \psi^{(r)}(s_0)\,(\theta'(t_0))^r + \sum_{i=1}^{r-1} \psi^{(i)}(s_0)\, f_{i,r}(t_0). \tag{7}$$

Soit $p = \min\left\{j \in \mathbb{N}^* \ ; \ \psi^{(j)}(t_0) \neq \mathbf{0}\right\}$. La tangente à la courbe paramétrée C au point $M_0 = \psi(s_0)$ est donc la droite passant par M_0 et dirigée par le vecteur $\psi^{(p)}(s_0)$. La relation (7) indique que

$$\forall r \in \{1, \ldots, p-1\} \quad \varphi^{(r)}(t_0) = \mathbf{0} \quad \text{et} \quad \varphi^{(p)}(t_0) = \psi^{(p)}(s_0)\,(\theta'(t_0))^p.$$

Comme θ est un difféomorphisme, sa dérivée ne s'annule pas ; on a $\theta'(t_0) \neq 0$ et par conséquent $\varphi^{(p)}(t_0) \neq \mathbf{0}$. On a $M_0 = \varphi(t_0)$ et les vecteurs $\varphi^{(p)}(t_0)$ et $\psi^{(p)}(s_0)$ sont colinéaires : la tangente à la courbe paramétrée C au point $M_0 = \varphi(t_0)$ est donc également la droite passant par M_0 et dirigée par le vecteur $\varphi^{(p)}(t_0)$. En d'autres termes, quel que soit le paramétrage de la courbe paramétrée C considéré, on obtient pour tangente à la courbe paramétrée C au point M_0 la même droite. □

Remarques

1. Si $M_0 = \varphi(t_0)$ est un point régulier de la courbe paramétrée C dont un paramétrage est $\varphi \in \mathcal{C}^1(I, \mathbb{R}^2)$ alors $\varphi'(t_0) \neq \mathbf{0}$ et l'entier p de la définition 11.12 vaut 1. La tangente $\mathcal{T}$ à la courbe en M_0 est la droite passant par M_0 et dirigée par le vecteur $\varphi'(t_0)$; on a

$$\mathcal{T} = \left\{\mathbf{x} \in \mathbb{R}^2 \ ; \ \exists \alpha \in \mathbb{R} \quad \mathbf{x} = \varphi(t_0) + \alpha\,\varphi'(t_0)\right\}.$$

2. Considérons une courbe paramétrée C qui coïncide avec le graphe d'une fonction f de $\mathbb{R}$ dans $\mathbb{R}$ qui est dérivable ; un paramétrage de la courbe C est $\varphi : t \in \mathbb{R} \mapsto (t, f(t))$. Tous les points de la courbe C sont réguliers car $\varphi'(t) = (1, f'(t))$ n'est jamais nul ; on a $p = 1$. On retrouve les résultats concernant la tangente au graphe d'une fonction de la variable réelle[47] : $f'(t_0)$ est la pente de la tangente au point $M = (t_0, f(t_0))$ et la remarque précédente indique que l'équation de la tangente est $y = f(t_0) + f'(t_0)\,(x - t_0)$.

---

[46] On pourra se reporter à la démonstration de la proposition 11.4 pour la forme explicite de cette relation dans le cas où $r = 1$ et où $r = 2$.

[47] Voir le *Cours de première année* p. 714.

3. Un point multiple M_0 d'une courbe paramétrée C peut posséder plusieurs tangentes distinctes. C'est le cas d'un point double $M_0 = \varphi(t_1) = \varphi(t_2)$, où $\varphi \in \mathcal{C}^\infty(I, \mathbb{R}^2)$ désigne un paramétrage de C et t_1, t_2 deux valeurs distinctes du paramètre, si les deux vecteurs $\varphi^{(p_1)}(t_1)$ et $\varphi^{(p_2)}(t_2)$ où $p_1 = \min\left\{j \in \mathbb{N}^* \ ; \ \varphi^{(j)}(t_1) \neq \mathbf{0}\right\}$ et $p_2 = \min\left\{j \in \mathbb{N}^* \ ; \ \varphi^{(j)}(t_2) \neq \mathbf{0}\right\}$ ne sont pas colinéaires. ◇

Exemples

1. Considérons la *quartique piriforme* dont un paramétrage est

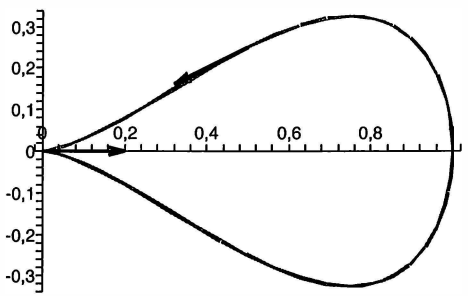
$$\varphi : t \in [0, \pi[\mapsto \left(\cos^2(t), \ \cos^3(t)\sin(t)\right) \in \mathbb{R}^2.$$

Le point $M_0 = \varphi(\frac{\pi}{4}) = (\frac{1}{2}, \frac{1}{4})$ est un point régulier de la courbe pour lequel on a $\varphi'(\frac{\pi}{4}) = (-1, -\frac{1}{2})$. La tangente à la quartique piriforme au point M_0 est donc la droite de vecteur directeur $(-1, -\frac{1}{2})$ passant par $(\frac{1}{2}, \frac{1}{4})$; elle a pour équation $-\frac{1}{2}x + y = 0$.

Déterminons la tangente en $\varphi(\frac{\pi}{2}) = (0, 0)$. On a $\varphi'(\frac{\pi}{2}) = \mathbf{0}$ et $\varphi''(\frac{\pi}{2}) = (2, 0)$. On en déduit que $p = \min\left\{j \in \mathbb{N}^* \ ; \ \varphi^{(j)}(\frac{\pi}{2}) \neq \mathbf{0}\right\} = 2$. La tangente à l'origine à la quartique piriforme est la droite passant par l'origine et de vecteur directeur le vecteur de composantes $(2, 0)$. Cette droite coïncide avec l'axe des abscisses.

On peut avoir recours à MAPLE pour représenter la quartique piriforme ainsi que les vecteurs directeurs aux tangentes en $M_0 = (\frac{1}{2}, \frac{1}{4})$ et à l'origine.

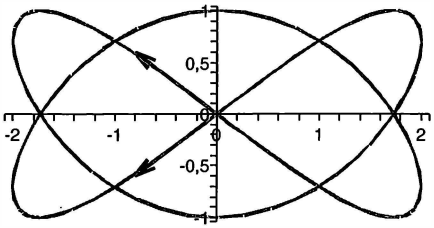
```
> with(plots):
> phi:=t->[cos(t)^2,cos(t)^3*sin(t)]:
> des1:=plot([phi(t)[1],phi(t)[2],t=0..Pi],scaling=constrained):
> t0:=Pi/4: v:=op(subs(t=t0,diff([phi(t)],t$1))): eval(%);
                                -1, -1/2
> des2 := arrow(<phi(t0)>,<v>, shape=arrow, length=0.2):
> t0:=Pi/2:  v:=op(subs(t=t0,diff([phi(t)],t$2))): eval(%);
                                 2, 0
> des3 := arrow(<phi(t0)>,<v>, shape=arrow, length=0.2):
> display(des1,des2,des3);
```



2. Considérons l'arc paramétré $\varphi : t \in [0, 4\pi[\mapsto (2\sin(t),\ \cos(\frac{3}{2}t))$ dont l'image C est une courbe de Lissajous dans $\mathbb{R}^2$. On a établi p. 744 que l'origine est un point double de la courbe qui est l'image par φ des valeurs π et 3π du paramètre. On a $\varphi'(\pi) = (-2, \frac{3}{2})$ et $\varphi'(3\pi) = (-2, -\frac{3}{2})$. Il existe donc deux tangentes à la courbe paramétrée C à l'origine qui ont pour équation $\frac{3}{2}x + 2y = 0$ et $-\frac{3}{2}x + 2y = 0$

On peut avoir recours à MAPLE pour représenter cette courbe de Lissajous ainsi que les vecteurs directeurs aux deux tangentes à l'origine.

```
> phi:=t->[2*sin(t),cos(3*t/2)]:
> dphi:=unapply(diff(phi(t),t),t):
> des1:=plot([phi(t)[1],phi(t)[2],t=0..4*Pi],scaling=constrained):
> des2:=arrow(<phi(Pi)>,<op(dphi(Pi))>,shape=arrow,length=1):
> des3:=arrow(<phi(3*Pi)>,<op(dphi(3*Pi))>,shape=arrow,length=1):
> display(des1,des2,des3);
```



Repère local d'une courbe paramétrée dans le plan

L'entier p étant donné par la définition 11.12, on suppose qu'il existe une dérivée $\varphi^{(j)}(t_0)$ (avec $j > p$) de φ telle que la famille engendrée par $\varphi^{(p)}(t_0)$ et cette nouvelle dérivée soit une famille libre[48]. Notons q l'entier défini par :

$$q = \min\left\{j \in \mathbb{N}^* \ ;\ \ j > p \quad \text{et} \quad \{\varphi^{(p)}(t_0), \varphi^{(j)}(t_0)\} \text{ est libre}\right\}.$$

Remarquons que puisque[49] l'espace $\mathbb{R}^2$ est de dimension 2, la famille libre $\{\varphi^{(p)}(t_0), \varphi^{(q)}(t_0)\}$ forme une base de $\mathbb{R}^2$.

[48] Voir la définition 8.8 p. 317 du *Cours de première année.*

[49] Voir la proposition 8.11 p. 330 du *Cours de première année.*

DÉFINITION 11.13 (Repère local d'une courbe dans le plan)
Soient $\varphi \in \mathcal{C}^\infty(I, \mathbb{R}^2)$ un arc paramétré du plan et C la courbe paramétrée qui lui est associée. Soit $t_0 \in I$ une valeur du paramètre tel que les entiers

$$p = \min\{j \in \mathbb{N}^* \ ; \ \varphi^{(j)}(t_0) \neq \mathbf{0}\} \quad \textit{et}$$
$$q = \min\left\{j \in \mathbb{N}^* \ ; \ j > p \quad \textit{et} \quad \{\varphi^{(p)}(t_0), \varphi^{(j)}(t_0)\} \ \textit{est libre}\right\}$$

soient définis. On appelle repère local *au point $M_0 = \varphi(t_0)$ de la courbe paramétrée C le repère d'origine M_0 dont les vecteurs de la base sont $\varphi^{(p)}(t_0)$ et $\varphi^{(q)}(t_0)$.*

Remarques

1. On montre[50] que la notion de repère local d'une courbe paramétrée est une notion intrinsèque : le repère ne dépend pas du paramétrage C^1-équivalent choisi.

2. On remarquera que d'après la définition 11.12, le premier vecteur de base du repère local est le vecteur directeur de la tangente à la courbe paramétrée au point considéré.

3. Si $M_0 = \varphi(t_0)$ est un point birégulier de la courbe paramétrée C dont un paramétrage est φ alors d'après la définition 11.11 on a $\varphi'(t_0) \neq \mathbf{0}$ et la famille $\{\varphi'(t_0), \varphi''(t_0)\}$ est libre ; on a donc dans ce cas $p = 1$ et $q = 2$. Ainsi, le repère local à la courbe paramétrée C en un point birégulier $M_0 = \varphi(t_0)$ est $(M_0; \varphi'(t_0), \varphi''(t_0))$. ◇

Le repère local d'une courbe paramétrée en un point n'est en général pas un repère orthonormé direct.

Exemples

1. Déterminons le repère local en $M_0 = \varphi(\pi) = (0, 0)$ à la courbe *cardioïde* dont un paramétrage est[51]

$$\varphi : t \in [-\pi, \pi] \longmapsto \big((1 + \cos(t))\cos(t), \quad (1 + \cos(t))\sin(t)\big).$$

Les dérivées successives de φ en π sont :

$$\varphi'(\pi) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \varphi''(\pi) = \begin{pmatrix} -1 \\ 0 \end{pmatrix} \quad \text{et} \quad \varphi'''(\pi) = \begin{pmatrix} 0 \\ -3 \end{pmatrix}.$$

(50) Ce résultat se démontre en utilisant un raisonnement très semblable à celui utilisé pour démontrer la proposition 11.4 ou la proposition 11.6.

(51) Le nom de cette courbe paramétrée est liée à sa forme et provient du terme grec « kardia » qui signifie « cœur ».

Avec les notations de la définition 11.13, on a $p = 2$. Le point $M_0 = (0,0)$ est donc un point singulier et la tangente à la courbe en M_0 est la droite passant par ce point et dirigée par le vecteur de composantes $(-1,0)$. Pour obtenir la valeur de l'entier q, déterminons si les vecteurs $\varphi''(\pi)$ et $\varphi'''(\pi)$ forment un système libre ou non. Pour cela[52], calculons le déterminant $\det(\varphi''(\pi), \varphi'''(\pi))$. On a

$$\det(\varphi''(\pi), \varphi'''(\pi)) = \begin{vmatrix} -1 & 0 \\ 0 & -3 \end{vmatrix} = 3.$$

Comme ce déterminant est non nul, les vecteurs $\varphi''(\pi)$ et $\varphi'''(\pi)$ forment un système libre, donc on a $q = 3$. Le repère local de la courbe paramétrée C au point $M_0 = (0,0)$ est donc donné par $(M_0; \varphi''(\pi), \varphi'''(\pi))$.

2. Intéressons-nous maintenant au point $M_0 = \varphi(\frac{1}{2}\pi) = (0,1)$. Les dérivées successives de φ en $\frac{\pi}{2}$ sont :

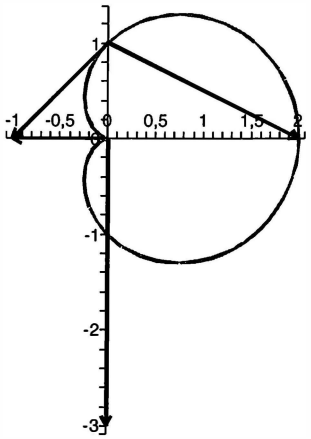
$$\varphi'(\tfrac{1}{2}\pi) = \begin{pmatrix} -1 \\ -1 \end{pmatrix} \quad \text{et} \quad \varphi''(\tfrac{1}{2}\pi) = \begin{pmatrix} 2 \\ -1 \end{pmatrix}.$$

On a donc $p = 1$ et $q = 2$ et le point $M_0 = (0,1)$ est un point birégulier. Le repère local de la courbe paramétrée C au point $(0,1)$ est $(M_0; \varphi'(\frac{1}{2}\pi), \varphi''(\frac{1}{2}\pi))$.

On peut avec MAPLE représenter la courbe paramétrée C ainsi que ses repères locaux aux points $(0,0)$ et $(0,1)$.

```
> phi:= t-> [(1+cos(t))*cos(t),(1+cos(t))*sin(t)]:
> Dn:= proc(f,n,t0)
  description "calcul la dérivée n-ieme de la fonction f en t0";
    local dnf;
    dnf:=diff(f(t),t$n);
    eval(subs(t=t0,dnf));
  end proc:
> des1:=plot([op(phi(t)),t=-Pi..Pi],scaling=constrained):
> t0:=Pi/2: des2:=arrow({phi(t0)},{Dn(phi,1,t0), Dn(phi,2,t0)}):
> t0:=Pi: des3:=arrow({phi(t0)},{Dn(phi,2,t0), Dn(phi,3,t0)}):
> display(des1,des2,des3);
```

[52] Voir le *Cours de première année* p. 465.



Exercice 11.4 Déterminer le repère local en $M_0 = \varphi(t_0)$ à la courbe paramétrée C dont un paramétrage est $\varphi : t \in [-\pi, \pi] \mapsto (\cos(t), t^3)$ pour $t_0 = 0$ et $t_0 = 2$.

Allure d'une courbe paramétrée au voisinage d'un de ses points

Nous allons voir que la connaissance des entiers p et q, donnés par la définition 11.13, pour un point d'une courbe paramétrée fournit des informations sur l'allure de cette courbe au voisinage du point considéré. Supposons que l'on connaisse le couple (p, q) en un point $M_0 = \varphi(t_0)$ d'une courbe paramétrée dont un paramétrage est $\varphi \in \mathcal{C}^\infty(I, \mathbb{R}^2)$; la formule de Taylor pour l'application φ à l'ordre q au voisinage de $t_0 \in I$ s'écrit[53]

$$\varphi(t) = \sum_{k=0}^{q} \frac{(t-t_0)^k}{k!} \varphi^{(k)}(t_0) + (t-t_0)^q \varepsilon(t)$$

où ε est une fonction de $\mathbb{R}$ dans $\mathbb{R}^2$ de limite nulle en t_0. Puisque toutes les dérivées d'ordre inférieur à p sont nulles en t_0 et que les dérivées d'ordre $j \in \{p+1, \ldots, q-1\}$ forment avec la dérivée d'ordre p un système lié, autrement dit puisque

$$\begin{aligned} &\forall i \in \{1, \ldots, p-1\} \quad \varphi^{(i)}(t_0) = \mathbf{0}, \\ \text{et} \quad &\forall j \in \{p+1, \ldots, q-1\} \quad \exists \lambda_j \in \mathbb{R} \quad \varphi^{(j)}(t_0) = \lambda_j \varphi^{(p)}(t_0), \end{aligned}$$

[53] L'application φ est une application de I à valeurs dans $\mathbb{R}^2$. Les deux applications composantes φ_1 et φ_2 sont deux applications de $I \subset \mathbb{R}$ dans $\mathbb{R}$ pour lesquels on peut utiliser la formule de Taylor établie dans le *Cours de première année*, p. 778.

la formule de Taylor prend la forme :

$$\varphi(t) = \varphi(t_0) + \left(\frac{(t-t_0)^p}{p!} + \sum_{j=p+1}^{q-1} \lambda_j \frac{(t-t_0)^j}{j!}\right) \varphi^{(p)}(t_0) + \frac{(t-t_0)^q}{q!} \varphi^{(q)}(t_0) + (t-t_0)^q \varepsilon(t).$$

Les coordonnées $(X(t), Y(t))$ d'un point $M(t) = \varphi(t)$ dans le repère local $(\varphi(t_0); \varphi^{(p)}(t_0), \varphi^{(q)}(t_0))$ de la courbe au point $M_0 = \varphi(t_0)$ vérifient donc

$$\begin{cases} X(t) = \dfrac{(t-t_0)^p}{p!} + \displaystyle\sum_{j=p+1}^{q-1} \lambda_j \frac{(t-t_0)^j}{j!} + o_{t_0}\big((t-t_0)^q\big) \\ Y(t) = \dfrac{(t-t_0)^q}{q!} + o_{t_0}\big((t-t_0)^q\big) \end{cases} \tag{8}$$

Comme le terme $\sum_{j=p+1}^{q-1} \lambda_j \frac{(t-t_0)^j}{j!}$ est négligeable[54] devant $\frac{(t-t_0)^p}{p!}$ lorsque t est au voisinage de t_0, la courbe paramétrée C, au voisinage du point $M_0 = \varphi(t_0)$, est « proche » de la courbe paramétrée dont un paramétrage dans le repère local $(\varphi(t_0); \varphi^{(p)}(t_0), \varphi^{(q)}(t_0))$ est

$$\widetilde{\varphi} \ : \ t \in J \longmapsto \left(\frac{(t-t_0)^p}{p!}, \frac{(t-t_0)^q}{q!}\right)$$

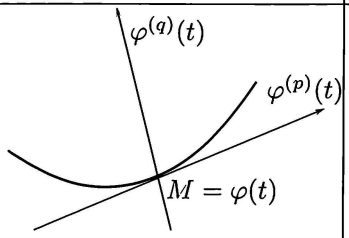
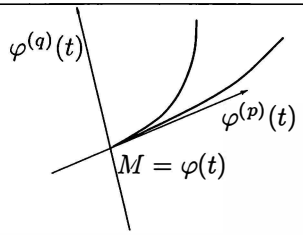
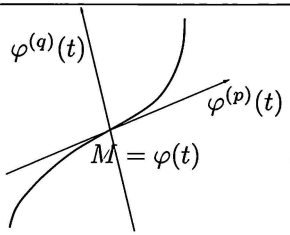
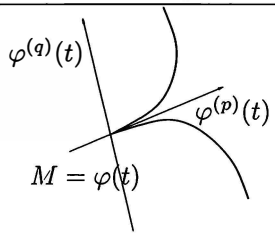
où J désigne un voisinage de t_0.

L'arc paramétré $\widetilde{\varphi}$ est $\mathcal{C}^\infty$-équivalent à l'arc paramétré $\psi : s \in K \mapsto (s^p, s^q)$ où K désigne un voisinage de 0 dans $\mathbb{R}$. L'intérêt de considérer l'arc paramétré ψ plutôt que $\widetilde{\varphi}$ est que son étude est plus facile que celle de l'arc paramétré $\widetilde{\varphi}$. En effet, en désignant par X et Y les coordonnées du point $\psi(t)$ dans le repère local $(\varphi(t_0); \varphi^{(p)}(t_0), \varphi^{(q)}(t_0))$ de la courbe paramétrée C en M_0, on a $X = t^p$, $Y = t^q$ et

- si p est impair alors $t = X^{1/p}$ et $Y = X^{q/p}$; la courbe C est — *dans le repère local en M_0 et au voisinage de M_0* — le graphe de la fonction $x \mapsto x^{q/p}$.
- si p est pair alors la relation $X = t^p$ est équivalente à $t = X^{1/p}$ ou à $t = -X^{1/p}$; la courbe C est — *dans le repère local en M_0 et au voisinage de M_0* — la réunion des graphes des deux fonctions $x \mapsto x^{q/p}$ et $x \mapsto (-x)^{q/p}$.

Ainsi, selon la parité des entiers p et q, on peut connaître l'allure de la courbe paramétrée au voisinage de M_0 dans le repère local de la courbe en ce point. Le tableau suivant indique les différentes allures possibles ainsi que la terminologie employée.

[54] Au sens de la comparaison locale des fonctions, voir la définition 15.1 p. 682 du *Cours de première année*.

	p impair	p pair
q pair	 Point ordinaire	 Point de rebroussement de seconde espèce
q impair	 Point d'inflexion	 Point de rebroussement de première espèce

Exemples

1. Considérons la cardioïde étudiée p. 759 dont un paramétrage est $\varphi : t \in [-\pi, \pi] \mapsto \big((1+\cos(t))\cos(t),\ (1+\cos(t))\sin(t)\big)$. Le point $(0,0)$ est un point de rebroussement de première espèce (on a vu que $p = 2$ et $q = 3$). Le point $(0,1)$ est un point ordinaire de cette cardioïde (on a vu que $p = 1$ et $q = 2$).

2. Le point $(0,0)$ est un point de rebroussement de première espèce de la quartique piriforme dont un paramétrage est $\varphi : t \in [0,\pi] \mapsto \big(\cos^2(t),\ \cos^3(t)\sin(t)\big)$, voir p. 757. On a en effet $\varphi'(\frac{\pi}{2}) = (0,0)$, $\varphi''(\frac{\pi}{2}) = (2,0)$ et $\varphi'''(\frac{\pi}{2}) = (0,-6)$ donc $p = 2$ et $q = 3$ car $\det(\varphi''(\frac{\pi}{2}), \varphi'''(\frac{\pi}{2})) = -6 \neq 0$.

Exercice 11.5 Soient $\varphi \in \mathcal{C}^2(I, \mathbb{R}^2)$ un arc paramétré du plan et $M_0 = \varphi(t_0)$ où $t_0 \in I$ un point régulier de la courbe paramétrée C associé à φ. On suppose que M_0 est un point d'inflexion de la courbe.

1 - a) Justifier que le système $\{\varphi'(t_0), \varphi''(t_0)\}$ est lié.

b) En déduire que t_0 vérifie $\varphi_1'(t_0)\,\varphi_2''(t_0) - \varphi_1''(t_0)\,\varphi_2'(t_0) = 0$ où φ_1 et φ_2 désignent les deux fonctions coordonnées de φ.

2 - En utilisant la propriété établie à la question précédente, déterminer parmi les points réguliers quels sont les points d'inflexion de la quartique piriforme dont un paramétrage est $\varphi : t \in [0,\pi] \mapsto (\cos^2(t),\ \cos^3(t)\sin(t))$.

PROPOSITION 11.7 (Concavité d'une courbe paramétrée)
Soient $\varphi \in \mathcal{C}^2(I, \mathbb{R}^2)$ un arc paramétré dans le plan et C la courbe paramétrée qui lui est associée. Si $M_0 = \varphi(t_0)$ avec $t_0 \in I$ est un point birégulier *de la courbe paramétrée C alors au voisinage de M_0 la courbe C se trouve dans le demi-plan limité par la tangente en M_0 et situé*[55] *du coté de $\varphi''(t_0)$.*

Démonstration Comme le point $M_0 = \varphi(t_0)$ est birégulier, on a $p = 1, q = 2$ et le repère local de la courbe paramétrée C en M_0 est $(M_0; \varphi'(t_0), \varphi''(t_0))$, voir la figure 8. D'après la relation (8) p. 762, les coordonnées $(X(t), Y(t))$ d'un point $M(t) = \varphi(t)$ dans ce repère local vérifient :

$$X(t) = t - t_0 + o_{t_0}\big((t-t_0)^2\big) \quad \text{et} \quad Y(t) = \frac{(t-t_0)^2}{2!} + o_{t_0}\big((t-t_0)^2\big).$$

Lorsque t est assez proche de t_0, $Y(t)$, qui est la coordonnée du point $M(t)$ relative au vecteur $\varphi''(t_0)$, est positive. Le point $M(t)$ se trouve donc dans le demi-plan limité par la droite passant par M_0 et de vecteur directeur $\varphi'(t_0)$ et situé du coté de $\varphi''(t_0)$. □

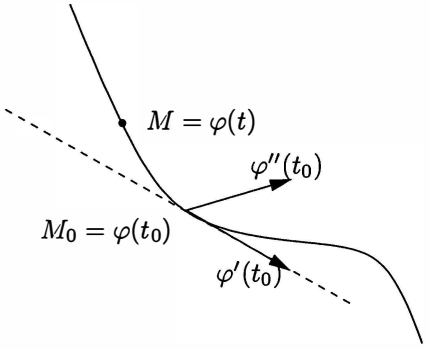


Fig. 8 En un point birégulier, une courbe paramétrée tourne sa concavité du coté de $\varphi''(t_0)$.

Branches infinies d'une courbe paramétrée plane

DÉFINITION 11.14 (Branche infinie)
Soient $\varphi \in \mathcal{C}^0(I, \mathbb{R}^2)$ un arc paramétré du plan et C la courbe paramétrée qui lui est associée. On dit que la courbe paramétrée C admet une branche infinie lorsque t tend vers t_0 où $t_0 \in \overline{I}$ si

$$\lim_{t \to t_0} \|\varphi(t)\| = +\infty.$$

[55] Voir la figure 8.

Détaillons les différents types de branches infinies possibles en $t_0 \in \overline{I}$ (t_0 est soit l'une des extrémités de l'intervalle I, soit $\pm\infty$). Notons $(x(t), y(t))$ les coordonnées du point $M = \varphi(t)$ de la courbe paramétrée C.

1. Si $\lim_{t \to t_0} y(t) = \pm\infty$ et $\lim_{t \to t_0} x(t) = \alpha$ alors la courbe paramétrée C possède pour branche infinie une droite asymptote verticale d'équation $x = \alpha$.
2. Si $\lim_{t \to t_0} y(t) = \beta$ et $\lim_{t \to t_0} x(t) = \pm\infty$ alors la courbe paramétrée C possède pour branche infinie une droite asymptote horizontale d'équation $y = \beta$.
3. Si $\lim_{t \to t_0} y(t) = \pm\infty$ et $\lim_{t \to t_0} x(t) = \pm\infty$ alors on s'intéresse, si elle existe, à la limite $a = \lim_{t \to t_0} \dfrac{y(t)}{x(t)}$.
   - Si $a = \pm\infty$ alors la courbe paramétrée C possède pour branche infinie une *branche parabolique* de direction l'axe des ordonnées, voir la figure 9.
   - Si $a = 0$ alors la courbe paramétrée C possède pour branche infinie une *branche parabolique* de direction l'axe des abscisses, voir la figure 9.

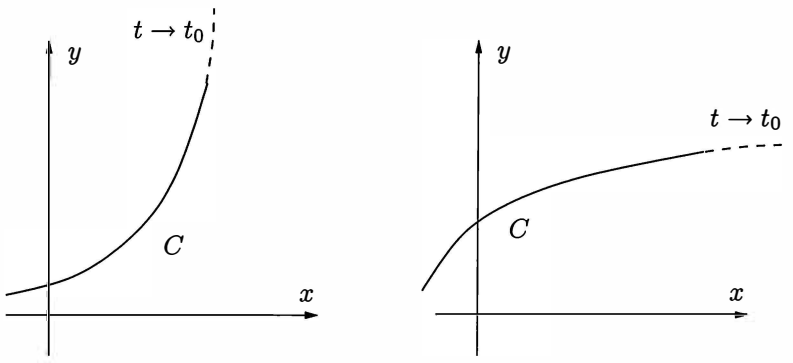


Fig. 9 Branche parabolique de direction l'axe des ordonnées (à gauche) et branche parabolique de direction l'axe des abscisses (à droite).

- Si $a \in \mathbb{R}^*$ alors on s'intéresse, si elle existe, à la limite $\lim_{t \to t_0} y(t) - ax(t)$.
  - Si cette limite vaut $\pm\infty$ alors la courbe paramétrée C possède pour branche infinie une *branche parabolique de direction asymptotique* la droite d'équation $y = ax$, voir la figure 10.
  - Si cette limite existe et vaut $b \in \mathbb{R}$ alors la courbe paramétrée C possède pour branche infinie une *droite asymptote* d'équation $y = ax + b$. De plus,
    - si $\lim_{t \to t_0} y(t) - (ax(t) + b) > 0$, la courbe est au-dessus de son asymptote ;
    - si $\lim_{t \to t_0} y(t) - (ax(t) + b) < 0$, la courbe est au-dessous de son asymptote.

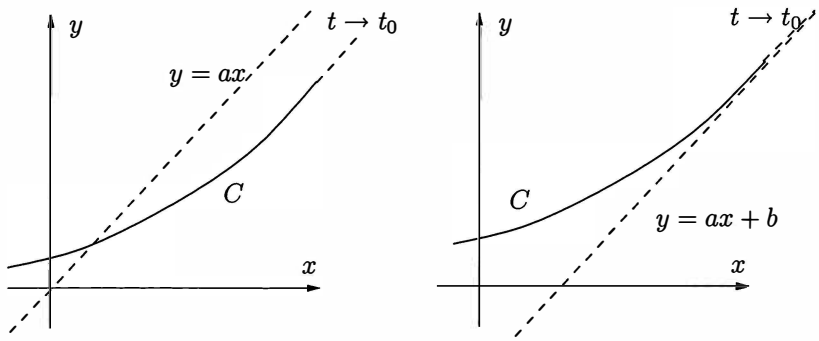


Fig. 10 Branche parabolique de direction asymptotique la droite d'équation $y = ax$ (à gauche) et droite asymptote d'équation $y = ax + b$ (à droite).

Ne pas confondre du point de vue terminologique, « branche infinie » et « branche parabolique » ; les asymptotes et les branches paraboliques sont des branches infinies.

Exemples

1. Considérons l'arc paramétré

$$\varphi : t \in \mathbb{R}_+^* \longmapsto \left(\mathrm{ch}(t) + \frac{1}{\mathrm{ch}(t)},\ \mathrm{sh}(t) + \frac{1}{\mathrm{sh}(t)} \right).$$

Comme $\|\varphi(t)\| = \sqrt{(\mathrm{ch}(t) + (\mathrm{ch}(t))^{-1})^2 + (\mathrm{sh}(t) + (\mathrm{sh}(t))^{-1})^2}$ admet pour limite $+\infty$ lorsque t tend vers 0 ou lorsque t tend vers $+\infty$, la courbe paramétrée C associée à φ admet deux branches infinies. Posons $x(t) = \mathrm{ch}(t) + (\mathrm{ch}(t))^{-1}$ et $y(t) = \mathrm{sh}(t) + (\mathrm{sh}(t))^{-1}$. On a $\mathrm{ch}(t) \underset{0}{\sim} 1$ et $\mathrm{sh}(t) \underset{0}{\sim} t$ donc

$$\frac{y(t)}{x(t)} \underset{0}{\sim} \frac{1}{2t} \quad \text{d'où} \quad \lim_{t \to 0+} \frac{y(t)}{x(t)} = +\infty.$$

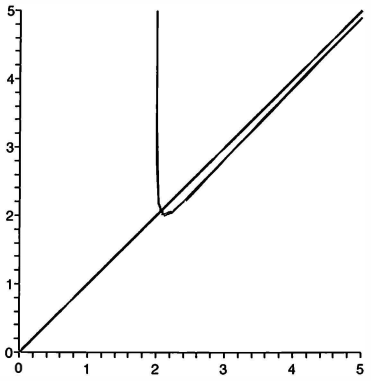
La courbe paramétrée C admet lorsque t tend vers 0 pour branche infinie une branche parabolique de direction l'axe des ordonnées. Puisque $\mathrm{ch}(t) \underset{+\infty}{\sim} \frac{1}{2}\mathrm{e}^t$ et $\mathrm{sh}(t) \underset{+\infty}{\sim} \frac{1}{2}\mathrm{e}^t$, on a $\lim_{t \to +\infty} y(t)/x(t) = 1$ et

$$\lim_{t \to +\infty} y(t) - x(t) = \lim_{t \to +\infty} \mathrm{sh}(t) - \mathrm{ch}(t) = \lim_{t \to +\infty} -\mathrm{e}^{-t} = 0.$$

La courbe paramétrée C admet, lorsque t tend vers $+\infty$, pour branche infinie une droite asymptote d'équation $y = x$.

On peut avoir recours à MAPLE pour illustrer la situation.

```
> phi:=t-> [cosh(t)+1/cosh(t),sinh(t)+1/sinh(t)]:
> des1:=plot([op(phi(t)),t=0..10], thickness=2,view=[0..5,0..5]):
> des2:=plot(x,x=0..5,thickness=2,linestyle=dot):
> display(des1,des2);
```



2. Considérons l'arc paramétré

$$\varphi : t \in \mathbb{R}_+^* \longmapsto \left(2t^2 + t,\ 4t^2 + \frac{5}{t^2}\right).$$

Comme $\|\varphi(t)\| = \sqrt{(2t^2+t)^2 + (4t^2+5/t^2)^2}$ admet pour limite $+\infty$ lorsque t tend vers 0 ou lorsque t tend vers $+\infty$, la courbe paramétrée C associée à φ admet deux branches infinies. Posons $x(t) = 2t^2+t$ et $y(t) = 4t^2+5/t^2$. Puisque

$$\frac{y(t)}{x(t)} = \frac{4t^4+5}{2t^4+t^3} \underset{0}{\sim} \frac{5}{t^3} \quad \text{on a} \quad \lim_{t\to 0^+} \frac{y(t)}{x(t)} = +\infty,$$

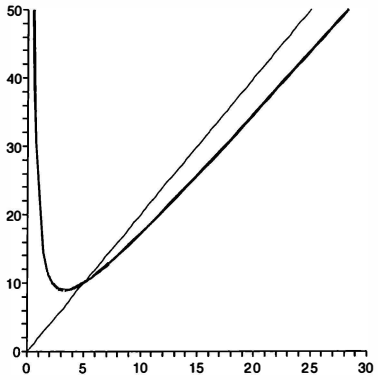
et la courbe paramétrée C admet lorsque t tend vers 0 pour branche infinie une branche parabolique de direction l'axe des ordonnées. Puisque

$$\lim_{t\to+\infty} \frac{y(t)}{x(t)} = 2 \quad \text{et} \quad \lim_{t\to+\infty} y(t) - 2x(t) = \lim_{t\to+\infty} \frac{5}{t^2} - 2t = -\infty,$$

la courbe paramétrée C admet lorsque t tend vers $+\infty$ pour branche infinie une branche parabolique de direction la droite d'équation $y = 2x$ mais n'admet pas de droite asymptote.

On peut avoir recours à MAPLE pour visualiser la situation.

```
> phi:=t-> [2*t^2+t,4*t^2+5/t^2]:
> des1:=plot([op(phi(t)),t=0..10],view=[0..30,0..50]):
> des2:=plot(2*x,x=0..30):
> display(des1,des2);
```



Remarque Si les fonctions coordonnées x et y tendent vers des limites finies x_0 et y_0 lorsque le paramètre t tend vers les extrémités de l'intervalle I alors le point (x_0, y_0) est appelé une *extrémité* ou un *point d'arrêt* de la courbe. ◇

Plan d'étude d'une courbe paramétrée dans le plan

L'étude de la courbe paramétrée dans le plan associée à l'arc paramétré $\varphi : t \in I \to (x(t), y(t)) \in \mathbb{R}^2$ inclut les étapes suivantes dont la mise en œuvre a été détaillée dans les paragraphes précédents.

- Réduire l'intervalle d'étude en étudiant les propriétés de périodicité et de parité de φ ou de symétrie de la courbe paramétrée C.
- Étudier les deux fonctions coordonnées de l'application φ et donner leur tableau de variation.
- Rechercher et étudier les points caractéristiques de la courbe : les points singuliers, les points multiples, etc. et évaluer les entiers p et q pour connaître l'allure de la courbe paramétrée C au voisinage de ces points.
- Étudier les branches infinies de la courbe paramétrée C.
- Tracer la courbe paramétrée C en faisant apparaître le renseignements obtenus précédemment.

Exemple Étudions la courbe paramétrée dans le plan associée à l'arc paramétré

$$\varphi : t \in \mathbb{R} \longmapsto \left(t^3(t-2),\ t^2(t-2)(t-1)\right).$$

On note $x : t \in \mathbb{R} \mapsto t^3(t-2)$ et $y : t \in \mathbb{R} \mapsto t^2(t-2)(t-1)$ les deux fonctions coordonnées de cette application.

$\trianglerighteq$ *Ensemble de définition, réduction de l'intervalle d'étude*
L'ensemble de définition de φ est $\mathbb{R}$ car les deux fonctions coordonnées x et y sont polynomiales. L'arc paramétré φ est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$. Aucune symétrie ou périodicité n'est décelable : on ne peut donc pas réduire l'intervalle d'étude.

▷ *Tableau de variation*
Les dérivées des deux fonctions coordonnées x et y en $t \in \mathbb{R}$ sont

$$x'(t) = 2t^2(2t-3) \quad \text{et} \quad y'(t) = t(4t^2 - 9t + 4).$$

On a $x'(t) = 0$ si $t = 0$ ou $t = 3/2$ et $y'(t) = 0$ si $t = 0$ ou si t est racine du polynôme $4X^2 - 9X + 4$. Le polynôme $4X^2 - 9X + 4$ admet deux racines réelles qui sont

$$r_1 = \frac{9+\sqrt{17}}{8} \quad \text{et} \quad r_2 = \frac{9-\sqrt{17}}{8}.$$

Le tableau de variation de φ est le suivant :

t	$-\infty$		0		r_2		$\frac{3}{2}$		r_1		$+\infty$
$x'(t)$		$-$	0		$-$		0		$+$		
$x(t)$	$+\infty$	$\searrow$	0	$\searrow$	$x(r_2)$	$\searrow$	$-\frac{27}{16}$	$\nearrow$	$x(r_1)$	$\nearrow$	$+\infty$
$y'(t)$		$-$	0	$+$	0	$-$			0	$+$	
$y(t)$	$+\infty$	$\searrow$	0	$\nearrow$	$y(r_2)$	$\searrow$	$-\frac{9}{16}$	$\searrow$	$y(r_1)$	$\nearrow$	$+\infty$

▷ *Points singuliers*

D'après la définition 11.11, les points singuliers de la courbe sont les points $M_0 = \varphi(t_0)$ tels que $\varphi'(t_0) = 0$. Le tableau de variation indique que la seule valeur du paramètre t_0 pour lesquels $x'(t_0) = y'(t_0) = 0$ est $t_0 = 0$. Il existe donc un unique point singulier qui a pour coordonnées $\varphi(0) = (0,0)$. Pour déterminer l'allure de la courbe C au voisinage de ce point, il faut calculer les entiers p et q définissant le repère local en ce point. Puisque

$$\varphi''(0) = \begin{pmatrix} 0 \\ 4 \end{pmatrix}, \quad \varphi'''(0) = \begin{pmatrix} -12 \\ -18 \end{pmatrix} \quad \text{et que} \quad \det(\varphi''(0), \varphi'''(0)) = 48 \neq 0,$$

on a $p = 2$ et $q = 3$. L'origine est donc un point de rebroussement de seconde espèce.

▷ *Points multiples*
D'après la définition 11.10, un point $M \in \mathbb{R}^2$ est un point multiple s'il existe deux valeurs différentes t_1 et t_2 du paramètre telles que $M = \varphi(t_1) = \varphi(t_2)$. Pour obtenir tous les points multiples de la courbe, déterminons les valeurs différentes t_1 et t_2 du paramètre telles que $\varphi(t_1) = \varphi(t_2)$, *i.e.* résolvons le système (formé de deux équations non linéaires) suivant :

$$\begin{cases} t_1^3(t_1-2) &= t_2^3(t_2-2) \\ t_1^2(t_1-2)(t_1-1) &= t_2^2(t_2-2)(t_2-1) \end{cases}.$$

En faisant la différence de la première équation multipliée au préalable par $(t_1 - 1)$ avec la seconde équation multipliée au préalable par t_1, on obtient que le couple (t_1, t_2) solution du système vérifie l'équation

$$(t_1 - 1)t_2^3(t_2 - 2) = t_1(t_2 - 2)(t_2 - 1)t_2^2.$$

Si $t_2 \neq 0$ et si $t_2 \neq 2$, cette équation se simplifie et est équivalente à $(t_1 - 1)t_2 = t_1(t_2 - 1)$, c'est-à-dire à $t_1 = t_2$. On en déduit que les valeurs du paramètre appartenant à $\mathbb{R} \setminus \{0, 2\}$ correspondent toutes à des points simples. Comme $\varphi(0) = \varphi(2) = (0, 0)$, on en déduit aussi que l'origine est l'unique point multiple ; il s'agit d'un point double.

$\trianglerighteq$ *Étude de l'allure de la courbe au voisinage de l'origine*
Pour la valeur $t_1 = 0$ du paramètre, nous avons déjà établi que l'origine est un point singulier (point de rebroussement de seconde espèce) et que la tangente en ce point est dirigée par $\varphi''(0) = (0, 4)$. Pour la valeur $t_2 = 2$ du paramètre, l'origine est un point régulier dont la tangente est dirigée par $\varphi'(2) = (8, 4)$. Il est important de noter que l'on a $\varphi''(2) = (24, 16)$ ce qui signifie que l'on a un point birégulier au voisinage duquel la concavité de la courbe est dirigée dans le sens du vecteur $(24, 16)$.

$\trianglerighteq$ *Étude des points d'arrêts ou branches infinies*
Intéressons-nous au comportement de la courbe lorsque le paramètre tend vers les bords de son ensemble de définition. Ici, l'ensemble de l'étude étant $\mathbb{R}$, on cherche à connaître le comportement de la courbe lorsque t tend vers $+\infty$ et $-\infty$. Comme

$$\lim_{t\to-\infty} x(t) = \lim_{t\to-\infty} y(t) = +\infty \quad \text{et} \lim_{t\to+\infty} x(t) = \lim_{t\to+\infty} y(t) = +\infty,$$

la courbe possède une branche infinie lorsque t tend vers $+\infty$ ou $-\infty$. De plus, $y(t)/x(t) = 1 - 1/t$ donc

$$\lim_{t\to-\infty} \frac{y(t)}{x(t)} = \lim_{t\to+\infty} \frac{y(t)}{x(t)} = 1$$

et $y(t) - x(t) = -t^2(t - 2)$ donc

$$\lim_{t\to-\infty} y(t) - x(t) = \lim_{t\to+\infty} y(t) - x(t) = -\infty.$$

On en déduit que lorsque t tend vers $+\infty$ ou $-\infty$, la courbe possède une branche parabolique de direction asymptotique la droite $y = x$.

▷ *Tracé*

Le tableau de variation permet de décrire le tracé de la courbe paramétrée C. La valeur $t = 0$ du paramètre correspond à l'origine. On a un point de rebroussement de seconde espèce. Lorsque t tend vers $-\infty$ la courbe s'éloigne de l'origine selon une branche parabolique de direction asymptotique la droite d'équation $y = x$, en passant par le point de coordonnées $(3, 6)$ correspondant à la valeur $t = -1$ du paramètre. La demi-tangente au point de rebroussement situé à l'origine est dirigée par le vecteur $\varphi''(0) = (0, 4)$. Lorsque t varie entre 0 et $r_2 \approx 0,61$, la courbe rejoint le point M_1 de coordonnées $\varphi(r_2) \approx (-0,31,\ 0,20)$ et en ce point la courbe possède une tangente horizontale. La courbe joint ensuite le point M_2 de coordonnées $\varphi(3/2) = (-27/16,\ -9/16) \approx (-1,69,\ -0,56)$ où la tangente est verticale, en coupant l'axe des abscisses en $\varphi(1) = (-1, 0)$. Lorsque t varie entre $3/2$ et r_1, la courbe va du point M_2 au point M_3 de coordonnées $\varphi(r_1) \approx (-1,59,\ -0,52)$. La tangente à la courbe en M_3 est horizontale. Lorsque t varie de r_1 à $+\infty$, la courbe repasse par l'origine pour la valeur $t = 2$ du paramètre et admet à l'origine une tangente de pente $1/2$. Elle s'éloigne ensuite de l'origine selon une branche parabolique de direction asymptotique la droite d'équation $y = x$ en passant par le point de coordonnées $(27, 18)$ pour la valeur $t = 3$ du paramètre.

La courbe paramétrée C est représentée à la figure 11.

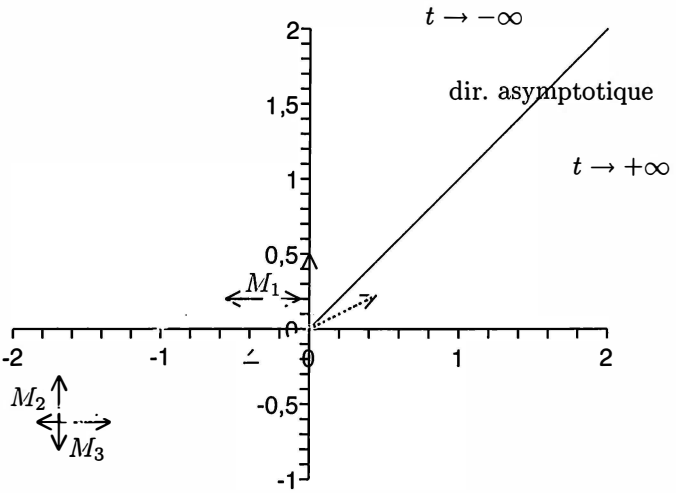


Fig. 11 Tracé de la courbe paramétrée autour de l'origine.

Exercice 11.6 Étudier la courbe de Lissajous C dont un paramétrage est[56] $\varphi : t \in \mathbb{R} \longmapsto (2\sin(t), \cos(3t/2))$.

[56] Cette courbe est une courbe de Lissajous définie p. 715 correspondant aux valeurs $a = 2$, $b = 1$, $\eta = 3/2$ et $\phi = \pi/2$.

11.3.4 Étude des arcs paramétrés dans l'espace

Une étude analogue à celle qui vient d'être menée pour les arcs paramétrés dans le plan peut être effectuée pour les arcs paramétrés dans l'espace. En étudiant les trois fonctions coordonnées d'un arc paramétré de $\mathbb{R}^3$, on est en mesure d'obtenir de nombreuses informations sur la courbe paramétrée. Une différence importante avec ce qui a été présenté dans la partie précédente pour les courbes paramétrées dans le plan est que la famille de deux vecteurs $\{\varphi^{(p)}(t_0), \varphi^{(q)}(t_0)\}$ n'est pas une base de $\mathbb{R}^3$; cette famille définit un sous-espace vectoriel de $\mathbb{R}^3$ de dimension 2 qui nous amène à définir la notion de *plan osculateur*. Par ailleurs, pour une courbe paramétrée dans l'espace, on est en général confronté à la difficulté de dessiner l'allure de la courbe à partir de l'information donnée par les tableaux de variation des fonctions coordonnées. En pratique, les logiciels de visualisation graphique se révèlent souvent indispensables pour visualiser une courbe dans l'espace. L'étude analytique de l'arc paramétré reste nécessaire pour établir les caractéristiques de la courbe.

Plan osculateur et repère local

DÉFINITION 11.15 (Plan osculateur)
Soient $\varphi \in \mathcal{C}^\infty(I, \mathbb{R}^3)$ un arc paramétré dans l'espace et C la courbe paramétrée qui lui est associée. Soit $t_0 \in I$ une valeur du paramètre tel que les entiers

$$p = \min\{j \in \mathbb{N}^* \ ; \ \varphi^{(j)}(t_0) \neq \mathbf{0}\} \quad \textit{et}$$
$$q = \min\left\{j \in \mathbb{N}^* \ ; \ j > p \quad \textit{et} \quad \{\varphi^{(p)}(t_0), \varphi^{(j)}(t_0)\} \ \textit{est libre}\right\}$$

soient définis. On appelle plan osculateur *à la courbe paramétrée C en $M_0 = \varphi(t_0)$, le plan passant par M_0 et dirigé par les vecteurs $\varphi^{(p)}(t_0)$ et $\varphi^{(q)}(t_0)$.*

Remarques

1. La notion de plan osculateur est une notion intrinsèque à la courbe paramétrée C ; il est le même[57] pour tous les arcs paramétrés $\mathcal{C}^1$-équivalents à l'arc paramétré φ choisi comme représentation paramétrique de la courbe.

2. Le plan osculateur en $M_0 = \varphi(t_0)$ à la courbe paramétrée C est, d'après la définition 11.15, l'ensemble

$$\mathcal{P} = \left\{\mathbf{x} \in \mathbb{R}^3 \ ; \ \exists(\lambda, \mu) \in \mathbb{R}^2 \quad \mathbf{x} = \varphi(t_0) + \lambda\varphi^{(p)}(t_0) + \mu\varphi^{(q)}(t_0)\right\}.$$

Il admet pour représentation paramétrique

$$\varphi : (u, v) \in \mathbb{R}^2 \longmapsto \varphi(t_0) + u\varphi^{(p)}(t_0) + v\varphi^{(q)}(t_0) \in \mathbb{R}^3$$

[57] La preuve de ce résultat reprend les arguments utilisés pour montrer que la notion de repère local à un arc paramétré du plan $\mathbb{R}^2$ est une notion intrinsèque, voir la remarque de la p. 759.

En désignant par (a, b, c) les composantes du produit vectoriel des vecteurs $\varphi^{(p)}(t_0)$ et $\varphi^{(q)}(t_0)$ et par (x_0, y_0, z_0) les coordonnées du point M_0, l'équation cartésienne du plan osculateur en M_0 est

$$a\,(x - x_0) + b\,(y - y_0) + c\,(z - z_0) = 0.$$

3. Lorsque $M_0 = \varphi(t_0)$ est un point birégulier de la courbe paramétrée C, la famille $\{\varphi'(t_0), \varphi''(t_0)\}$ est libre et $p = 1$, $q = 2$. Le plan osculateur à la courbe paramétrée C en M_0 est alors le plan passant par M_0 et dirigé par les vecteurs $\varphi'(t_0)$ et $\varphi''(t_0)$. ◇

Exemples

1. Déterminons le plan osculateur au point $(2, 0, 0)$ à la courbe paramétrée C de l'espace dont un paramétrage est

$$\varphi : t \in [0, 2\pi] \longmapsto (\cos(t) + \cos(3t),\ \sin(t) - \sin(3t),\ 2\sin(2t)).$$

On a $(2,0,0) = \varphi(0)$ et

$$\varphi'(0) = \begin{pmatrix} 0 \\ -2 \\ 4 \end{pmatrix} \quad \text{et} \quad \varphi''(0) = \begin{pmatrix} -10 \\ 0 \\ 0 \end{pmatrix}.$$

Les deux vecteurs $\varphi'(0)$ et $\varphi''(0)$ forment très clairement une famille libre car $\varphi''(0)$ est colinéaire au troisième vecteur de la base canonique de $\mathbb{R}^3$ alors que $\varphi'(0)$ est une combinaison linéaire des deux premiers vecteurs de la base. Le plan osculateur au point $(2, 0, 0)$ admet pour représentation paramétrique

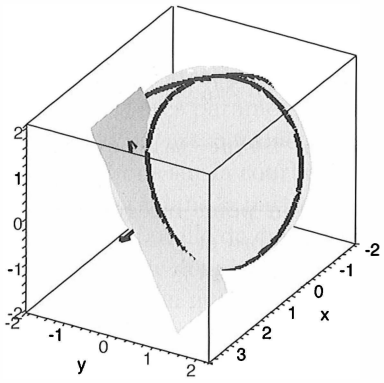
$$(u, v) \in \mathbb{R}^2 \longmapsto (-10v,\quad -2u,\quad 4u)$$

et pour équation cartésienne $z = -2y$.
On pourra par ailleurs vérifier que $\varphi_1^2(t)+\varphi_2^2(t)+\varphi_3^2(t) = 4$ pour tout $t \in [0, 2\pi]$, ce qui indique que la courbe paramétrée C appartient à la sphère centrée de centre $(0, 0, 0)$ et de rayon 2.

On peut avoir recours à MAPLE pour représenter[58] la courbe paramétrée C et son plan osculateur au point de coordonnées $(2, 0, 0)$.

```
> fy:=t->[cos(t)+cos(3*t),sin(t)-sin(3*t),2*sin(2*t)]:
> des1:=spacecurve(fy(t),t=0..2*Pi,scaling=constrained):
> des2:=plot3d(-2*y,x=1..3,y=-1..1):
> des3:=plot3d([2*cos(u)*sin(v),2*sin(u)*sin(v),2*cos(v)],
                u=0..2*Pi,v=0..Pi):
> des4:=arrow({fy(t0)},{Dn(fy,1,t0),-Dn(fy,2,t0)}):
# des1 : la courbe; des2 : le plan osculateur; des3 : la sphere;
# des4 : les vecteurs directeurs du plan osculateur
> display(des1,des2,des3,des4);
```

[58] La procédure MAPLE Dn est décrite p. 760.



2. Déterminons le plan osculateur en chacun de ses points à la courbe paramétrée C de l'espace dont un paramétrage est

$$\varphi : t \in [0, 2\pi] \longmapsto (\sqrt{3}\cos(t),\ \cos(t),\ \sqrt{3}\sin(t))$$

Pour tout $t \in [0, 2\pi]$, on a

$$\varphi'(t) = \begin{pmatrix} -\sqrt{3}\sin(t) \\ -\sin(t) \\ \sqrt{3}\cos(t) \end{pmatrix} \quad \text{et} \quad \varphi''(t) = \begin{pmatrix} -\sqrt{3}\cos(t) \\ -\cos(t) \\ -\sqrt{3}\sin(t) \end{pmatrix}.$$

La courbe est régulière et $p = 1$ car les fonctions sinus et cosinus ne s'annulant jamais simultanément, on a $\varphi'(t) \neq 0$ pour tout $t \in [0, 2\pi]$. Montrons que la courbe paramétrée C est une courbe birégulière, c'est-à-dire que pour tout $t \in [0, 2\pi]$ les vecteurs $\varphi'(t)$ et $\varphi''(t)$ sont libres. Pour cela, on peut montrer que le rang de la famille $\{\varphi'(t), \varphi''(t)\}$ vaut 2 en montrant que l'on peut extraire de la matrice $\mathrm{A} \in \mathcal{M}_{3,2}(\mathbb{R})$ formée des vecteurs colonnes $\varphi'(t)$ et $\varphi''(t)$ un déterminant d'ordre 2 non nul[59] ; on a

$$\mathrm{A} = \begin{pmatrix} -\sqrt{3}\sin(t) & -\sqrt{3}\cos(t) \\ -\sin(t) & -\cos(t) \\ \sqrt{3}\cos(t) & -\sqrt{3}\sin(t) \end{pmatrix} \quad \text{et} \quad \begin{vmatrix} -\sin(t) & -\cos(t) \\ \sqrt{3}\cos(t) & -\sqrt{3}\sin(t) \end{vmatrix} = \sqrt{3} \neq 0.$$

Une autre façon d'établir que pour tout $t \in [0, 2\pi]$ les vecteurs $\varphi'(t)$ et $\varphi''(t)$ sont libres, consiste à calculer le produit vectoriel $\varphi'(t) \wedge \varphi''(t)$ et à montrer que ce vecteur n'est jamais nul ; on a

$$\varphi'(t) \wedge \varphi''(t) = \begin{pmatrix} -\sqrt{3}\sin(t) \\ -\sin(t) \\ \sqrt{3}\cos(t) \end{pmatrix} \wedge \begin{pmatrix} -\sqrt{3}\cos(t) \\ -\cos(t) \\ -\sqrt{3}\sin(t) \end{pmatrix} = \begin{pmatrix} \sqrt{3} \\ -3 \\ 0 \end{pmatrix}.$$

[59] Voir la définition 8.15 p. 332 et la proposition 11.6 p. 485 du *Cours de première année*.

On en déduit que $q = 2$ et que le plan osculateur à la courbe paramétré C en un point $M = \varphi(t)$ de la courbe paramétrée C est le plan passant par M et dirigé par les vecteurs $\varphi'(t)$ et $\varphi''(t)$. Il admet pour représentation paramétrique

$$(u, v) \in \mathbb{R}^2 \longmapsto \begin{pmatrix} -\sqrt{3}\,\sin(t)\,u - \sqrt{3}\,\cos(t)\,v \\ -\sin(t)\,u - \cos(t)\,v \\ \sqrt{3}\,\cos(t)\,u - \sqrt{3}\,\sin(t)\,v \end{pmatrix}$$

et pour équation cartésienne

$$\sqrt{3}(x - \sqrt{3}\cos(t)) - 3(y - \cos(t)) = 0.$$

On peut remarquer que cette équation cartésienne se réécrit $\sqrt{3}y - x = 0$, ce qui indique que la courbe paramétrée C admet le même plan osculateur en chacun de ses points.

Comme pour les courbes paramétrées dans le plan, on peut définir la notion de *repère local* d'une courbe paramétrée C dans l'espace en un point M_0. L'espace étant de dimension 3, il faut adjoindre aux deux vecteurs $\varphi^{(p)}(t_0)$ et $\varphi^{(q)}(t_0)$ de la définition 11.15, un troisième vecteur qui leur est linéairement indépendant pour obtenir une base de $\mathbb{R}^3$.

DÉFINITION 11.16 (Repère local d'une courbe dans l'espace)
Soient $\varphi \in \mathcal{C}^\infty(I, \mathbb{R}^3)$ un arc paramétré dans l'espace et C la courbe paramétrée qui lui est associée. Soit $t_0 \in I$ une valeur du paramètre tel que les entiers

$$p = \min\{j \in \mathbb{N}^* \ ; \ \varphi^{(j)}(t_0) \neq \mathbf{0}\},$$
$$q = \min\left\{j \in \mathbb{N}^* \ ; \ j > p \quad et \quad \{\varphi^{(p)}(t_0), \varphi^{(j)}(t_0)\} \ est\ libre\right\} \quad et$$
$$r = \min\left\{j \in \mathbb{N}^* \ ; \ j > q \quad et \quad \{\varphi^{(p)}(t_0), \varphi^{(q)}(t_0), \varphi^{(j)}(t_0)\} \ est\ libre\right\}$$

soient définis. On appelle repère local *au point $M_0 = \varphi(t_0)$ de la courbe paramétrée C, le repère d'origine M_0 dont les vecteurs de la base sont $\varphi^{(p)}(t_0)$, $\varphi^{(q)}(t_0)$ et $\varphi^{(r)}(t_0)$.*

Remarques

1. Le repère local en un point d'une courbe paramétrée dans l'espace est formé des deux vecteurs définissant le plan osculateur, voir la définition 11.15, et d'un troisième vecteur qui n'appartient pas au plan osculateur. Ce troisième vecteur n'est pas nécessairement normal au plan osculateur.

2. La famille $\{\varphi^{(p)}(t_0), \varphi^{(q)}(t_0), \varphi^{(j)}(t_0)\}$ est libre si et seulement si le déterminant $\det(\varphi^{(p)}(t_0), \varphi^{(q)}(t_0), \varphi^{(j)}(t_0))$ est non nul. Dans ce cas les deux vecteurs $\varphi^{(p)}(t_0)$ et $\varphi^{(q)}(t_0)$ sont également libres. ◇

Exemples

1. Déterminons le repère local au point $(2,0,0)$ de la courbe paramétrée C de l'espace dont un paramétrage est

$$\varphi : t \in [0, 2\pi] \longmapsto (\cos(t) + \cos(3t),\ \sin(t) - \sin(3t),\ 2\sin(2t)).$$

On a $(2,0,0) = \varphi(0)$ et

$$\varphi'(0) = \begin{pmatrix} 0 \\ -2 \\ 4 \end{pmatrix}, \quad \varphi''(0) = \begin{pmatrix} -10 \\ 0 \\ 0 \end{pmatrix} \quad \text{et} \quad \varphi'''(0) = \begin{pmatrix} 0 \\ 26 \\ -16 \end{pmatrix}.$$

Le déterminant de la matrice dont les vecteurs colonnes sont $\varphi'(0)$, $\varphi''(0)$ et $\varphi'''(0)$ vaut -720. Puisqu'il est non nul, les trois vecteurs $\varphi'(0)$, $\varphi''(0)$ et $\varphi'''(0)$ sont libres. On a donc $p = 1$, $q = 2$, $r = 3$.

2. Déterminons à présent le repère local au point $(0,0,2)$ de cette même courbe paramétrée C. On a $(0,0,2) = \varphi(\frac{1}{4}\pi)$ et

$$\varphi'(\tfrac{1}{4}\pi) = \begin{pmatrix} -2\sqrt{2} \\ 2\sqrt{2} \\ 0 \end{pmatrix}, \quad \varphi''(\tfrac{1}{4}\pi) = \begin{pmatrix} 4\sqrt{2} \\ 4\sqrt{2} \\ -8 \end{pmatrix} \quad \text{et} \quad \varphi'''(\tfrac{1}{4}\pi) = \begin{pmatrix} 14\sqrt{2} \\ -14\sqrt{2} \\ 0 \end{pmatrix}.$$

Le déterminant de la matrice dont les vecteurs colonnes sont $\varphi'(\frac{1}{4}\pi)$, $\varphi''(\frac{1}{4}\pi)$ et $\varphi'''(\frac{1}{4}\pi)$ vaut

$$\begin{vmatrix} -2\sqrt{2} & 4\sqrt{2} & 14\sqrt{2} \\ 2\sqrt{2} & 4\sqrt{2} & -14\sqrt{2} \\ 0 & -8 & 0 \end{vmatrix} = 8 \begin{vmatrix} -2\sqrt{2} & 14\sqrt{2} \\ 2\sqrt{2} & -14\sqrt{2} \end{vmatrix} = 0.$$

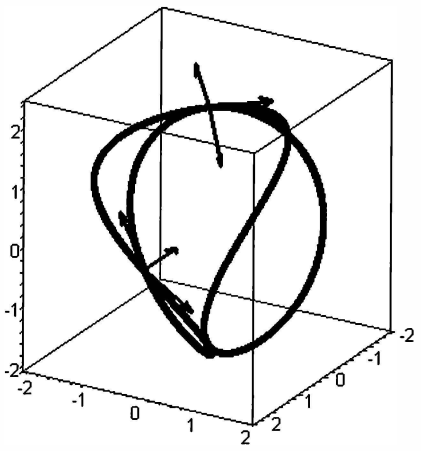
Puisque ce déterminant est nul, les trois vecteurs $\varphi'(\frac{1}{4}\pi)$, $\varphi''(\frac{1}{4}\pi)$ et $\varphi'''(\frac{1}{4}\pi)$ sont liés. On peut facilement voir que le déterminant d'ordre 2 extrait de ce déterminant en conservant les deux premières lignes et deux premières colonnes vaut -32, ce qui indique que le rang de la famille $\{\varphi'(\frac{1}{4}\pi), \varphi''(\frac{1}{4}\pi)\}$ vaut 2 et par conséquent que les vecteurs $\varphi'(\frac{1}{4}\pi)$ et $\varphi''(\frac{1}{4}\pi)$ forment une famille libre. On a $\varphi^{(4)}(\frac{1}{4}\pi) = (-40\sqrt{2},\ -40\sqrt{2},\ 32)$ et le déterminant de la matrice dont les vecteurs colonnes sont $\varphi'(\frac{1}{4}\pi)$, $\varphi''(\frac{1}{4}\pi)$ et $\varphi^{(4)}(\frac{1}{4}\pi)$ vaut

$$\begin{vmatrix} -2\sqrt{2} & 4\sqrt{2} & -40\sqrt{2} \\ 2\sqrt{2} & 4\sqrt{2} & -40\sqrt{2} \\ 0 & -8 & 32 \end{vmatrix} = 256 \begin{vmatrix} -1 & 2 & -5 \\ 1 & 2 & -5 \\ 0 & -1 & 1 \end{vmatrix}.$$

En remplaçant la première ligne par la première ligne à laquelle on a retranché la seconde, puis en développant le déterminant selon cette nouvelle première ligne, on établit qu'il vaut 1536. Puisqu'il est non nul, les trois vecteurs $\varphi'(\frac{1}{4}\pi)$, $\varphi''(\frac{1}{4}\pi)$ et $\varphi^{(4)}(\frac{1}{4}\pi)$ sont libres On a donc $p = 1$, $q = 2$, $r = 4$.

On peut avoir recours à MAPLE pour représenter[60] la courbe paramétrée C ainsi que les repères locaux à cette courbe aux points de coordonnées $(2,0,0)$ et $(0,0,2)$.

```
> fy:=t->[cos(t)+cos(3*t),sin(t)-sin(3*t),2*sin(2*t)]:
> des1:=spacecurve(fy(t),t=0..2*Pi,scaling=constrained):
> t:=0: des2:=arrow({fy(t)},{Dn(fy,1,t),Dn(fy,2,t),Dn(fy,3,t)}):
> t:=Pi/4: des3:=arrow({fy(t)},{Dn(fy,1,t),Dn(fy,2,t),Dn(fy,4,t)}):
> display(des1,des2,des3);
```



La formule de Taylor-Young appliquée à l'arc paramétré φ au voisinage de la valeur t_0 du paramètre permet, comme pour les courbes paramétrées dans le plan, de préciser l'allure de la courbe paramétrée C au voisinage du point $M_0 = \varphi(t_0)$ et d'effectuer une classification des points de la courbe. Nous ne détaillons pas cet aspect.

Courbes planes dans l'espace

L'étude et la représentation d'une courbe paramétrée dans l'espace peuvent être notablement simplifiées si cette courbe appartient à une surface connue. Par exemple, il est beaucoup plus facile de tracer l'hélice circulaire dont un paramétrage est

$$\varphi : t \in \mathbb{R} \longmapsto \big(|a|\cos(t),\ a\sin(t),\ bt\big) \qquad (a, b \in \mathbb{R}^*)$$

lorsque l'on a remarqué que $\varphi_1^2(t) + \varphi_2^2(t) = a^2$, ce qui établit que cette courbe paramétrée est portée par le cylindre d'axe $(0, \mathbf{k})$ et de rayon a, voir la figure 12.

[60] La procédure MAPLE Dn est décrite p. 760.

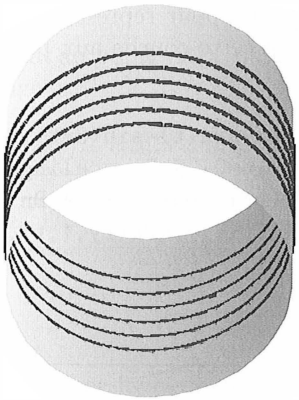


Fig. 12 Hélice circulaire tracée sur un cylindre.

Un cas particulièrement important en pratique est celui où la courbe paramétrée appartient à un plan ; on qualifie une telle courbe de *courbe plane*. La proposition suivante donne une caractérisation des courbes planes birégulières à l'aide du plan osculateur[61]. Elle indique qu'une courbe paramétrée birégulière est plane si et seulement si le plan osculateur est le même[62] en tout point de la courbe. La courbe est alors contenue dans ce plan osculateur.

PROPOSITION 11.8 *Soit C une courbe paramétrée birégulière de l'espace dont un paramétrage est $\varphi \in \mathcal{C}^2(I, \mathbb{R}^3)$. La courbe C est une courbe plane si et seulement si*

$$\exists t_0 \in I \quad \forall t \in I \quad Vect\{\varphi'(t), \varphi''(t)\} = Vect\{\varphi'(t_0), \varphi''(t_0)\} \tag{9}$$

où $Vect\{\varphi'(t), \varphi''(t)\}$ désigne l'espace vectoriel engendré par les vecteurs $\varphi'(t)$ et $\varphi''(t)$.

Démonstration $\trianglerighteq$ Montrons que si la condition (9) est satisfaite alors la courbe paramétrée C est contenue dans le plan osculateur (il s'agira donc d'une courbe plane). Comme la courbe paramétrée C est birégulière, le plan osculateur à la courbe paramétrée C en un point $M_0 = \varphi(t_0)$, où t_0 est fixé dans I, est l'ensemble, voir la définition 11.15,

$$\mathcal{P} = \left\{\mathbf{x} \in \mathbb{R}^3 \ ; \ \exists(\lambda, \mu) \in \mathbb{R}^2 \quad \mathbf{x} = \varphi(t_0) + \lambda\varphi'(t_0) + \mu\varphi''(t_0)\right\}.$$

[61] Rappelons que le plan osculateur d'une courbe paramétrée birégulière dont un paramétrage est $\varphi \in \mathcal{C}^2(I, \mathbb{R}^3)$ en $M_0 = \varphi(t_0)$ est un plan dirigé par les vecteurs $\varphi'(t_0)$ et $\varphi''(t_0)$ et passant par le point M_0.

[62] On commet ici un abus de langage, on devrait dire, dans un premier temps, qu'une courbe est plane si et seulement si la direction vectorielle du plan osculateur est constante.

Pour montrer que la courbe paramétrée C qui est l'ensemble $\{\varphi(t) \in \mathbb{R}^3 \ ; \ t \in I\}$ est contenue dans le plan osculateur $\mathcal{P}$, nous allons montrer que

$$\forall t \in I \qquad \exists (\lambda, \mu) \in \mathbb{R}^2 \quad \varphi(t) = \varphi(t_0) + \lambda \varphi'(t_0) + \mu \varphi''(t_0). \tag{10}$$

La condition (9) implique que pour tout $t \in I$, $\varphi'(t) \in \text{Vect}\{\varphi'(t_0), \varphi''(t_0)\}$ ce qui signifie qu'il existe deux réels ω_t et σ_t tels que

$$\varphi'(t) = \omega_t \varphi'(t_0) + \sigma_t \varphi''(t_0). \tag{11}$$

Commençons par étudier la régularité des applications $\Omega : t \in I \mapsto \omega_t \in \mathbb{R}$ et $\Sigma : t \in I \mapsto \sigma_t \in \mathbb{R}$ et montrons qu'elles sont de classe $\mathcal{C}^1$. Désignons par $\mathbf{w}$ un vecteur orthogonal à $\varphi''(t_0)$ mais non orthogonal à $\varphi'(t_0)$ (un tel vecteur existe car le point t_0 étant birégulier, les vecteurs $\varphi'(t_0)$ et $\varphi''(t_0)$ ne sont pas colinéaires). Pour tout $t \in I$, le produit scalaire de $\varphi'(t)$ avec $\mathbf{w}$ vaut, compte tenu de la relation (11),

$$\varphi'(t) \cdot \mathbf{w} = \omega_t \varphi'(t_0) \cdot \mathbf{w} + \sigma_t \underbrace{\varphi''(t_0) \cdot \mathbf{w}}_{=0} = \omega_t \varphi'(t_0) \cdot \mathbf{w}.$$

On en déduit que

$$\forall t \in I \qquad \Omega(t) = \frac{\varphi'(t) \cdot \mathbf{w}}{\varphi'(t_0) \cdot \mathbf{w}},$$

ce qui implique, puisque l'application φ est de classe $\mathcal{C}^2$ sur I, que l'application Ω est de classe $\mathcal{C}^1$ sur I. En considérant un vecteur $\mathbf{w}$ qui cette fois-ci est orthogonal à $\varphi'(t_0)$ mais non orthogonal à $\varphi''(t_0)$, on montre de la même manière que l'application Σ est de classe $\mathcal{C}^1$ sur I.

Pour tout $t \in I$, on a[63] $\varphi(t) = \varphi(t_0) + \int_{t_0}^t \varphi'(s) \ \mathrm{d}s$ et en utilisant la relation (11), on obtient

$$\varphi(t) = \varphi(t_0) + \left(\int_{t_0}^t \Omega(s) \ \mathrm{d}s \right) \varphi'(t_0) + \left(\int_{t_0}^t \Sigma(s) \ \mathrm{d}s \right) \varphi''(t_0).$$

Notons que comme les applications Ω et Σ sont de classe $\mathcal{C}^1$ sur I, elles sont Riemann-intégrables sur $[t_0, t]$ pour tout $t \in I$ et par conséquent les deux intégrales $\lambda = \int_{t_0}^t \Omega(s) \ \mathrm{d}s$ et $\mu = \int_{t_0}^t \Sigma(s) \ \mathrm{d}s$ existent. La relation (10) est établie et on peut conclure que la courbe paramétrée C est contenue dans le plan osculateur $\mathcal{P}$.

$\trianglerighteq$ Réciproquement, montrons que si une courbe plane C dont un paramétrage est $\varphi \in \mathcal{C}^2(I, \mathbb{R}^3)$ est une courbe birégulière alors la condition (9) est satisfaite. Par hypothèse, la courbe paramétrée C étant plane, elle appartient à un plan de $\mathbb{R}^3$ que nous noterons $\mathcal{P}_C$. Soit $t \in I$; comme φ est dérivable en t, on a[64]

$$\varphi'(t) = \lim_{h \to 0} \frac{\varphi(t+h) - \varphi(t)}{h}$$

---

[63] Pour des fonctions de $\mathbb{R}$ dans $\mathbb{R}$, cette relation a été établie dans le *Cours de première année*, p. 851. Ici l'application φ' est à valeurs dans $\mathbb{R}^3$. La quantité $\int_{t_0}^t \varphi'(s) \ \mathrm{d}s$ doit être comprise comme l'intégrale de chacune des composantes de φ'.

[64] Cette égalité définit la dérivée de φ en t, voir la définition 16.1 p. 713 du *Cours de première année*.

et par conséquent en considérant des valeurs de h de la forme $h = \frac{1}{n}$, où $n \in \mathbb{N}^*$, on obtient

$$\lim_{n \to +\infty} n\big(\varphi(t + \tfrac{1}{n}) - \varphi(t)\big) = \varphi'(t).$$

Puisque la courbe C est incluse dans le plan $\mathcal{P}_C$, pour tout $n \in \mathbb{N}^*$ tel que $t + \frac{1}{n} \in I$, on a $\varphi(t + \frac{1}{n}) \in \mathcal{P}_C$ et par conséquent $n(\varphi(t + \frac{1}{n}) - \varphi(t)) \in \mathcal{P}_C$. Le plan $\mathcal{P}_C$ étant un sous-ensemble fermé de l'espace $\mathbb{R}^3$, toute limite d'une suite d'éléments de $\mathcal{P}_C$ appartient[65] à $\mathcal{P}_C$; on a donc $\varphi'(t) \in \mathcal{P}_C$.

En effectuant un raisonnement analogue à partir de la relation

$$\varphi''(t) = \lim_{h \to 0} \frac{\varphi'(t+h) - \varphi'(t)}{h},$$

on montre que $\varphi''(t)$ appartient lui aussi à $\mathcal{P}_C$. Par ailleurs, comme la courbe paramétrée C est birégulière, les vecteurs $\varphi'(t)$ et $\varphi''(t)$ sont libres et engendrent un sous-espace vectoriel de dimension 2 dans $\mathbb{R}^3$. On en conclut que

$$\forall t \in I \quad \mathrm{Vect}\{\varphi'(t), \varphi''(t)\} = \mathcal{P}_C.$$

En désignant par t_0 un élément de I, on obtient

$$\forall t \in I \quad \mathrm{Vect}\{\varphi'(t), \varphi''(t)\} = \mathrm{Vect}\{\varphi'(t_0), \varphi''(t_0)\}$$

ce qui prouve l'assertion (9). □

Exemples

1. Montrons que la courbe paramétrée C de l'espace dont un paramétrage est

$$\varphi : t \in [0, 2\pi] \longmapsto (\sqrt{3}\cos(t),\ \cos(t),\ \sqrt{3}\sin(t))$$

est une courbe plane. Nous avons établi p. 774 que cette courbe était birégulière et que pour tout $t \in [0, 2\pi]$, on a

$$\varphi'(t) = \begin{pmatrix} -\sqrt{3}\sin(t) \\ -\sin(t) \\ \sqrt{3}\cos(t) \end{pmatrix} \quad \text{et} \quad \varphi''(t) = \begin{pmatrix} -\sqrt{3}\cos(t) \\ -\cos(t) \\ -\sqrt{3}\sin(t) \end{pmatrix}.$$

D'après la proposition 11.8, pour montrer que la courbe C est plane, on va montrer que

$$\forall t \in [0, 2\pi] \qquad \mathrm{Vect}\{\varphi'(t), \varphi''(t)\} = \mathrm{Vect}\{\varphi'(0), \varphi''(0)\}.$$

[65] Voir le théorème 7.6 p. 410.

Soient $t \in [0, \pi]$ et $\mathbf{u}$ un vecteur de l'espace vectoriel $\mathrm{Vect}\{\varphi'(t), \varphi''(t)\}$ engendré par $\varphi'(t)$ et $\varphi''(t)$; il existe $(\lambda, \mu) \in \mathbb{R}^2$ tel que

$$\begin{aligned}\mathbf{u} = \lambda\varphi'(t) + \mu\varphi''(t) &= \begin{pmatrix} -\lambda\sqrt{3}\sin(t) - \mu\sqrt{3}\cos(t) \\ -\lambda\sin(t) - \mu\cos(t) \\ \lambda\sqrt{3}\cos(t) - \mu\sqrt{3}\sin(t) \end{pmatrix} \\ &= (\lambda\cos(t) - \mu\sin(t)) \begin{pmatrix} 0 \\ 0 \\ \sqrt{3} \end{pmatrix} + (\lambda\sin(t) + \mu\cos(t)) \begin{pmatrix} -\sqrt{3} \\ -1 \\ 0 \end{pmatrix} \\ &= (\lambda\cos(t) - \mu\sin(t))\ \varphi'(0)\ + (\lambda\sin(t) + \mu\cos(t))\ \varphi''(0).\end{aligned}$$

Le vecteur $\mathbf{u}$ s'exprimant comme combinaison linéaire des vecteurs $\varphi'(0)$ et $\varphi''(0)$, il appartient à l'espace vectoriel $\mathrm{Vect}\{\varphi'(0), \varphi''(0)\}$. On en déduit que pour tout $t \in [0, 2\pi]$, on a $\mathrm{Vect}\{\varphi'(t), \varphi''(t)\} \subset \mathrm{Vect}\{\varphi'(0), \varphi''(0)\}$. L'égalité de ces deux espaces vectoriels découle du fait qu'ils sont tous les deux de dimension 2. La courbe paramétrée C est donc une courbe plane contenue dans le plan passant par le point de coordonnées $\varphi(0) = (\sqrt{3}, 1, 0)$ et dirigé par les vecteurs $\varphi'(0) = (0, 0, \sqrt{3})$ et $\varphi''(0) = (-\sqrt{3}, -1, 0)$.

On peut noter que, pour tout $t \in [0, 2\pi]$, on a

$$\varphi(t) = \big(\sqrt{3}\cos(t),\quad \cos(t),\quad \sqrt{3}\sin(t)\big)$$

ce qui indique que la courbe paramétrée C est une ellipse[66] dans le plan d'équation $y = x/\sqrt{3}$.

2. Montrons que quels que soient $A = (a_0, a_1, a_2) \in \mathbb{R}^3$, $\mathbf{u} = (b_0, b_1, b_2) \in \mathbb{R}^3$ et $\mathbf{v} = (c_0, c_1, c_2) \in \mathbb{R}^3$ non nuls, la courbe paramétrée dont un paramétrage est

$$\varphi : t \in [0, 1] \longmapsto \begin{pmatrix} a_0 + b_0 t + c_0 t^2 \\ a_1 + b_1 t + c_1 t^2 \\ a_2 + b_2 t + c_2 t^2 \end{pmatrix}$$

est une courbe plane. Pour $t \in [0, 1]$, le plan osculateur à la courbe paramétrée en $M = \varphi(t)$ est dirigé par les vecteurs

$$\varphi'(t) = \begin{pmatrix} b_0 + 2c_0 t \\ b_1 + 2c_1 t \\ b_2 + 2c_2 t \end{pmatrix} \quad \text{et} \quad \varphi''(t) = \begin{pmatrix} 2c_0 \\ 2c_1 \\ 2c_2 \end{pmatrix}.$$

Puisque $\varphi'(t) = \mathbf{u} + 2t\,\mathbf{v}$ et que $\varphi''(t) = 2\mathbf{v}$, on a

$$\forall t \in [0, 1] \quad \mathrm{Vect}\{\varphi'(t), \varphi''(t)\} = \mathrm{Vect}\{\mathbf{u}, \mathbf{v}\} = \mathrm{Vect}\{\varphi'(0), \varphi''(0)\}.$$

On déduit de la proposition 11.8 que la courbe est plane.

---

[66] Voir p. 725 la représentation paramétrique d'une ellipse.

Remarque Les courbes planes interviennent dans la description de nombreux phénomènes physiques. Citons à titre d'exemple :
– la *première loi de Kepler*[67] affirme que les planètes décrivent une courbe plane qui est une ellipse dont le soleil occupe un des foyers ;
– un fil pesant flexible infiniment mince homogène inextensible suspendu entre deux points et placé dans un champ de pesanteur uniforme décrit une courbe plane appelée chaînette ;
– les lignes de champ du champ magnétique créé par deux fils parallèles parcourus par un courant de même intensité et de même sens sont des courbes planes, contenues dans des plans orthogonaux aux fils, appelées *ovales de Cassini*, voir p. 719. ◇

Projection d'une courbe paramétrée sur un plan

Nous avons vu dans la partie précédente comment établir qu'une courbe paramétrée est une courbe plane. Lorsqu'une courbe de $\mathbb{R}^3$ n'est pas plane, il peut être intéressant pour son étude de la projeter sur certains plans choisis de manière idoine. Nous allons voir dans cette partie comment obtenir une représentation paramétrique du projeté d'une courbe paramétrée sur un plan donné.

Soient C une courbe paramétrée de $\mathbb{R}^3$ et $\mathcal{P}$ un plan de $\mathbb{R}^3$ dont l'équation cartésienne est $ax+by+cz = d$. Un vecteur normal au plan $\mathcal{P}$ est le vecteur $\mathbf{n} = (a, b, c)$. Considérons un point M de la courbe paramétrée C de coordonnées (x_0, y_0, z_0) et notons (x_1, y_1, z_1) les coordonnées du projeté orthogonal P du point M sur le plan $\mathcal{P}$, voir la figure 13. Le point P est caractérisé par le fait que le vecteur $\overrightarrow{MP}$ est colinéaire au vecteur $\mathbf{n}$ (i.e. il existe $\lambda \in \mathbb{R}$ tel que $\overrightarrow{MP} = \lambda\mathbf{n}$) et par le fait que le point P appartient au plan $\mathcal{P}$. Ces deux conditions impliquent que les coordonnées de P vérifient :

$$\begin{cases} x_1 - x_0 = \lambda a \\ y_1 - y_0 = \lambda b \\ z_1 - z_0 = \lambda c \end{cases} \quad \text{et} \quad ax_1 + by_1 + cz_1 = d.$$

On en déduit que le réel λ est donné par $\lambda = d - (ax_0 + by_0 + cz_0)$ et on a

$$\begin{cases} x_1 = x_0 + \lambda a \\ y_1 = y_0 + \lambda b \\ z_1 = z_0 + \lambda c \end{cases}. \tag{12}$$

[67] JOHANNES KEPLER (1571-1630) est un astronome resté célèbre pour avoir découvert que les planètes n'ont pas des orbites circulaires autour du soleil mais des orbites elliptiques et qui a proposé les relations mathématiques (dites *lois de Kepler*) qui régissent leur mouvement.

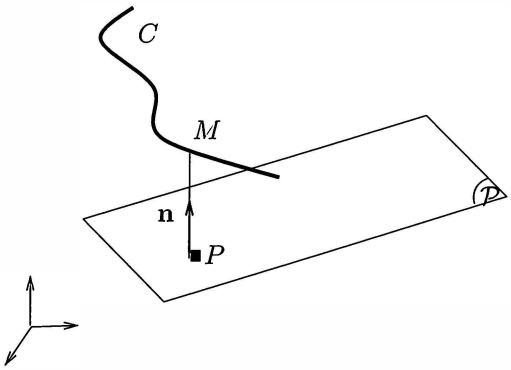


Fig. 13 Projection d'un point sur un plan.

On obtient le projeté de la courbe paramétrée C sur le plan $\mathcal{P}$ en projetant chacun de ses points. Si $\varphi : t \in I \mapsto (\varphi_1(t), \varphi_2(t), \varphi_3(t))$ est un paramétrage de la courbe paramétrée C alors, d'après les relations (12), le projeté de la courbe paramétrée C sur le plan $\mathcal{P}$ est la courbe paramétrée dont un paramétrage est

$$\psi : t \in I \longmapsto (\psi_1(t),\ \psi_2(t),\ \psi_3(t))$$

$$\text{où} \quad \begin{cases} \psi_1(t) = \varphi_1(t) + a\,\lambda(t) \\ \psi_2(t) = \varphi_2(t) + b\,\lambda(t) \\ \psi_3(t) = \varphi_3(t) + c\,\lambda(t) \end{cases} \quad \text{avec} \quad \lambda(t) = d - \big(a\varphi_1(t) + b\varphi_1(t) + c\varphi_1(t)\big).$$

On notera que ces relations indiquent que le projeté d'une courbe paramétrée de classe $\mathcal{C}^k$ sur un plan est une courbe paramétrée plane de classe $\mathcal{C}^k$. En pratique, plutôt que de retenir ces formules, il est préférable de refaire le raisonnement ayant permis de les établir dans chaque cas particulier rencontré.

Exemples

1. Déterminons le projeté orthogonal dans le plan d'équation $z = 0$ de la spirale dont une représentation paramétrique est

$$\varphi : t \in \mathbb{R} \longmapsto \big(\cos(t),\ \sin(t),\ t\big) \in \mathbb{R}^3.$$

Un vecteur normal au plan d'équation $z = 0$ est le vecteur $\mathbf{n} = (0, 0, 1)$. On a donc $c = 1$, $a = b = d = 0$. On en déduit que le projeté orthogonal de la spirale sur le plan d'équation $z = 0$ est la courbe paramétrée dont une représentation paramétrique est

$$\psi : t \in \mathbb{R} \longmapsto \big(\cos(t),\ \sin(t),\ 0\big) \in \mathbb{R}^3.$$

Il s'agit d'un cercle centré en $(0, 0, 0)$ et de rayon 1.

2. Intéressons-nous à présent à la projection de cette même spirale sur le plan $\mathcal{P}$ d'équation $x+y+z=1$. Un vecteur normal au plan $\mathcal{P}$ est le vecteur $\mathbf{n}=(1,1,1)$. Désignons par $\psi : t \in \mathbb{R} \mapsto (\psi_1(t),\psi_2(t),\psi_3(t)) \in \mathbb{R}^3$ un paramétrage du projeté de la spirale sur le plan $\mathcal{P}$. Pour tout $t \in \mathbb{R}$, le projeté du point $M=\varphi(t)=(\cos(t),\sin(t),t)$ est le point $P=\psi(t)$ de $\mathcal{P}$ tel que le vecteur $\overrightarrow{MP}$ est colinéaire à $\mathbf{n}$. Déterminer ce point revient à trouver $\lambda \in \mathbb{R}$ et $(\psi_1(t),\psi_2(t),\psi_3(t)) \in \mathbb{R}^3$ tels que

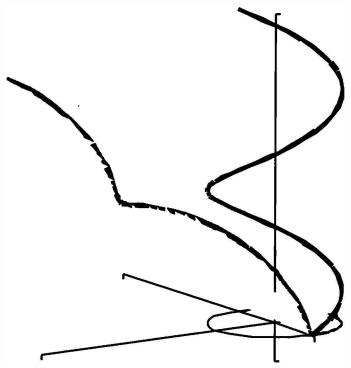
$$\begin{cases} \psi_1(t)=\cos(t)+\lambda \\ \psi_2(t)=\sin(t)+\lambda \\ \psi_3(t)=t+\lambda \end{cases} \quad \text{et} \quad \psi_1(t)+\psi_2(t)+\psi_3(t)=1.$$

On obtient $\lambda=\frac{1}{3}(1-\cos(t)-\sin(t)-t)$, ce qui donne pour paramétrage du projeté de la spirale sur le plan d'équation $x+y+z=1$:

$$\psi : t \in \mathbb{R} \longmapsto \frac{1}{3}\begin{pmatrix} 1+2\cos(t)-\sin(t)-t \\ 1-\cos(t)+2\sin(t)-t \\ 1-\cos(t)-\sin(t)+2t \end{pmatrix} \in \mathbb{R}^3.$$

On peut représenter graphiquement la spirale et son projeté dans le plan $\mathcal{P}$ avec MAPLE.

```
> des1:=spacecurve([cos(t),sin(t),t],t=0..3*Pi):
> des2:=spacecurve([cos(t),sin(t),0],t=0..3*Pi):
> des3:=plot3d(1-x-y,x=-4..1,y=-4..1,transparency=0.5):
> des4:=spacecurve([(1+2*cos(t)-sin(t)-t)/3,(1-cos(t)+2*sin(t)-t)/3,
                    (1-cos(t)-sin(t)+2*t)/3],t=0..3*Pi):
> display3d({des1,des2,des3,des4});
```



Représentation paramétrique d'une courbe plane dans un système de coordonnées adapté

Une courbe dans l'espace, même plane, n'est pas toujours simple à visualiser sans logiciel de visualisation graphique (considérer l'exemple précédent où la spirale est projetée sur le plan d'équation $x + y + z = 1$). En pratique, étant donné une courbe paramétrée plane, il peut être intéressant d'en obtenir une représentation paramétrique dans un système de coordonnées adapté au plan contenant la courbe. La méthode pour obtenir une telle représentation paramétrique consiste dans un premier temps à déterminer une base orthogonale de $\mathbb{R}^3$ dans laquelle le plan $\mathcal{P}$ a pour équation $Z = 0$. On procède pour cela de la manière suivante.

Considérons un plan $\mathcal{P}$ d'équation $ax + by + cz = d$ dans le repère canonique $(O; \mathbf{i}, \mathbf{j}, \mathbf{k})$ de $\mathbb{R}^3$ avec[68] $(a, b, c, d) \in \mathbb{R}^4$, $a^2 + b^2 + c^2 = 1$ et $c \geqslant 0$. On construit une base orthogonale $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ dans laquelle le plan $\mathcal{P}$ aura pour équation $Z = 0$ en choisissant $\mathbf{e}_3 = \mathbf{n} = (a, b, c)$, voir la figure 14. Le vecteur $\mathbf{e}_3$ est donc un vecteur normal au plan $\mathcal{P}$ de norme euclidienne égale à 1. On choisit ensuite $\mathbf{e}_1$ et $\mathbf{e}_2$ de sorte que $\mathbf{e}_1$ et $\mathbf{e}_2$ soient orthogonaux à $\mathbf{e}_3$ et orthogonaux entre eux. On peut choisir par exemple

$$\mathbf{e}_2 = \frac{1}{\sqrt{a^2+b^2}} \begin{pmatrix} -b \\ a \\ 0 \end{pmatrix} \quad \text{puis} \quad \mathbf{e}_1 = \mathbf{e}_2 \wedge \mathbf{e}_3 = \frac{1}{\sqrt{a^2+b^2}} \begin{pmatrix} ac \\ bc \\ -a^2-b^2 \end{pmatrix}.$$

Par construction, dans le repère orthonormé direct $(\Omega; \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ où Ω est un point quelconque du plan $\mathcal{P}$ de coordonnées $(x_\Omega, y_\Omega, z_\Omega)$ choisi pour origine, le plan $\mathcal{P}$ a pour équation $Z = 0$.

Considérons une courbe paramétrée C, contenue dans le plan $\mathcal{P}$, dont un paramétrage dans le repère canonique $(0; \mathbf{i}, \mathbf{j}, \mathbf{k})$ est $\varphi : t \in I \mapsto (\varphi_1(t), \varphi_2(t), \varphi_3(t))$ et déterminons un paramétrage de cette courbe dans le repère $(\Omega; \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ où $\Omega = \varphi(t_0)$ (avec $t_0 \in I$) est un point de la courbe paramétrée C.

Soit M un point de la courbe paramétrée C dont les coordonnées dans le repère $(O; \mathbf{i}, \mathbf{j}, \mathbf{k})$ sont (x, y, z) et dont les coordonnées dans le repère $(\Omega; \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ sont (X, Y, Z), voir la figure 14. On a $\overrightarrow{OM} = \overrightarrow{O\Omega} + \overrightarrow{\Omega M}$, c'est-à-dire

$$x\,\mathbf{i} + y\,\mathbf{j} + z\,\mathbf{k} = \overrightarrow{O\Omega} + X\,\mathbf{e}_1 + Y\,\mathbf{e}_2 + Z\,\mathbf{e}_3.$$

Dans la base canonique $(\mathbf{i}, \mathbf{j}, \mathbf{k})$, cette égalité s'écrit

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \varphi_1(t_0) \\ \varphi_2(t_0) \\ \varphi_3(t_0) \end{pmatrix} + X \frac{1}{\sqrt{a^2+b^2}} \begin{pmatrix} ac \\ bc \\ -a^2-b^2 \end{pmatrix} + Y \frac{1}{\sqrt{a^2+b^2}} \begin{pmatrix} -b \\ a \\ 0 \end{pmatrix} + Z \begin{pmatrix} a \\ b \\ c \end{pmatrix}.$$

[68] Notons que le plan $\mathcal{P}$ reste inchangé si on multiplie tous les coefficients de son équation cartésienne par un nombre réel non nul. Ainsi lorsque $a^2+b^2+c^2 \neq 1$ on peut diviser tous ses coefficients par cette quantité et lorsque $c \leqslant 0$ on peut multiplier tous les coefficients par -1. La condition imposée ici n'est donc nullement restrictive.

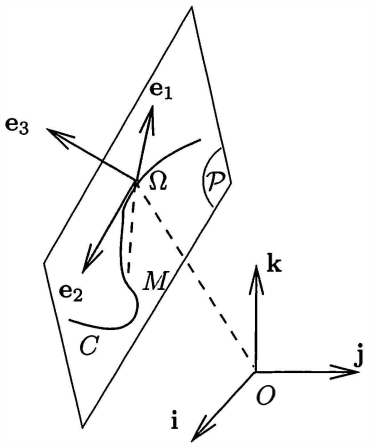


Fig. 14 Dans le repère $(\Omega; \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$, le plan $\mathcal{P}$ a pour équation $Z = 0$.

On dispose donc de l'expression des coordonnées (x, y, z) en fonction de (X, Y, Z). Cette relation permet également d'exprimer les coordonnées (X, Y, Z) en fonction de (x, y, z) en résolvant le système linéaire associé. On obtient[69], compte tenu du fait que $a^2 + b^2 + c^2 = 1$ et que le point M appartient au plan $\mathcal{P}$:

$$\begin{cases} X = \dfrac{1}{\sqrt{a^2+b^2}} \left(ac\,(x - \varphi_1(t_0)) + bc\,(y - \varphi_2(t_0)) - (a^2+b^2)(z - \varphi_3(t)) \right) \\ Y = \dfrac{1}{\sqrt{a^2+b^2}} \left(-b\,(x - \varphi_1(t_0)) + a\,(y - \varphi_2(t_0)) \right) \\ Z = 0 \end{cases}$$

On en conclut que la courbe paramétrée C admet pour paramétrage dans le repère $(\Omega; \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$

$$\psi : t \in I \longmapsto (X(t), Y(t), 0) \in \mathbb{R}^3 \tag{13}$$

où
$$\begin{cases} X(t) = \dfrac{1}{\sqrt{a^2+b^2}} \Big(ac\,(\varphi_1(t) - \varphi_1(t_0)) + bc\,(\varphi_2(t) - \varphi_2(t_0)) \\ \qquad\qquad - (a^2+b^2)(\varphi_3(t) - \varphi_3(t)) \Big) \\ Y(t) = \dfrac{1}{\sqrt{a^2+b^2}} \Big(-b\,(\varphi_1(t) - \varphi_1(t_0)) + a\,(\varphi_2(t) - \varphi_2(t_0)) \Big) \end{cases}$$

[69] On remarquera que ce résultat n'a de sens que si $a^2 + b^2 \neq 0$. Lorsque $a^2 + b^2 = 0$ alors le plan $\mathcal{P}$ a déjà pour équation $z = 0$ dans le repère initial.

Exemple Reprenons l'étude de la projection orthogonale sur le plan $\mathcal{P}$ d'équation $x + y + z = 1$ de la spirale effectuée p. 784. Dans le repère canonique $(O; \mathbf{i}, \mathbf{j}, \mathbf{k})$, un paramétrage du projeté de la spirale est,

$$\varphi : t \in \mathbb{R} \longmapsto \frac{1}{3} \begin{pmatrix} 1 + 2\cos(t) - \sin(t) - t \\ 1 - \cos(t) + 2\sin(t) - t \\ 1 - \cos(t) - \sin(t) + 2t \end{pmatrix} \in \mathbb{R}^3.$$

Déterminons un paramétrage du projeté de la spirale dans un repère local au plan $\mathcal{P}$. On a $a = b = c = 1/\sqrt{3}$. Choisissons pour origine le point $\Omega = \varphi(3\pi) = (-\frac{1}{3} - \pi, \frac{2}{3} - \pi, \frac{2}{3} + 2\pi)$ et pour vecteurs de la base du repère du plan $\mathcal{P}$ les vecteurs

$$\mathbf{e}_3 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{e}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \quad \text{et} \quad \mathbf{e}_1 = \mathbf{e}_2 \wedge \mathbf{e}_3 = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}.$$

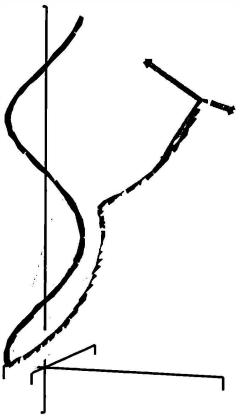
Un paramétrage du projeté de la spirale dans le repère $(\Omega; \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ est donné d'après (13) par

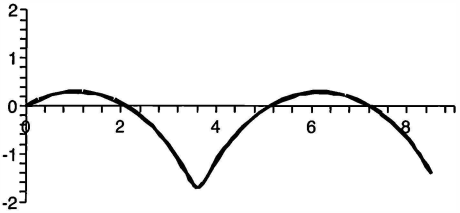
$$\psi : t \in \mathbb{R} \longmapsto \begin{pmatrix} \frac{1}{\sqrt{6}}(\sin(t) + \cos(t) + 1 - 2t + 6\pi) \\ \frac{1}{\sqrt{2}}(\sin(t) - \cos(t) - 1) \\ 0 \end{pmatrix}.$$

Bien entendu, dans la pratique on refera le raisonnement ayant conduit au paramétrage (13) sur chaque cas particulier rencontré.

On peut avoir recours à MAPLE pour illustrer la situation. Les instructions suivantes permettent de représenter d'une part la spirale, son projeté dans le plan $\mathcal{P}$ ainsi que le repère $(\Omega; \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ et d'autre part la courbe paramétrée associée à ψ.

```
> a:=1/sqrt(3):b:=1/sqrt(3):c:=1/sqrt(3):cste:=1/sqrt(a^2+b^2):
> e1:=[a*c*cste,b*c*cste,-(a^2+b^2)*cste]:
> e2:=[-b*cste,a*cste,0]:
> e3:=[a,b,c];
> des1:=spacecurve([cos(t),sin(t),t], t=0..3*Pi):
> des2:=arrow({fy(t0)},{e1,e2,e3}):
> des3:=plot3d(1-x-y,x=-4..1,y=-4..1,transparency=0.5):
> des4:=spacecurve([(1+2*cos(t)-sin(t)-t)/3,
    (1-cos(t)+2*sin(t)-t)/3,(1-cos(t)-sin(t)+2*t)/3],t=0..3*Pi):
> display3d({des1,des2,des3,des4},scaling=constrained);
> plot([(1+cos(t)+sin(t)-2*t+6*Pi)/sqrt(6),
       (-1-cos(t)+sin(t))/sqrt(2),t=0..3*Pi]);
```





11.4 Étude des nappes paramétrées

Rappelons[70] qu'une nappe paramétrée de classe $\mathcal{C}^k$ ($k \in \mathbb{N}^* \cup \{\infty\}$) de $\mathbb{R}^n$ ($n \in \mathbb{N}, n \geqslant 2$) est une application φ de classe $\mathcal{C}^k$ définie sur un sous-ensemble U de $\mathbb{R}^2$ et à valeurs dans $\mathbb{R}^n$. Les ensembles U sur lesquels sont définies les nappes paramétrées étudiées ici sont supposés connexes[71]. Une surface paramétrée correspond à l'image $\varphi(U) \subset \mathbb{R}^n$ d'une nappe paramétrée $\varphi \in \mathcal{C}^0(U, \mathbb{R}^n)$.

Souvent, une surface n'est pas donnée comme une image d'une nappe paramétrée mais est définie de manière descriptive. Par exemple : « la surface S est formée d'un segment horizontal s'appuyant sur l'axe d'un cylindre d'une part, et sur une hélice tracée sur ce cylindre d'autre part. Le pas de l'hélice étant de 2π et le rayon du cylindre $R = 1$ ». La surface S est représentée figure 15.

On peut toutefois dans de nombreux cas trouver une nappe paramétrée représentant la dite surface. Détaillons cette approche sur l'exemple cité. Tout d'abord, un paramétrage d'une hélice H de pas 2π et de rayon 1 est

$$t \in \mathbb{R} \longmapsto \big(\cos(t),\ \sin(t),\ t\big) \in \mathbb{R}^3.$$

D'après l'énoncé, un point M de coordonnées $(x, y, z) \in \mathbb{R}^3$ appartient à la surface S s'il est sur un segment d'extrémités N et P, N étant un point de l'hélice H et P le projeté de N sur l'axe $(0; \mathbf{k})$. Autrement dit,

$$M \in S \iff \exists N \in H \quad \exists \lambda \in [0,1] \quad \overrightarrow{OM} = \lambda \overrightarrow{ON} + (1-\lambda)\overrightarrow{OP}.$$

Le point N appartient à l'hélice H signifie qu'il existe $t \in \mathbb{R}$ tel que $N = (\cos(t), \sin(t), t)$ et le projeté P de N sur l'axe $(O; \mathbf{k})$ a pour coordonnées $(0, 0, t)$. On a donc,

$$M \in S \iff \exists t \in \mathbb{R} \quad \exists \lambda \in [0,1] \quad (x, y, z) = (\lambda \cos(t), \lambda \sin(t), t).$$

[70] Voir la définition 11.3 p. 716.

[71] Voir à ce sujet la remarque de la p. 717.

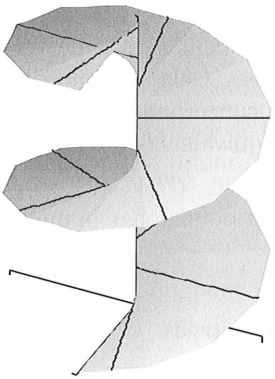


Fig. 15 Escalier hélicoïdal.

Une représentation paramétrique de la surface S est par conséquent

$$\varphi : (t, \lambda) \in \mathbb{R} \times [0, 1] \longmapsto \big(\lambda \cos(t),\ \lambda \sin(t),\ t\big) \in \mathbb{R}^3.$$

Exercice 11.7 Dans un repère orthonormé $(O; \mathbf{i}, \mathbf{j}, \mathbf{k})$, on note A le point de coordonnées $(1, 0, 1)$ et B celui de coordonnées $(0, 0, 1)$. Donner une représentation paramétrique de la surface S obtenue en faisant tourner le segment $[AB]$ autour de l'axe $(O; \mathbf{k})$.

11.4.1 Classification des nappes paramétrées

À l'instar des courbes paramétrées qui peuvent être représentées par plusieurs arcs paramétrés, plusieurs nappes paramétrées peuvent représenter une même surface paramétrée. On est alors amené à considérer une relation d'équivalence[72] sur l'ensemble des nappes paramétrées. Cette relation consiste à essayer de regrouper les nappes décrivant une même surface.

Définition 11.17 (Nappes paramétrées $\mathcal{C}^k$-équivalentes)
*Soient U et V deux ouverts connexes de $\mathbb{R}^2$, n un entier supérieur ou égal à 2 et $k \in \mathbb{N}^* \cup \{\infty\}$. On dit que la nappe paramétrée $\psi \in \mathcal{C}^0(V, \mathbb{R}^n)$ est $\mathcal{C}^k$-*équivalente *à la nappe paramétrée $\varphi \in \mathcal{C}^0(U, \mathbb{R}^n)$, s'il existe un $\mathcal{C}^k$-difféomorphisme θ de U sur V tel que $\varphi = \psi \circ \theta$.*

Si la nappe paramétrée $\psi \in \mathcal{C}^0(V, \mathbb{R}^n)$ est $\mathcal{C}^k$-équivalente à la nappe paramétrée $\varphi \in \mathcal{C}^0(U, \mathbb{R}^n)$, alors on dit que ψ se déduit de φ par le *changement de paramétrage* θ.

[72] Voir le *Cours de première année* p. 55.

Remarques

1. La relation « être $\mathcal{C}^k$-équivalent » est une relation d'équivalence sur l'ensemble des nappes paramétrées. Cela se démontre d'une manière très similaire à la proposition 11.2 p. 736 concernant les arcs paramétrés. On appelle *nappe géométrique* toute classe d'équivalence pour cette relation.

2. Si $\varphi \in \mathcal{C}^0(U, \mathbb{R}^n)$ et $\psi \in \mathcal{C}^0(V, \mathbb{R}^n)$ sont deux nappes paramétrées $\mathcal{C}^k$-équivalentes alors elles ont pour image la même surface paramétrée puisque $\varphi(U) = \psi(\theta(U)) = \psi(V)$. ◇

Orientation des nappes paramétrées

Un difféomorphisme étant une application dont le jacobien ne s'annule pas, le jacobien d'un difféomorphisme est une application à valeurs dans $\mathbb{R}$ qui ne change pas de signe[73]. On peut ainsi, comme on l'a fait pour les arcs paramétrés (voir la définition 11.8 p. 739), classer les nappes qui sont $\mathcal{C}^k$-équivalentes selon le signe du jacobien du changement de paramétrage.

Définition 11.18 (Nappes paramétrées $\mathcal{O}^k$-équivalentes)
Soient U et V deux ouverts connexes de $\mathbb{R}^2$, n un entier supérieur ou égal à 2 et $k \in \mathbb{N}^* \cup \{\infty\}$. On dit que les deux nappes paramétrées $\mathcal{C}^k$-équivalentes $\varphi \in \mathcal{C}^0(U, \mathbb{R}^n)$ et $\psi \in \mathcal{C}^0(V, \mathbb{R}^n)$ sont $\mathcal{O}^k$-équivalentes *s'il existe un $\mathcal{C}^k$-difféomorphisme θ de U sur V à jacobien positif tel que $\varphi = \psi \circ \theta$.*

On montre, comme pour les arcs paramétrés, voir le lemme 11.1 p. 740, qu'il existe au plus deux classes d'équivalence pour la relation $\mathcal{R}_{\mathcal{O}}$: « être $\mathcal{O}^k$-équivalent » parmi les nappes $\mathcal{C}^k$-équivalentes à une nappe donnée. On dit qu'une nappe paramétrée est *orientable* lorsque sa classe d'équivalence (i.e. la nappe géométrique qui lui est associée) pour la relation $\mathcal{R}$: « être $\mathcal{C}^k$-équivalent » possède deux classes d'équivalence pour la relation $\mathcal{R}_{\mathcal{O}}$. *Orienter* une nappe paramétrée c'est choisir l'une de ces deux classes.

Les propositions 11.3 et 11.5 affirment que tous les arcs paramétrés simples et que tous les arcs paramétrés réguliers sont orientables. On peut légitimement se demander s'il en est de même pour les nappes paramétrées. Nous allons voir que toutes les nappes paramétrées simples sont orientables mais qu'il existe des nappes paramétrées régulières qui ne le sont pas.

Proposition 11.9 *Toute nappe paramétrée simple[74] est orientable.*

---

[73] Cette propriété est vraie uniquement si le difféomorphisme est défini sur un ensemble connexe. Dans tout ce chapitre, nous considérons des nappes paramétrées définies sur des ensembles connexes, voir p. 717.

[74] Rappelons qu'une nappe paramétrée φ est dite simple si l'application φ est injective.

Démonstration La preuve de ce résultat, qui s'effectue en utilisant un raisonnement analogue à celui effectué pour établir la proposition 11.3, est laissée en exercice. □

Exercice 11.8 En s'inspirant de la preuve de la proposition 11.3, montrer que toute nappe paramétrée simple est orientable.

11.4.2 Régularité des nappes paramétrées

Comme pour les courbes paramétrées, on peut définir une notion de point régulier d'une surface paramétrée.

DÉFINITION 11.19 (Point régulier d'une surface paramétrée)
Soient $\varphi \in \mathcal{C}^1(U, \mathbb{R}^n)$ une nappe paramétrée et S la surface paramétrée de classe $\mathcal{C}^1$ qui lui est associée.

✗ *On dit que le point M_0 de S tel que $M_0 = \varphi(u_0, v_0)$ avec $(u_0, v_0) \in U$ est un point* régulier *de S pour le paramétrage φ si la famille*

$$\left\{ \frac{\partial \varphi}{\partial u}(u_0, v_0),\ \frac{\partial \varphi}{\partial v}(u_0, v_0) \right\}$$

est une famille libre.

✗ *Un point qui n'est pas régulier est appelé un* point singulier.

Remarques

1. Une surface paramétrée dont tous les points sont réguliers est appelée une *surface régulière.*

2. La matrice jacobienne de l'application $\varphi \in \mathcal{C}^1(U, \mathbb{R}^n)$ en $(u_0, v_0) \in U$ est la matrice $\mathrm{J}_\varphi(u_0, v_0) \in \mathcal{M}_{n,2}(\mathbb{R})$ définie par

$$\mathrm{J}_\varphi(u_0, v_0) = \begin{pmatrix} \partial_u \varphi_1(u_0, v_0) & \partial_v \varphi_1(u_0, v_0) \\ \partial_u \varphi_2(u_0, v_0) & \partial_v \varphi_2(u_0, v_0) \\ \vdots & \vdots \\ \partial_u \varphi_n(u_0, v_0) & \partial_v \varphi_n(u_0, v_0) \end{pmatrix}.$$

D'après la définition 11.19, un point $M_0 = \varphi(u_0, v_0)$ où $(u_0, v_0) \in U$ est un point régulier si et seulement si la matrice jacobienne $\mathrm{J}_\varphi(u_0, v_0)$ est de rang 2.

3. Dans le cas d'une surface paramétrée S dans l'espace (cas où $n = 3$), on peut également déterminer les points réguliers et les points singuliers de S à l'aide du produit vectoriel. Le point $M_0 = \varphi(u_0, v_0)$ avec $(u_0, v_0) \in U$ est un point singulier de S pour le paramétrage φ si la famille $\{\partial_u \varphi(u_0, v_0),\ \partial_v \varphi(u_0, v_0)\}$ est liée, autrement dit si les deux vecteurs $\partial_u \varphi(u_0, v_0)$ et $\partial_v \varphi(u_0, v_0)$ sont colinéaires. Or, dans l'espace, deux vecteurs sont colinéaires si et seulement si leur produit vectoriel est nul. Par conséquent, si

$$\frac{\partial \varphi}{\partial u}(u_0, v_0) \wedge \frac{\partial \varphi}{\partial v}(u_0, v_0) = \mathbf{0}$$

alors le point $M_0 = \varphi(u_0, v_0)$ est un point singulier et dans le cas contraire, il s'agit d'un point régulier. ◇

Comme toutes les notions définies à partir d'un paramétrage, il est important d'établir que la notion de point régulier ne dépend pas de la nappe paramétrée choisie comme paramétrage de la surface.

PROPOSITION 11.10 *Soient φ et ψ deux nappes paramétrées de classe $\mathcal{C}^1$ qui sont $\mathcal{C}^1$-équivalentes et S la surface paramétrée qui est leur est associée. Un point $M_0 \in S$ est régulier pour le paramétrage φ si et seulement si il est régulier pour le paramétrage ψ.*

Démonstration Soient $\varphi \in \mathcal{C}^1(U, \mathbb{R}^n)$ et $\psi \in \mathcal{C}^1(V, \mathbb{R}^n)$ deux nappes paramétrées $\mathcal{C}^1$-équivalentes et S la surface paramétrée qui est leur image commune ($S = \varphi(U) = \psi(V)$). D'après la définition 11.17, il existe un $\mathcal{C}^k$-difféomorphisme θ de U dans V tel que $\varphi = \psi \circ \theta$. Notons θ_1 et θ_2 les deux composantes de l'application θ ; pour tout $(u, v) \in U$ on a

$$\varphi(u, v) = \psi(\theta_1(u, v), \theta_2(u, v)).$$

En dérivant cette égalité par rapport aux variables u et v, on obtient[75] :

$$\left\{ \begin{array}{rcl} \dfrac{\partial \varphi}{\partial u}(u, v) & = & \dfrac{\partial \theta_1}{\partial u}(u, v)\ \dfrac{\partial \psi}{\partial s}(\theta(u, v)) + \dfrac{\partial \theta_2}{\partial u}(u, v)\ \dfrac{\partial \psi}{\partial t}(\theta(u, v)) \\ \dfrac{\partial \varphi}{\partial v}(u, v) & = & \dfrac{\partial \theta_1}{\partial v}(u, v)\ \dfrac{\partial \psi}{\partial s}(\theta(u, v)) + \dfrac{\partial \theta_2}{\partial v}(u, v)\ \dfrac{\partial \psi}{\partial t}(\theta(u, v)) \end{array} \right. . \tag{14}$$

Considérons un point M_0 de S tel que $M_0 = \varphi(u_0, v_0)$ avec $(u_0, v_0) \in U$ et tel que $M_0 = \psi(s_0, t_0)$ avec $(s_0, t_0) \in V$. On a $\theta(u_0, v_0) = (s_0, t_0)$ et les relations (14) évaluées en (u_0, v_0) s'écrivent

$$\left\{ \begin{array}{rcl} \dfrac{\partial \varphi}{\partial u}(u_0, v_0) & = & a\ \dfrac{\partial \psi}{\partial s}(s_0, t_0) + b\ \dfrac{\partial \psi}{\partial t}(s_0, t_0) \\ \dfrac{\partial \varphi}{\partial v}(u_0, v_0) & = & c\ \dfrac{\partial \psi}{\partial s}(s_0, t_0) + d\ \dfrac{\partial \psi}{\partial t}(s_0, t_0) \end{array} \right. \tag{15}$$

où $a = \partial_u \theta_1(u_0, v_0)$, $b = \partial_u \theta_2(u_0, v_0)$, $c = \partial_v \theta_1(u_0, v_0)$ et $d = \partial_v \theta_2(u_0, v_0)$.

Supposons que le point $M_0 = \varphi(u_0, v_0)$ est un point singulier de S pour le paramétrage φ. D'après la définition 11.19, la famille $\{\partial_u \varphi(u_0, v_0), \partial_v \varphi(u_0, v_0)\}$ est liée :

$$\exists (\lambda, \mu) \in \mathbb{R}^2_* \qquad \lambda \frac{\partial \varphi}{\partial u}(u_0, v_0) + \mu \frac{\partial \varphi}{\partial v}(u_0, v_0) = 0.$$

On déduit des relations (15) que

$$(\lambda a + \mu c)\ \frac{\partial \psi}{\partial s}(s_0, t_0) + (\lambda b + \mu d)\ \frac{\partial \psi}{\partial t}(s_0, t_0) = 0.$$

[75] On a noté $\frac{\partial \psi}{\partial s}$ (resp. $\frac{\partial \psi}{\partial t}$) la dérivée de l'application ψ par rapport au premier (resp. second) vecteur de la base canonique de $\mathbb{R}^2$.

Comme θ est un difféomorphisme de U sur V, son jacobien ne s'annule pas sur U. Par conséquent, les réels a, b, c et d ne sont pas tous nuls, ce qui implique que la famille $\{\partial_s\psi(s_0,t_0), \partial_t\psi(s_0,t_0)\}$ est liée. Le point M_0 est donc également un point singulier de S pour le paramétrage ψ.

On en conclut que tout point de S qui est singulier pour le paramétrage φ, l'est également pour le paramétrage ψ (et inversement). La contraposée de cette assertion indique que tout point de S qui est régulier pour le paramétrage ψ, l'est également pour le paramétrage φ (et inversement) □

Remarque La proposition 11.10 peut également être démontrée en utilisant des résultats d'algèbre linéaire et la caractérisation d'un point régulier à l'aide du rang, donnée dans la remarque de la p. 791. Si $\varphi \in \mathcal{C}^1(U,\mathbb{R}^n)$ et $\psi \in \mathcal{C}^1(V,\mathbb{R}^n)$ sont deux nappes paramétrées $\mathcal{C}^k$-équivalentes telles que $\varphi = \psi\circ\theta$ où θ est un $\mathcal{C}^k$-difféomorphisme de U sur V alors pour tout $(u_0,v_0) \in U$ on a

$$\mathrm{J}_\varphi(u_0,v_0) = \mathrm{J}_\psi(s_0,t_0) \times \mathrm{J}_\theta(u_0,v_0)$$

où $(s_0,t_0) = \theta(u_0,v_0)$ et où la matrice[(76)] $\mathrm{J}_\theta(u_0,v_0) \in \mathrm{GL}_2(\mathbb{R})$. On en déduit que les matrices $\mathrm{J}_\varphi(u_0,v_0)$ et $\mathrm{J}_\psi(s_0,t_0)$ ont même image et par conséquent que leurs rangs sont égaux. ◇

Illustrons par un exemple les deux méthodes pour déterminer les points réguliers d'une nappe paramétrée qui ont été mentionnées dans la remarque de la p. 791.

Exemples

1. L'escalier hélicoïdal est la surface S de $\mathbb{R}^3$ dont un paramétrage est[(77)]

$$\varphi : (u,v) \in \mathbb{R}\times[0,1] \longmapsto \big(v\cos(u),\ v\sin(u),\ u\big) \in \mathbb{R}^3.$$

Considérons le point $M_0 = \varphi(0,\frac{1}{2}) = (\frac{1}{2},0,0)$ de S. On a

$$\frac{\partial\varphi}{\partial u}(0,\tfrac{1}{2}) = \begin{pmatrix}0\\ \frac{1}{2}\\ 1\end{pmatrix}, \quad \frac{\partial\varphi}{\partial v}(0,\tfrac{1}{2}) = \begin{pmatrix}1\\ 0\\ 0\end{pmatrix} \quad \text{et} \quad J_\varphi(0,\tfrac{1}{2}) = \begin{pmatrix}0 & 1\\ \frac{1}{2} & 0\\ 1 & 0\end{pmatrix}.$$

Le déterminant d'ordre 2 correspondant à la matrice extraite de $J_\varphi(0,\frac{1}{2})$ en conservant les deux premières lignes et les deux colonnes vaut $-\frac{1}{2}$. Il est non nul, ce qui signifie que les deux vecteurs $\partial_u\varphi(0,\frac{1}{2})$ et $\partial_v\varphi(0,\frac{1}{2})$ sont linéairement indépendants et par conséquent que le point M_0 est un point régulier de S.

---

(76) Remarquer que la matrice $\mathrm{J}_\theta(u_0,v_0)$ correspond à la matrice $\begin{pmatrix}a & c\\ b & d\end{pmatrix}$ où les coefficients a, b, c et d sont définis dans la preuve de la proposition 11.10.

(77) Voir l'exemple de la p. 788.

Plus généralement, considérons un point $M = \varphi(u,v)$ avec $(u,v) \in \mathbb{R} \times [0,1]$ de S ; on a

$$\frac{\partial \varphi}{\partial u}(u,v) = \begin{pmatrix} -v\sin(u) \\ v\cos(u) \\ 1 \end{pmatrix}, \qquad \frac{\partial \varphi}{\partial v}(u,v) = \begin{pmatrix} \cos(u) \\ \sin(u) \\ 0 \end{pmatrix}$$

$$\text{et} \quad J_\varphi(u,v) = \begin{pmatrix} -v\sin(u) & \cos(u) \\ v\cos(u) & \sin(u) \\ 1 & 0 \end{pmatrix}.$$

Le déterminant correspondant à la matrice carrée d'ordre 2 extraite de $J_\varphi(u,v)$ en conservant les deux premières lignes et les deux colonnes vaut $-v$. Si $v \neq 0$ alors ce déterminant n'est pas nul et par conséquent le point $M = \varphi(u,v)$ est régulier. Si $v = 0$, alors les déterminants des trois matrices carrées d'ordre 2 extraites de $J_\varphi(u,v)$ en conservant successivement les lignes 1, 2, puis 2, 3 et 1, 3 valent

$$\begin{vmatrix} 0 & \cos(u) \\ 0 & \sin(u) \end{vmatrix} = 0, \quad \begin{vmatrix} 0 & \sin(u) \\ 1 & 0 \end{vmatrix} = -\sin(u), \quad \begin{vmatrix} 0 & \cos(u) \\ 1 & 0 \end{vmatrix} = -\cos(u).$$

Comme les fonctions sinus et cosinus ne s'annulent jamais simultanément, quel que soit la valeur de u, il existe au moins un de ces déterminants qui est non nul. On en déduit que tous les points de la surface paramétrée S sont réguliers.

2. Considérons la nappe paramétrée

$$\varphi : (u,v) \in [0, 2\pi[\times \mathbb{R}^+ \longmapsto \big(v\cos(u),\ v\sin(u),\ v\big) \in \mathbb{R}^3.$$

dont la surface paramétrée associée est un cône de sommet l'origine du repère et d'axe $(O;\mathbf{k})$. Pour tout $(u,v) \in [0, 2\pi[\times \mathbb{R}$, on a

$$\frac{\partial \varphi}{\partial u}(u,v) \wedge \frac{\partial \varphi}{\partial v}(u,v) = \begin{pmatrix} -v\sin(\theta) \\ v\cos(\theta) \\ 0 \end{pmatrix} \wedge \begin{pmatrix} \cos(u) \\ \sin(u) \\ 1 \end{pmatrix} = \begin{pmatrix} v\cos(u) \\ v\sin(u) \\ -v \end{pmatrix}.$$

Ce vecteur est nul si et seulement si $v = 0$. Comme pour tout $u \in [0, 2\pi[$, on a $\varphi(u,0) = (0,0,0)$, on en déduit que le sommet du cône est un point multiple et qu'il s'agit de l'unique point singulier du cône.

La proposition 11.5 indique que tout arc paramétré régulier est un arc orientable. La proposition suivante établit que cette propriété est mise en défaut pour les nappes paramétrées.

PROPOSITION 11.11
Il existe des nappes paramétrées régulières qui ne sont pas orientables.

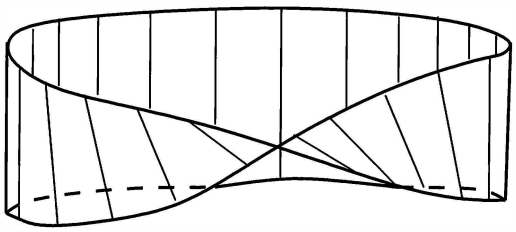


Fig. 16 Ruban de Möbius

Démonstration Pour prouver cette proposition, il suffit de trouver une nappe paramétrée qui ne soit pas orientable. Un exemple bien connu de nappe paramétrée non orientable est le *ruban de Möbius* représenté à la figure 16 dont un paramétrage est

$$\varphi : (u,v) \in U \longmapsto \begin{pmatrix} R\cos(2u) + v\sin(u)\cos(2u) \\ R\sin(2u) + v\sin(u)\sin(2u) \\ v\cos(u) \end{pmatrix}$$

où $U = \mathbb{R} \times [-\frac{1}{2}R, \frac{1}{2}R]$ et $R \in \mathbb{R}_+^*$. Montrons que la nappe paramétrée φ est non orientable et qu'elle est régulière.

$\trianglerighteq$ L'application $\vartheta : (u,v) \in U \mapsto (\pi + u, -v) \in U$ est un $\mathcal{C}^1$-difféomorphisme dont le jacobien vaut

$$\begin{vmatrix} \partial_u\vartheta_1(u,v) & \partial_v\vartheta_1(u,v) \\ \partial_u\vartheta_2(u,v) & \partial_v\vartheta_2(u,v) \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ 0 & -1 \end{vmatrix} = -1.$$

Les deux nappes φ et $\phi = \varphi \circ \vartheta$ sont, par définition, $\mathcal{C}^1$-équivalentes. De plus, pour tout $(u,v) \in U$ on a

$$\phi(u,v) = \varphi(\vartheta(u,v)) = \varphi(\pi + u, -v) = \varphi(u,v),$$

ce qui signifie que $\phi = \varphi$. Pour montrer que le ruban de Möbius n'est pas orientable, nous allons montrer que toute nappe paramétrée $\mathcal{C}^1$-équivalente à φ est $\mathcal{O}^1$-équivalente à φ. Cela établira qu'il existe une seule classe d'équivalence pour la relation $\mathcal{R}_{\mathcal{O}}$: « être $\mathcal{O}^1$-équivalent » parmi toutes les nappes paramétrées $\mathcal{C}^1$-équivalentes à φ et donc que φ n'est pas orientable.

Considérons une nappe paramétrée $\psi \in \mathcal{C}^1(V, \mathbb{R}^3)$ qui est $\mathcal{C}^1$-équivalente à φ ; il existe un $\mathcal{C}^1$-difféomorphisme θ de U sur V tel que $\psi = \varphi \circ \theta$. On distingue deux cas :

- Si θ a un jacobien qui est toujours positif, alors d'après la définition 11.18, ψ est $\mathcal{O}^1$-équivalente à φ.
- Si θ a un jacobien qui est toujours négatif, alors, comme $\psi = \varphi \circ \theta$ et $\phi = \varphi \circ \vartheta$, on a $\psi = \phi \circ \vartheta^{-1} \circ \theta$. L'application $\vartheta^{-1} \circ \theta$ qui est la composée de deux $\mathcal{C}^1$-difféomorphismes à jacobien négatif est un $\mathcal{C}^1$-difféomorphisme à jacobien positif. On en déduit que ψ est $\mathcal{O}^1$-équivalente à ϕ. Mais puisque $\phi = \varphi$, ψ est aussi $\mathcal{O}^1$-équivalente à φ.

MÖBIUS, August-Ferdinand (1790, Schulpforta - 1868, Leipzig).



August-Ferdinand Möbius, astronome et mathématicien allemand, fut directeur de l'observatoire de Leipzig. En astronomie, Möbius s'intéressa à la mécanique céleste et publia des travaux relatifs aux occultations des planètes. En mathématiques, les travaux de Möbius portent essentiellement sur la géométrie où il apporta une contribution importante à la géométrie projective[78]. Par ses études de surfaces atypiques, comme le ruban qui porte son nom et qui est une surface « à un seul côté », il fut un des précurseurs de la topologie. Le « ruban de Möbius » fut découvert indépendamment par Johann Benedict Listing (mathématicien allemand (1808-1882)) à peu près à la même époque.

Dans tous les cas, la nappe paramétrée ψ est $\mathcal{C}^1$-équivalente à φ. Il n'existe donc qu'une seule classe d'équivalence pour la relation $\mathcal{R}_{\mathcal{C}}$. Le ruban de Möbius n'est pas orientable.

$\trianglerighteq$ Montrons à présent que le ruban de Möbius est une nappe paramétrée régulière. Pour tout $(u,v) \in \mathbb{R}^2$, le produit vectoriel $\mathbf{N}(u,v)$ entre les deux vecteurs $\partial_u\varphi(u,v)$ et $\partial_v\varphi(u,v)$ vaut

$$\begin{pmatrix} -2R\sin(2u) + v\cos(u)\cos(2u) - 2v\sin(u)\sin(2u) \\ -2R\cos(2u) + v\cos(u)\sin(2u) + 2v\sin(u)\cos(2u) \\ -v\sin(u) \end{pmatrix} \wedge \begin{pmatrix} \sin(u)\cos(2u) \\ \sin(u)\sin(2u) \\ \cos(u) \end{pmatrix}$$
$$= \begin{pmatrix} 2R\cos(u)\cos(2u) + 2v\sin(u)\cos(u)\cos(2u) + v\sin(2u) \\ 2R\cos(u)\sin(2u) + 2v\sin(u)\cos(u)\sin(2u) - v\cos(2u) \\ -2R\sin(u) - 2v\sin(u)^2 \end{pmatrix}.$$

La nappe paramétrée φ est une nappe régulière si et seulement si pour tout $(u,v) \in \mathbb{R}^2$ le vecteur $\mathbf{N}(u,v)$ est non nul, autrement dit, si et seulement si sa norme euclidienne est non nulle. Après simplification, on obtient

$$\|\mathbf{N}(u,v)\|^2 = 4(R + v\sin(u))^2 + v^2,$$

ce qui prouve que pour tout $(u,v) \in U$ on a $\|\mathbf{N}(u,v)\| \neq 0$ car $R > 0$. On en conclut que le ruban de Möbius est une surface paramétrée régulière. □

Concrètement, on peut « fabriquer » un ruban de Möbius à l'aide d'une bande de papier dont on recolle les deux extrémités après avoir leur avoir fait subir un demi-tour. Contrairement à la même bande de papier recollée sans être vrillée, ce ruban n'a ni recto, ni verso. On s'en convainc aisément en traçant un trait au crayon sur le long du ruban de Möbius : en redécollant le ruban de Möbius, on s'aperçoit que le trait figure des deux côtés de la bande.

[78] La géométrie projective est le domaine des mathématiques qui modélise les notions intui-

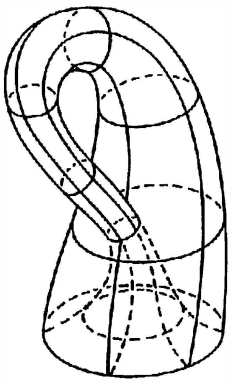


Fig. 17 Bouteille de Klein.

Notons qu'une autre surface non orientable, également très connue, est la surface appelée *bouteille de Klein*[79]. Elle peut être obtenue par recollement de deux anneaux de Möbius le long de leurs bords.

Il est néanmoins possible de comprendre la structure de la bouteille de Klein assez simplement. Imaginons un individu vivant dans un monde plat, à deux dimensions et essayons de lui expliquer ce qu'est un nœud. Pour cela, on lui dessine un nœud sur le plan : il ne voit qu'une courbe qui s'auto-intersecte. On lui explique alors que ce ne sont pas des points d'intersection qu'il voit, mais que la courbe passe au-dessus et au-dessous. Notre individu est interloqué puisque vivant dans un monde plat, il ne comprend pas ce qu'est le dessus ni ce qu'est le dessous. Il lui manque une dimension (le haut et le bas) pour pouvoir visualiser le nœud. Nous rencontrons le même problème lorsque nous essayons de visualiser la bouteille de Klein, puisque nous voyons une surface qui s'auto-intersecte, voir la figure 17. Néanmoins, si nous raisonnons avec une quatrième dimension, il suffit d'imaginer qu'aux endroits d'auto-intersection, la bouteille passe dessus et dessous au sens de cette quatrième dimension.

11.4.3 Étude des surfaces paramétrées dans l'espace

Intéressons-nous à présent plus spécifiquement aux surfaces paramétrées dans l'espace ($n = 3$).

---

tives de perspective et d'horizon. Un espace projectif est défini en mathématiques comme l'ensemble des droites vectorielles d'un espace vectoriel ; on peut imaginer l'œil d'un observateur placé sur l'origine d'un espace vectoriel, et chaque élément de l'espace projectif correspond à une direction de son regard.

[79] KLEIN, Felix (1849, Düsseldorf - 1925, Göttingen). Mathématicien allemand, connu pour ses travaux en théorie des groupes, en géométrie non euclidienne et en analyse. L'appellation « bouteille » proviendrait d'une erreur d'un traducteur qui aurait confondu en allemand « kleinsche Fläche » (surface de Klein) et « kleinsche Flasche » (bouteille de Klein).

DÉFINITION 11.20 (Plan tangent)
Soient $\varphi \in \mathcal{C}^1(U, \mathbb{R}^3)$ une nappe paramétrée et S la surface paramétrée de classe $\mathcal{C}^1$ qui lui est associée. On appelle plan tangent *à la surface paramétrée S en un point régulier $M_0 = \varphi(u_0, v_0)$ avec $(u_0, v_0) \in U$ de S, le plan passant par M_0 et de vecteurs directeurs $\partial_u\varphi(u_0, v_0)$ et $\partial_v\varphi(u_0, v_0)$.*

D'après la définition 11.20, le plan tangent à la surface paramétrée S en $M_0 = \varphi(u_0, v_0)$ est l'ensemble des points $M \in \mathbb{R}^3$ tels que

$$\exists (\alpha, \beta) \in \mathbb{R}^2 \quad \overrightarrow{MM_0} = \alpha\, \partial_u\varphi(u_0, v_0) + \beta\, \partial_v\varphi(u_0, v_0).$$

C'est donc l'ensemble

$$\mathcal{P} = \left\{ \mathbf{x} \in \mathbb{R}^3 \ ; \ \exists (\alpha, \beta) \in \mathbb{R}^2 \quad \mathbf{x} = \varphi(u_0, v_0) + \alpha\, \partial_u\varphi(u_0, v_0) + \beta\, \partial_v\varphi(u_0, v_0) \right\}.$$

L'espace vectoriel engendré par les vecteurs $\partial_u\varphi(u_0, v_0)$ et $\partial_v\varphi(u_0, v_0)$ coïncidant avec $\mathrm{Im}(\mathrm{J}_\varphi(u_0, v_0))$, l'image de la matrice jacobienne de φ en (u_0, v_0), on a aussi

$$\mathcal{P} = \left\{ \mathbf{x} \in \mathbb{R}^3 \ ; \ \exists\, \mathbf{y} \in \mathrm{Im}(\mathrm{J}_\varphi(u_0, v_0)) \quad \mathbf{x} = \varphi(u_0, v_0) + \mathbf{y} \right\}.$$

La proposition suivante indique que la notion de plan tangent en un point régulier d'une surface paramétrée est indépendante du paramétrage $\mathcal{C}^1$-équivalent choisi, autrement dit que l'on a le même plan tangent pour toutes les nappes paramétrées $\mathcal{C}^1$-équivalentes représentant une surface S.

PROPOSITION 11.12 *Soient $\varphi \in \mathcal{C}^1(U, \mathbb{R}^n)$ et $\psi \in \mathcal{C}^1(V, \mathbb{R}^n)$ deux nappes paramétrés $\mathcal{C}^1$-équivalentes et S la surface paramétrée qui est leur est associée. Si M_0 est un point régulier de S tel que $M_0 = \varphi(u_0, v_0)$ avec $(u_0, v_0) \in U$ et $M_0 = \psi(s_0, t_0)$ avec $(s_0, t_0) \in V$ alors le plan passant par M_0 et de vecteurs directeurs $\partial_u\varphi(u_0, v_0)$ et $\partial_v\varphi(u_0, v_0)$ est identique au plan passant par M_0 et de vecteurs directeurs $\partial_s\psi(s_0, t_0)$ et $\partial_t\varphi(s_0, t_0)$.*

Démonstration Puisque les deux nappes paramétrées $\varphi \in \mathcal{C}^1(U, \mathbb{R}^3)$ et $\psi \in \mathcal{C}^1(V, \mathbb{R}^3)$ sont $\mathcal{C}^k$-équivalentes, d'après la définition 11.17, il existe un $\mathcal{C}^k$-difféomorphisme θ de U dans V tel que $\varphi = \psi \circ \theta$ et on a d'après les relations (15) p. 792

$$\begin{cases} \partial_u\varphi(u_0, v_0) &= a\, \partial_s\psi(s_0, t_0) + b\, \partial_t\psi(s_0, t_0) \\ \partial_v\varphi(u_0, v_0) &= c\, \partial_s\psi(s_0, t_0) + d\, \partial_t\psi(s_0, t_0) \end{cases}$$

où $a = \partial_u\theta_1(u_0, v_0)$, $b = \partial_u\theta_2(u_0, v_0)$, $c = \partial_v\theta_1(u_0, v_0)$ et $d = \partial_v\theta_2(u_0, v_0)$.

Ces relations indiquent que l'espace vectoriel $\mathrm{Im}(\mathrm{J}_\varphi(u_0, v_0))$ engendré par les deux vecteurs $\partial_u\varphi(u_0, v_0)$ et $\partial_v\varphi(u_0, v_0)$ est égal à l'espace vectoriel $\mathrm{Im}(\mathrm{J}_\psi(s_0, t_0))$ engendré par les deux vecteurs $\partial_s\psi(s_0, t_0)$ et $\partial_t\psi(s_0, t_0)$. On en déduit que les deux plans tangents sont égaux et on peut en conclure que le plan tangent ne dépend pas du paramétrage choisi. □

La notion de plan tangent à une surface paramétrée n'est définie qu'en des points réguliers d'une surface paramétrée de classe $\mathcal{C}^1$. Nous avons vu qu'une façon de déterminer les points réguliers d'une surface paramétrée de classe $\mathcal{C}^1$ dans l'espace est de calculer le produit vectoriel

$$\mathbf{N}(u,v) = \partial_u\varphi(u,v) \wedge \partial_v\varphi(u,v)$$

où $\varphi \in \mathcal{C}^1(U,\mathbb{R}^3)$ désigne un paramétrage de la surface. Le point $M_0 = \varphi(u_0,v_0)$ est régulier si et seulement si $\mathbf{N}(u_0,v_0) \neq 0$. Cette façon de procéder pour déterminer les points réguliers d'une surface paramétrée possède l'avantage de fournir l'équation du plan tangent. En effet, si $M_0 = \varphi(u_0,v_0)$ est un point régulier alors $\mathbf{N}(u_0,v_0)$ est un vecteur normal au plan tangent. Le plan tangent en M_0 a donc pour équation cartésienne

$$a(x-x_0) + b(y-y_0) + c(z-z_0) = 0 \tag{16}$$

où (a,b,c) désignent les composantes de $\mathbf{N}(u_0,v_0)$ et (x_0,y_0,z_0) les coordonnées de M_0.

Remarque La notion de plan tangent pour des surfaces paramétrées coïncide avec la notion de plan tangent pour des surfaces cartésiennes (donnée à la définition 11.6). En effet, une surface cartésienne dans $\mathbb{R}^3$ d'équation $z = f(x,y)$ où f est une application définie sur un ouvert U de $\mathbb{R}^2$ et à valeurs dans $\mathbb{R}$ est aussi une surface paramétrée dont un paramétrage est $\varphi : (u,v) \in U \mapsto (u,v,f(u,v)) \in \mathbb{R}^3$. Pour cette nappe paramétrée, on a pour tout $(u,v) \in U$

$$\mathbf{N}(u,v) = \frac{\partial\varphi}{\partial u}(u,v) \wedge \frac{\partial\varphi}{\partial v}(u,v) = \begin{pmatrix} 1 \\ 0 \\ \partial_u f(u,v) \end{pmatrix} \wedge \begin{pmatrix} 0 \\ 1 \\ \partial_v f(u,v) \end{pmatrix} = \begin{pmatrix} -\partial_u f(u,v) \\ -\partial_v f(u,v) \\ 1 \end{pmatrix}.$$

Le vecteur $\mathbf{N}(u,v)$ n'est jamais nul, ce qui montre que tous les points d'une surface cartésienne sont réguliers au sens de la définition 11.19. D'après la relation (16), l'équation du plan tangent à la surface paramétrée au point $M_0 = \varphi(u_0,v_0) = (u_0,v_0,\varphi(u_0,v_0))$ est

$$(x-u_0)\,\partial_u f(u_0,v_0) + (y-v_0)\,\partial_v f(u_0,v_0) + (z-\varphi(u_0,v_0)) = 0.$$

On retrouve l'équation du plan tangent à la surface cartésienne d'équation $z = f(x,y)$ donnée p. 731. ◇

Exemples

1. Considérons la nappe paramétrée

$$\varphi : (u,v) \in \mathbb{R} \times [0,1] \longmapsto \big(v\cos(u),\ v\sin(u),\ u\big) \in \mathbb{R}^3.$$

Nous avons vu p. 793 que l'image de cette nappe paramétrée est l'escalier hélicoïdal et qu'il s'agit d'une surface paramétrée régulière.
Pour $(u,v) \in \mathbb{R} \times [0,1]$, on a

$$\frac{\partial\varphi}{\partial u}(u,v) \wedge \frac{\partial\varphi}{\partial v}(u,v) = \begin{pmatrix} -v\sin(u) \\ v\cos(u) \\ 1 \end{pmatrix} \wedge \begin{pmatrix} \cos(u) \\ \sin(u) \\ 0 \end{pmatrix} = \begin{pmatrix} -\sin(u) \\ \cos(u) \\ -v \end{pmatrix}.$$

Le plan tangent à l'escalier hélicoïdal en un point $M = \varphi(u,v)$ avec $(u,v) \in \mathbb{R} \times [0,1]$ a donc pour équation cartésienne

$$-\sin(u)\,(x - v\,\cos(u)) + \cos(u)\,(y - v\,\sin(u)) - v(z-u) = 0.$$

Au point M_0 de coordonnées $\varphi(0,\frac{1}{2}) = (\frac{1}{2},0,0)$, le plan tangent à l'escalier hélicoïdal a pour équation cartésienne $y - \frac{1}{2}z = 0$.

2. Cherchons à déterminer une équation du plan tangent en chaque point régulier d'un cône S dont un paramétrage est

$$\varphi : (u,v) \in [0,2\pi[\times\mathbb{R} \longmapsto \big(v\cos(u),\ v\sin(u),\ v\big) \in \mathbb{R}^3.$$

Nous avons établi p. 793 que tous les points du cône sont des points réguliers à l'exception de son sommet qui ici est l'origine et qui correspond à $\varphi(u,0)$, avec $u \in [0,2\pi[$. Pour tout $(u,v) \in [0,2\pi[\times\mathbb{R}^*$ on a

$$\mathbf{N}(u,v) = \frac{\partial\varphi}{\partial u}(u,v) \wedge \frac{\partial\varphi}{\partial v}(u,v) = \begin{pmatrix} -v\sin(u) \\ v\cos(u) \\ 0 \end{pmatrix} \wedge \begin{pmatrix} \cos(u) \\ \sin(u) \\ 1 \end{pmatrix} = \begin{pmatrix} v\cos(u) \\ v\sin(u) \\ -v \end{pmatrix}.$$

L'équation du plan tangent au cône en un point $M_0 = \varphi(u,v)$ distinct de l'origine est :

$$v\cos(u)\,(x - v\cos(u)) + v\sin(u)\,(y - v\sin(u)) - v\,(z - v) = 0$$

c'est-à-dire : $v\cos(u)\,x + v\sin(u)\,y - v\,z = 0$.

Exercice 11.9 On considère la nappe paramétrée[80]

$$\varphi : (u,v) \in [0,2\pi[\times[0,1] \longmapsto \begin{pmatrix} v\cos(u) - (1-v)\sin(u) \\ v\sin(u) + (1-v)\cos(u) \\ v \end{pmatrix}.$$

1 - Montrer que la surface paramétrée S associée à φ est régulière.
2 - Déterminer l'équation du plan tangent à S en chacun de ses points.

Remarque Si une nappe paramétrée $\varphi \in \mathcal{C}^1(U,\mathbb{R}^n)$ est orientable alors l'application $\mathbf{N} : (u,v) \in U \mapsto \partial_u\varphi(u,v) \wedge \partial_v\varphi(u,v)$ est continue et on a

$$\varphi(u,v) = \varphi(\widetilde{u},\widetilde{v}) \quad \Longrightarrow \quad \mathbf{N}(u,v) = \mathbf{N}(\widetilde{u},\widetilde{v})$$

pour presque[81] tout $((u,v),(\widetilde{u},\widetilde{v})) \in U^2$. Autrement dit, si une surface est orientable alors on peut associer à chaque point $M \in S$ un vecteur normal à la surface S et cette association définit une application continue. ◇

[80] Voir l'exercice 11.7 p. 789.

[81] C'est-à-dire pour tout $((u,v),(\widetilde{u},\widetilde{v})) \in U^2$ sauf éventuellement pour $((u,v),(\widetilde{u},\widetilde{v})) \in V^2$ où V est un sous-ensemble de U d'intérieur vide.

11.4.4 Arcs tracés sur une nappe

Dans ce chapitre, nous avons défini d'une part les courbes comme des objets à un paramètre t et d'autre part les surfaces comme des objets à deux paramètres u et v indépendants. Une question naturelle qui se pose est la suivante : que se passe-t-il si on fige l'un des paramètres définissant une surface ou si on lie les deux paramètres définissant cette surface ? On obtient une courbe (il n'existe plus qu'un seul paramètre libre) et cette courbe est tracée sur la surface initiale.

Un exemple classique de courbes tracées sur une surface est celui des *méridiens* et des *parallèles* qui sont des courbes tracées sur la surface du globe terrestre assimilé à une sphère. En prenant pour paramétrage de cette surface

$$\varphi : (u,v) \in [0,2\pi[\times[0,\pi[\longmapsto \big(R\cos(u)\sin(v),\ R\sin(u)\sin(v),\ R\cos(v)\big)\,,$$

les méridiens, qui sont des demi-cercles joignant les deux pôles, correspondent aux courbes obtenues en figeant u (autrement dit en supposant que u prend une valeur fixée $u_0 \in [0,2\pi[$) et ont pour paramétrage

$$\varphi_{u_0} : v \in [0,\pi[\longmapsto \big(R\cos(u_0)\sin(v),\ R\sin(u_0)\sin(v),\ R\cos(v)\big)\,,$$

alors que les parallèles, qui sont des cercles parallèles à l'équateur, correspondent aux courbes obtenues en figeant v (autrement dit en supposant que u prend une valeur fixée $v_0 \in [0,\pi[$) et ont pour paramétrage

$$\varphi_{v_0} : u \in [0,2\pi[\longmapsto \big(R\cos(u)\sin(v_0),\ R\sin(u)\sin(v_0),\ R\cos(v_0)\big)\,,$$

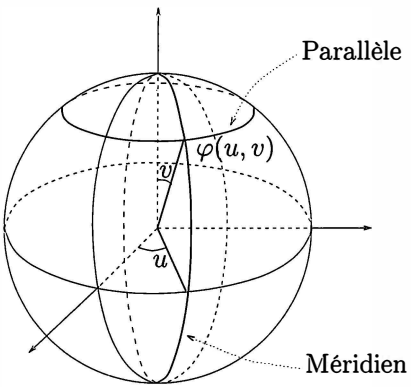


Fig. 18 Exemple de courbes tracées sur une surface : les méridiens et parallèles.

DÉFINITION 11.21 (Courbe tracée sur une surface)
Soient $\varphi \in \mathcal{C}^0(U, \mathbb{R}^3)$ une nappe paramétrée et S la surface paramétrée qui lui est associée. On appelle courbe tracée sur la surface S, *toute courbe paramétrée de $\mathbb{R}^3$ dont un paramétrage est*

$$\xi : t \in I \longmapsto \varphi(f(t)) \in \mathbb{R}^3$$

où I est un intervalle de $\mathbb{R}$ et f une application continue de I dans U.

Remarques

1. Figer l'un des paramètres comme cela a été fait pour les méridiennes et parallèles, revient à considérer des courbes paramétrées où l'application f est respectivement de la forme $f : t \mapsto (u_0, t)$ ou $f : t \mapsto (t, v_0)$ avec (u_0, v_0) fixé dans U. Ces lignes particulières sont appelées *lignes coordonnées* tracées sur la surface.

2. Les surfaces cartésiennes peuvent être considérées comme des surfaces paramétrées comme cela a été indiqué p. 722. Pour ces surfaces, la définition 11.21 coïncide avec la définition des lignes coordonnées donnée p. 728. ◇

La notion de lignes coordonnées sur une surface paramétrée dépend du paramétrage choisi. Si deux nappes paramétrées $\mathcal{C}^1$-équivalentes (distinctes) ont même image, alors les lignes coordonnées obtenues pour l'une des nappes paramétrées ne coïncident pas nécessairement avec les lignes coordonnées obtenues pour l'autre nappe paramétrée.

Ainsi, la nappe paramétrée

$$\varphi : (u, v) \in \mathbb{R}^2 \longmapsto \big(R\cos(u+v)\sin(v),\ R\sin(u+v)\sin(v),\ R\cos(v)\big)$$

est un paramétrage de la sphère de centre $(0, 0, 0)$ et de rayon $R \in \mathbb{R}_+^*$, dont les lignes coordonnées obtenues en figeant $u = u_0 \in \mathbb{R}$ ne sont ni des méridiens, ni des parallèles.

La proposition suivante indique que la droite tangente en un point régulier M_0 d'une courbe C tracée sur une surface paramétrée S est incluse dans le plan tangent à la surface en ce point si celui-ci est également un point régulier de la surface, voir la figure 19.

PROPOSITION 11.13 *Soient C une courbe tracée sur une surface paramétrée S de $\mathbb{R}^3$ et M_0 un point de C. Si le point M_0 est à la fois un* point régulier *de la courbe C et un* point régulier *de la surface S alors la tangente à la courbe C en M_0 est incluse dans le plan tangent à la surface S en M_0.*

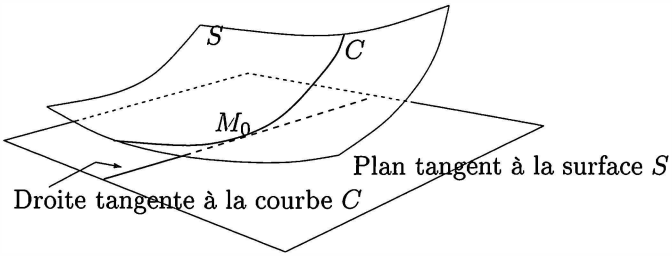


Fig. 19 La droite tangente en M_0 est incluse dans le plan tangent en M_0.

Démonstration Soient $\varphi \in \mathcal{C}^1(U, \mathbb{R}^3)$ un paramétrage de la surface S et $\xi = \varphi \circ f$, où $f \in \mathcal{C}^1(I, \mathbb{R}^2)$ est une application d'un intervalle I de $\mathbb{R}$ à valeurs dans $U \in \mathbb{R}^2$, un paramétrage de la courbe C tracée sur S. Considérons un point M_0 qui est à la fois un point régulier de la courbe C et un point régulier de la surface S. D'une part, il existe $t_0 \in I$ tel que $M_0 = \xi(t_0)$ avec $\xi'(t_0) \neq 0$ et d'autre part, il existe $(u_0, v_0) \in U$ tel que $M_0 = \varphi(u_0, v_0)$ et la famille $\{\partial_u\varphi(u_0, v_0), \partial_v\varphi(u_0, v_0)\}$ est une famille libre.

D'après la la définition 11.12, la tangente à la courbe C en M_0 est

$$\mathcal{T} = \left\{\mathbf{x} \in \mathbb{R}^3 \ ; \ \exists\lambda \in \mathbb{R} \quad \mathbf{x} = \xi(t_0) + \lambda\,\xi'(t_0)\right\}$$

et d'après la définition 11.19, le plan tangent à la surface S en M_0 est

$$\mathcal{P} = \left\{\mathbf{x} \in \mathbb{R}^3 \ ; \ \exists(\alpha, \beta) \in \mathbb{R}^2 \quad \mathbf{x} = \varphi(u_0, v_0) + \alpha\,\partial_u\varphi(u_0, v_0) + \beta\,\partial_v\varphi(u_0, v_0)\right\}.$$

Avec ces notations, démontrer la proposition revient à monter que $\mathcal{T} \subset \mathcal{P}$. Considérons un élément $\mathbf{x}$ de $\mathcal{T}$ et montrons que $\mathbf{x}$ appartient à $\mathcal{P}$. Puisque $\mathbf{x} \in \mathcal{T}$, il existe $\lambda \in \mathbb{R}$ tel que $\mathbf{x} = \xi(t_0) + \lambda\,\xi'(t_0)$. Désignons par f_1 et f_2 les deux applications coordonnées de f et explicitons $\xi'(t_0)$. Pour tout $t \in I$, on a $\xi(t) = \varphi(f_1(t), f_2(t))$ d'où en utilisant la formule de dérivation composée[82]

$$\xi'(t) = \frac{\partial f_1}{\partial t}(t)\,\frac{\partial\varphi}{\partial u}(f_1(t), f_2(t)) + \frac{\partial f_2}{\partial t}(t)\,\frac{\partial\varphi}{\partial v}(f_1(t), f_2(t)).$$

On en déduit que

$$\xi'(t_0) = \frac{\partial f_1}{\partial t}(t_0)\,\frac{\partial\varphi}{\partial u}(u_0, v_0) + \frac{\partial f_2}{\partial t}(t_0)\,\frac{\partial\varphi}{\partial v}(u_0, v_0)$$

puis que

$$\begin{aligned}\mathbf{x} &= \xi(t_0) + \lambda\frac{\partial f_1}{\partial t}(t_0)\,\frac{\partial\varphi}{\partial u}(u_0, v_0) + \lambda\frac{\partial f_2}{\partial t}(t_0)\,\frac{\partial\varphi}{\partial v}(u_0, v_0)\\ &= \varphi(u_0, v_0) + \alpha\,\frac{\partial\varphi}{\partial u}(u_0, v_0) + \beta\,\frac{\partial\varphi}{\partial v}(u_0, v_0)\end{aligned}$$

où on a noté $\alpha = \partial_t f_1(t_0)$ et $\beta = \partial_t f_2(t_0)$. On a donc bien $\mathbf{x} \in \mathcal{P}$. □

[82] Voir le chap. 10 p. 637.

Remarque Sous les hypothèses de la proposition 11.13, on peut montrer que le plan tangent à S en M_0 est la réunion de toutes les tangentes aux courbes régulières tracées sur S et passant par M_0. ◇

11.5 Théorème d'inversion locale en géométrie

11.5.1 Application du théorème d'inversion locale à l'étude des courbes paramétrées

Nous avons vu qu'il existait pour une même courbe plusieurs descriptions possibles : sous forme paramétrée, sous forme cartésienne et sous forme implicite, et nous avons déjà mentionné p. 722 que toute courbe cartésienne admettait une représentation paramétrique. Dans cette partie, nous allons établir à quelles conditions une courbe paramétrée admet une représentation cartésienne. Commençons par illustrer la problématique sur un exemple en considérant la courbe paramétrée C du plan dont un paramétrage est

$$\varphi : t \in [0,1] \longmapsto \big(t^2,\ \sin(t)\big) \in \mathbb{R}^2.$$

Si un point $(x,y) \in \mathbb{R}^2$ appartient à la courbe paramétrée C alors il existe une valeur du paramètre $t \in [0,1]$ tel que $x = t^2 \in [0,1]$ et $y = \sin(t) \in [0, \sin(1)]$. Ainsi, on a $t = \sqrt{x}$ et une relation liant les coordonnées x et y est : $y = \sin(\sqrt{x})$. La courbe paramétrée C admet donc une représentation cartésienne : c'est le graphe de l'application $f : x \in [0,1] \mapsto \sin(\sqrt{x})$.

On peut se demander si cette approche peut toujours être mise en œuvre ou à quelles conditions elle peut l'être. Le théorème d'inversion locale[83] est un outil permettant d'aborder ce problème. Dans la situation présente, nous allons avoir recours au théorème d'inversion locale en dimension 1 qui prend la forme suivante.

THÉORÈME 11.1 (Théorème d'inversion locale en dimension 1)
Soient I un intervalle de $\mathbb{R}$, t_0 un élément de I et f une application de I dans $\mathbb{R}$ de classe $\mathcal{C}^1$. Si $f'(t_0) \neq 0$ alors il existe un voisinage I_0 de t_0 et un voisinage J_0 de $f(t_0)$ tels que f soit un $\mathcal{C}^1$-difféomorphisme de I_0 sur J_0.

Regardons comment le théorème d'inversion locale permet d'établir des conditions d'existence d'une représentation cartésienne pour une courbe paramétrée. Considérons un arc paramétré $\varphi : t \in I \mapsto (\varphi_1(t), \varphi_2(t)) \in \mathbb{R}^2$ de classe $\mathcal{C}^1$ et un point $M_0 = \varphi(t_0)$ où $t_0 \in I$ de la courbe paramétrée C associée à φ.

Si on suppose que $\varphi_1'(t_0) \neq 0$ alors, d'après le théorème d'inversion locale, il existe un voisinage I_0 de t_0 et un voisinage J_0 de $\varphi_1(t_0)$ tel que φ_1 soit un $\mathcal{C}^1$-difféomorphisme de I_0 sur J_0. On en déduit que la fonction φ_1 est « localement inversible » au voisinage de t_0, c'est-à-dire que

$$\forall t \in I_0 \quad \exists x \in J_0 \quad t = \varphi_1^{-1}(x).$$

[83] Ce théorème est énoncé en toute généralité p. 665.

Ainsi, pour tout $t \in I_0$ on a $\varphi_2(t) = \varphi_2(\varphi_1^{-1}(x))$, ce qui signifie qu'au voisinage de $M_0 = \varphi(t_0)$ la courbe paramétrée C admet une représentation cartésienne d'équation $y = \varphi_2(\varphi_1^{-1}(x))$; la courbe paramétrée C est donc le graphe de l'application $\varphi_2 \circ \varphi_1^{-1}$ sur J_0.

Notons que le même raisonnement peut être effectué en inversant le rôle de φ_1 et φ_2 si $\varphi_2'(t_0) \neq 0$. On obtient dans ce cas que la courbe paramétrée C admet au voisinage de $M_0 = \varphi(t_0)$ une représentation cartésienne d'équation $x = \varphi_1(\varphi_2^{-1}(y))$. La courbe paramétrée C est donc le graphe de l'application $\varphi_1 \circ \varphi_2^{-1}$ sur J_0 considérée relativement au repère $(O, -\mathbf{j}, \mathbf{i})$.

On en conclut qu'il est possible, au voisinage du point $M_0 = \varphi(t_0)$, d'obtenir une représentation cartésienne de l'arc paramétré $\varphi \in \mathcal{C}^1(I, \mathbb{R}^2)$ si $\varphi_1'(t_0) \neq 0$ ou si $\varphi_2'(t_0) \neq 0$, autrement dit si M_0 est un point régulier de la courbe paramétrée.

On est donc en mesure d'énoncer le résultat suivant.

PROPOSITION 11.14 *Soit C une courbe paramétrée de $\mathbb{R}^2$. Si $M_0 \in C$ est un point régulier de C alors la courbe paramétrée C admet,* au voisinage du point M_0, *une représentation cartésienne.*

Exemple Considérons l'arc paramétré $\varphi : t \in [0, 2\pi[\mapsto (\cos(t), \sin(t))$ dont l'image est le cercle C centré à l'origine et de rayon 1. Pour $t \in [0, 2\pi[$ on a $\varphi'(t) = (-\sin(t), \cos(t))$ et ce vecteur ne s'annule jamais, ce qui permet d'établir que tous les points du cercle sont des points réguliers pour le paramétrage φ. On en déduit que la courbe paramétrée C admet une représentation cartésienne au voisinage de chacun de ses points. Considérons les différentes situations.

Pour $t \in]0, \pi[$, on a $\varphi_1'(t) = \sin(t) > 0$ et

$$\varphi_2(t) = \sin(t) = \sqrt{1 - \cos^2(t)} = \sqrt{1 - \varphi_1(t)^2}.$$

Au voisinage des points M du cercle tels que $M = \varphi(t)$ avec $t \in]0, \pi[$, la courbe paramétrée C est le graphe de l'application $f : x \in]-1, 1[\mapsto \sqrt{1 - x^2}$.

Pour $t \in]\pi, 2\pi[$, on a $\varphi_1'(t) = \sin(t) < 0$ et

$$\varphi_2(t) = \sin(t) = -\sqrt{1 - \cos^2(t)} = -\sqrt{1 - \varphi_1(t)^2}.$$

Au voisinage des points M du cercle tels que $M = \varphi(t)$ avec $t \in]\pi, 2\pi[$, la courbe paramétrée C est le graphe de l'application $f : x \in]-1, 1[\mapsto -\sqrt{1 - x^2}$.

Pour $t = \pi$, on a $\varphi_1'(\pi) = 0$ mais $\varphi_2'(\pi) = 1$ et pour tout $t \in]\frac{1}{2}\pi, \frac{3}{2}\pi[$ on a

$$\varphi_1(t) = \cos(t) = -\sqrt{1 - \sin^2(t)} = -\sqrt{1 - \varphi_2(t)^2}.$$

On en déduit qu'au voisinage du point $M = \varphi(\pi) = (-1, 0)$, la courbe paramétrée C est le graphe de l'application $f : x \mapsto -\sqrt{1 - x^2}$ relativement au repère $(O, -\mathbf{j}, \mathbf{i})$. On établit de la même manière qu'au voisinage du point $M = \varphi(0) = (1, 0)$, la courbe paramétrée C est le graphe de l'application $f : x \mapsto \sqrt{1 - x^2}$ relativement au repère $(O, -\mathbf{j}, \mathbf{i})$.

Remarques

1. La proposition 11.14 assure l'existence d'une représentation cartésienne pour une courbe paramétrée au voisinage d'un point régulier. La détermination effective de l'équation cartésienne n'est pas toujours aussi aisée que dans les exemples traités ici. Il s'agit généralement d'une tâche relativement difficile.

2. La proposition 11.14 donne une condition suffisante pour qu'une courbe paramétrée admette une représentation cartésienne au voisinage d'un de ses points. Cette condition n'est pas nécessaire : la courbe paramétrée dont un paramétrage est $\varphi : t \in \mathbb{R} \mapsto (t^2,\ t^6) \in \mathbb{R}^2$ est le graphe de l'application $f : x \in \mathbb{R} \mapsto x^3$ bien que l'origine soit un point singulier de la courbe.

3. La proposition 11.14 assure l'existence d'une représentation cartésienne au voisinage d'un point régulier. Cette représentation n'est que « locale » et dans la plupart des cas il n'existe pas de représentation cartésienne « globale » valable pour toute la courbe paramétrée. C'est le cas par exemple du cercle ; il admet pour paramétrage $\varphi : t \in [0, 2\pi[\mapsto (\cos(t), \sin(t))$ mais n'est le graphe d'aucune fonction de $\mathbb{R}$ dans $\mathbb{R}$ bien que tous ses points soient des points réguliers pour φ. ◇

Le théorème d'inversion globale énoncé p. 665 donne une condition suffisante pour qu'une courbe paramétrée de classe $\mathcal{C}^1$ dont un paramétrage est $\varphi : t \in I \mapsto (\varphi_1(t), \varphi_2(t))$ admette une même représentation cartésienne pour toute la courbe : il suffit que l'application φ_1 soit injective sur I et que φ_1' ne s'annule pas sur I ou que φ_2 soit injective sur I et que φ_2' ne s'annule pas sur I. Comme une fonction réelle continue qui ne s'annule pas sur un intervalle est de signe constant, on peut affirmer que, si φ_1 est de classe $\mathcal{C}^1$ et si φ_1' ne s'annule pas alors φ_1 est strictement monotone et par conséquent injective. Ainsi, une condition suffisante pour qu'une courbe paramétrée par φ admette une représentation cartésienne globale est que la dérivée d'une des deux fonctions composantes φ_1 ou φ_2 ne s'annule pas sur I.

Exemple Considérons l'arc paramétré $\varphi : t \in]-\frac{1}{2}\pi, \frac{1}{2}\pi[\mapsto (\tan(t),\ \cos(t))$ dont l'image est une courbe paramétrée connue sous le nom de *quartique de Külp*[84]. Tous les points de la quartique de Külp sont des points réguliers car pour tout $t \in]-\frac{1}{2}\pi, \frac{1}{2}\pi[$, on a

$$\varphi'(t) = (1 + \tan^2(t),\ -\sin(t)) \quad \text{et} \quad \varphi_1'(t) > 0.$$

On en déduit que la quartique de Külp admet une représentation cartésienne de la forme $y = f(x)$ au voisinage de chacun de ses points. Comme φ_1' est strictement positive sur $]-\frac{1}{2}\pi, \frac{1}{2}\pi[$, on peut affirmer qu'il existe une représentation cartésienne globale de la quartique de Külp.

Compte tenu de la relation $1 + \tan^2(t) = 1/\cos^2(t)$ et du fait que pour tout $t \in]-\frac{1}{2}\pi, \frac{1}{2}\pi[$ on a $\cos(t) > 0$, on en déduit que

$$\forall t \in]-\tfrac{1}{2}\pi, \tfrac{1}{2}\pi[\qquad \varphi_2(t) = \cos(t) = \big(1 + \tan^2(t)\big)^{-1/2} = \big(1 + \varphi_1^2(t)\big)^{-1/2}.$$

[84] Edmund KÜLP (1800-1862), mathématicien allemand.

La quartique de Külp admet donc une représentation cartésienne globale : elle est le graphe de l'application $f : x \in \mathbb{R} \mapsto 1/\sqrt{1+x^2}$.

11.5.2 Application du théorème d'inversion locale à l'étude des surfaces paramétrées

On peut avoir une démarche analogue à celle qui vient d'être développée pour les courbes paramétrées avec les surfaces paramétrées. Le théorème d'inversion locale permet d'établir le résultat suivant qui joue un rôle analogue à celui de la proposition 11.14 pour les surfaces paramétrées de $\mathbb{R}^3$.

PROPOSITION 11.15 *Si M_0 est un point régulier d'une surface paramétrée S de $\mathbb{R}^3$, alors la surface paramétrée S admet une représentation cartésienne au voisinage du point M_0.*

Démonstration Désignons par $\varphi \in \mathcal{C}^1(U, \mathbb{R}^3)$, où U désigne un ouvert connexe de $\mathbb{R}^2$, un paramétrage de la surface paramétrée S et considérons un point M_0 de S tel que $M_0 = \varphi(u_0, v_0)$ avec $(u_0, v_0) \in U$, qui est régulier pour le paramétrage φ. Puisque le point M_0 est régulier, la matrice jacobienne de φ en (u_0, v_0)

$$\mathrm{J}_\varphi(u_0, v_0) = \begin{pmatrix} \partial_u\varphi_1(u_0, v_0) & \partial_v\varphi_1(u_0, v_0) \\ \partial_u\varphi_2(u_0, v_0) & \partial_v\varphi_2(u_0, v_0) \\ \partial_u\varphi_3(u_0, v_0) & \partial_v\varphi_3(u_0, v_0) \end{pmatrix}$$

est de rang 2. Cela signifie que l'on peut en extraire de $\mathrm{J}_\varphi(u_0, v_0)$ une matrice carrée d'ordre 2 dont le déterminant est non nul. Supposons dans un premier temps que cette matrice soit

$$\mathrm{A} = \begin{pmatrix} \partial_u\varphi_1(u_0, v_0) & \partial_v\varphi_1(u_0, v_0) \\ \partial_u\varphi_2(u_0, v_0) & \partial_v\varphi_2(u_0, v_0) \end{pmatrix},$$

et considérons l'application $\psi : (u, v) \in U \mapsto (\varphi_1(u, v), \varphi_2(u, v)) \in \mathbb{R}^2$. La matrice jacobienne de ψ en (u_0, v_0) est la matrice A et par hypothèse le jacobien de ψ est non nul. D'après le théorème d'inversion locale[85], il existe un voisinage $\widetilde{U}$ de (u_0, v_0) dans U telle que l'application ψ soit un difféomorphisme de $\widetilde{U}$ sur $\widetilde{V} = \psi(\widetilde{U})$ et pour tout $(u, v) \in \widetilde{U}$ on a

$$(u, v) = \psi^{-1}(\psi(u, v)) = \psi^{-1}(\varphi_1(u, v), \varphi_2(u, v)).$$

On en déduit qu'au voisinage de $M_0 = \varphi(u_0, v_0)$, les points $M = \varphi(u, v)$ de la surface paramétrée S de coordonnées $(x, y, z) = (\varphi_1(u, v), \varphi_2(u, v), \varphi_3(u, v))$ vérifient $\varphi_3(u, v) = \varphi_3(\psi^{-1}(\varphi_1(u, v), \varphi_2(u, v)))$. Ils appartiennent donc à la surface cartésienne d'équation $z = \varphi_3(\psi^{-1}(x, y))$.

[85] Voir le théorème 10.4 p. 665.

Si ce ne sont pas les deux premières lignes de la matrice jacobienne $\mathrm{J}_\varphi(u_0, v_0)$ qui forment une matrice de déterminant non nul, alors ce sont soit les deux dernières lignes, soit la première et la dernière ligne. Dans ces deux cas, on peut faire le même raisonnement que celui décrit précédemment, en intervertissant les rôles de φ_1, φ_2 et φ_3. □

Remarques

1. Dans la démonstration de la proposition 11.15, nous avons exploité le fait que $M_0 = \varphi(t_0)$ est régulier si et seulement si le rang de la matrice jacobienne de φ en (u_0, v_0) en égal à 2. Cette équivalence permet de reformuler la proposition 11.15 sous la forme suivante : si $M_0 = \varphi(u_0, v_0)$ est un point d'une surface paramétrée dont un paramétrage est φ et si $\mathrm{rg}(\mathrm{J}_\varphi(u_0, v_0)) = 2$, alors la surface S admet une représentation cartésienne au voisinage du point M_0.

2. Tout comme la proposition 11.14, la proposition 11.15 assure l'existence d'une représentation cartésienne pour une surface paramétrée mais la détermination explicite de l'équation cartésienne peut être difficile. On notera là encore que la condition est une condition suffisante mais non nécessaire et qu'en général la représentation cartésienne de la surface paramétrée n'est que « locale » ; dans la plupart des cas il n'existe pas de représentation cartésienne « globale » valable pour toute la courbe paramétrée. ◇

Exemple Considérons la surface paramétrée S dont un paramétrage est

$$\varphi : (u, v) \in \mathbb{R}^2 \longmapsto \big(3u\cos(v),\ 2v^2u,\ \sin(v)\big)$$

et intéressons-nous à l'existence d'une représentation cartésienne pour S au voisinage d'un point de la droite $D = \{(x, y, z) \in \mathbb{R}^3 \ ;\ y = 0 \text{ et } z = 0\}$. Remarquons tout d'abord que la droite D est incluse dans la surface paramétrée S car tout point de D est de la forme $(x, 0, 0)$ avec $x \in \mathbb{R}$ et $(x, 0, 0) = \varphi(\frac{1}{3}x, 0, 0)$.

Pour tout $(u, v) \in \mathbb{R}^2$, la matrice jacobienne de l'application φ en (u, v) est

$$\mathrm{J}_\varphi(u, v) = \begin{pmatrix} 3\cos(v) & -3u\sin(v) \\ 2v^2 & 4uv \\ 0 & \cos(v) \end{pmatrix}.$$

On a en particulier,

$$\mathrm{J}_\varphi\left(\tfrac{1}{3}x, 0\right) = \begin{pmatrix} 3 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{et} \quad \begin{vmatrix} 3 & 0 \\ 0 & 1 \end{vmatrix} = 3 \neq 0$$

donc le rang de la matrice jacobienne $\mathrm{J}_\varphi\left(\frac{x}{3}, 0\right)$ est égal à 2 puisque l'on peut extraire de la matrice $\mathrm{J}_\varphi\left(\frac{x}{3}, 0\right)$ une matrice d'ordre 2 à déterminant non nul. On en déduit que tous les points de la droite D sont réguliers et par conséquent qu'il est possible, au voisinage de tout point de la droite D, de décrire la surface paramétrée S à l'aide d'une représentation cartésienne.

On peut être un peu plus précis : parmi les trois déterminants d'ordre 2 que l'on peut extraire du jacobien $J_\varphi(\frac{1}{3}x, 0)$, un seul est non nul : c'est celui où l'on conserve la première et la troisième ligne. La démonstration de la proposition 11.15 indique que l'on peut par conséquent exprimer les paramètres u et v en fonction des coordonnées x et z. La surface paramétrée S admet donc une représentation cartésienne d'équation $y = f(x, z)$.

On peut à partir de ces informations, tenter d'établir la représentation cartésienne de la surface. Cherchons à exprimer les paramètres u et v en fonction des coordonnées cartésiennes x et z. On a $v = \arcsin(z) \in [-\frac{1}{2}\pi, \frac{1}{2}\pi]$ et par conséquent $\cos(v) = \sqrt{1 - \sin^2(v)} = \sqrt{1 - z^2}$. On en déduit que $x = 3u\cos(v) = 3u\sqrt{1 - z^2}$ d'où

$$u = \frac{x}{3\sqrt{1 - z^2}} \quad \text{et} \quad y = 2v^2u = \frac{2x \arcsin^2(z)}{3\sqrt{1 - z^2}}.$$

Au voisinage des points de la droite D, la surface paramétrée S admet une représentation cartésienne d'équation $y = \frac{2}{3}x \arcsin^2(z)/\sqrt{1 - z^2}$.

11.6 Théorème des fonctions implicites

Nous avons vu comment on pouvait à l'aide du théorème d'inversion locale établir l'existence (localement) d'une représentation cartésienne pour une courbe paramétrée de $\mathbb{R}^2$ ou une surface paramétrée de $\mathbb{R}^3$. Nous allons dans cette partie énoncer un résultat, le théorème des fonctions implicites, qui permet d'établir l'existence d'une représentation cartésienne pour une courbe ou pour une surface définie sous forme implicite.

11.6.1 Théorème des fonctions implicites pour des fonctions de $\mathbb{R}^2$ dans $\mathbb{R}$

Considérons une courbe implicite C du plan d'équation $F(x, y) = 0$, où F est une application continue d'un ouvert U de $\mathbb{R}^2$ à valeurs dans $\mathbb{R}$, et intéressons-nous au problème de déterminer si cette courbe implicite possède une représentation cartésienne.

Le cas du cercle est assez représentatif des différentes situations que l'on peut rencontrer. Une équation implicite du cercle de centre $(0, 0)$ et de rayon 1 dans le plan est $x^2 + y^2 = 1$. Si ce cercle admettait une représentation cartésienne de la forme $y = f(x)$ où f désigne une fonction de $\mathbb{R}$ dans $\mathbb{R}$, il serait la représentation graphique de la fonction f et un point $\alpha \in]-1, 1[$ aurait deux images $\sqrt{1 - \alpha^2}$ et $-\sqrt{1 - \alpha^2}$ par f, voir la figure 20. Cela est en contradiction avec la définition d'une fonction[86]. Géométriquement, la représentation graphique d'une fonction admet au plus un point d'intersection avec une droite parallèle à l'axe des ordonnées ; ce n'est pas le cas ici.

[86] Voir la définition 2.12 p. 35 du *Cours de première année*.

Toutefois, le demi-cercle « supérieur » représenté par le sous-ensemble $\{(x,y) \in \mathbb{R}^2 \ ; \ x^2+y^2=1 \text{ et } y \geqslant 0\}$ admet une représentation cartésienne : il est le graphe de l'application $f : x \in [-1,1] \mapsto \sqrt{1-x^2}$. On a une situation analogue pour le demi-cercle « inférieur » $\{(x,y) \in \mathbb{R}^2 \ ; \ x^2+y^2=1 \text{ et } y \leqslant 0\}$ qui est le graphe de l'application $f : x \in [-1,1] \mapsto -\sqrt{1-x^2}$. On notera cependant que pour un arc de cercle contenant le point $(1,0)$ ou le point $(-1,0)$, il n'existe pas de représentation cartésienne de la forme $y = f(x)$. Ici encore la condition d'existence d'une représentation cartésienne sera une condition locale.

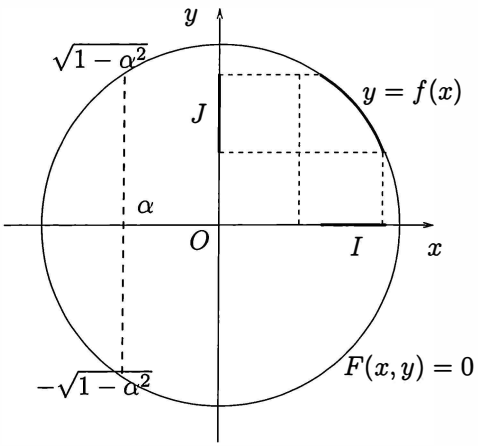


Fig. 20 Illustration du théorème des fonctions implicites sur un cercle.

Le théorème suivant indique à quelles conditions une courbe implicite admet au voisinage d'un de ses points une représentation cartésienne.

THÉORÈME 11.2 (Théorème des fonctions implicites)
Soient U un ouvert de $\mathbb{R}^2$, F une application de U dans $\mathbb{R}$ de classe C^k $(k \in \mathbb{N}^* \cup \{\infty\})$ et (a,b) un élément de U.
Si $F(a,b) = 0$ et $\partial_y F(a,b) \neq 0$ alors il existe :
- un intervalle ouvert I de $\mathbb{R}$ contenant a,
- un intervalle ouvert J de $\mathbb{R}$ contenant b tel que $I \times J \subset U$,
- et une unique application $f \in C^k(I,J)$ tels que pour $(x,y) \in I \times J$

$$F(x,y) = 0 \iff y = f(x).$$

En particulier, on a $f(a) = b$ et $F(x,f(x)) = 0$ pour tout $x \in I$.

Démonstration $\trianglerighteq$ Considérons l'application

$$G : (x,y) \in U \longmapsto (x, F(x,y)) \in \mathbb{R}^2.$$

D'après les hypothèses du théorème, cette application est de classe $\mathcal{C}^k$. Sa matrice jacobienne en $(x,y) \in U$ est :

$$\mathrm{J}_G(x,y) = \begin{pmatrix} 1 & \partial_x F(x,y) \\ 0 & \partial_y F(x,y) \end{pmatrix}.$$

Le jacobien de l'application G en (a,b) vaut donc $\partial_y F(a,b)$ et, par hypothèse, il est non nul. Le théorème d'inversion locale[87] indique qu'il existe un voisinage V de (a,b) tel que l'application G soit un $\mathcal{C}^k$-difféomorphisme de V sur $G(V)$. Comme V est un voisinage de (a,b), il existe $r \in \mathbb{R}_+^*$ tel que la boule ouverte de centre (a,b) et de rayon r pour la norme infinie $B((a,b),r) =]a-r,a+r[\times]b-r,b+r[$ soit incluse dans V. Notons $I =]a-r,a+r[$ et $J =]b-r,b+r[$; l'application G est un $\mathcal{C}^k$-difféomorphisme de $I \times J$ sur $G(I \times J) = I \times F(I \times J)$.

D'après la définition de G, pour tout $(x,y) \in I \times J$ on a

$$F(x,y) = 0 \iff (x, F(x,y)) = (x,0) \iff G(x,y) = (x,0).$$

De plus, comme G est un $\mathcal{C}^k$-difféomorphisme de $I \times J$ sur $G(I \times J)$, on a

$$\begin{aligned} G(x,y) = (x,0) &\iff (x,y) = G^{-1}(x,0) \\ &\iff (x,y) = (G_1^{-1}(x,0), G_2^{-1}(x,0)) \end{aligned}$$

où G^{-1} désigne la bijection réciproque de G, et G_1^{-1} et G_2^{-1} les deux applications composantes de l'application G^{-1}.
Remarquons que d'après la définition de l'application G, on a pour $x \in I$,

$$G(G^{-1}(x,0)) = (G_1^{-1}(x,0), F(G^{-1}(x,0))).$$

D'autre part, comme G et G^{-1} sont bijection réciproque l'une de l'autre, on a aussi $G(G^{-1}(x,0)) = (x,0)$. On a donc $G_1^{-1}(x,0) = x$ et $F(G^{-1}(x,0)) = 0$.
On déduit de tout ce qui précède que pour tout $(x,y) \in I \times J$ on a

$$F(x,y) = 0 \iff (x,y) = (x, G_2^{-1}(x,0)) \iff y = G_2^{-1}(x,0).$$

Définissons l'application $f : x \in I \mapsto G_2^{-1}(x,0) \in J$. D'après l'équivalence précédente on a $F(x,y) = 0$ si et seulement si $y = f(x)$. La première affirmation du théorème est donc établie.

$\trianglerighteq$ Comme $G(a,b) = (a,0)$, on a

$$(a,b) = G^{-1}(G(a,b)) = G^{-1}(a,0) = (G_1^{-1}(a,0), G_2^{-1}(a,0)).$$

On en déduit que $b = G_2^{-1}(a,0)$ c'est-à-dire que $b = f(a)$.
Par ailleurs, on a vu que pour tout $x \in I$ on a $F(G^{-1}(x,0)) = 0$. Or,

$$F(G^{-1}(x,0)) = F(G_1^{-1}(x,0), G_2^{-1}(x,0)) = F(x, G_2^{-1}(x,0)) = F(x, f(x)).$$

On en déduit que $F(x,f(x)) = 0$ pour tout $x \in I$. □

[87] Voir le théorème 10.4 p. 665.

Le théorème des fonctions implicites ne donne pas explicitement la représentation cartésienne de la surface implicite ; il en assure l'existence. Néanmoins, à partir de la relation $F(x, f(x)) = 0$ qui est valable pour tout $x \in I$, on peut obtenir de nombreuses informations sur l'application f. Ainsi, la dérivée de la fonction f en a peut s'obtenir en fonction des dérivées partielles de F en (a, b). En effet, en dérivant l'égalité $F(x, f(x)) = 0$ par rapport à x, on obtient[88] :

$$\partial_1 F(x, f(x)) + f'(x)\, \partial_2 F(x, f(x)) = 0,$$

ce qui fournit pour $x = a$ l'égalité

$$\boxed{f'(a) = -\frac{\partial_x F(a,b)}{\partial_y F(a,b)}.} \tag{17}$$

On peut, de la même manière, en dérivant successivement par rapport à x la relation $F(x, f(x)) = 0$, obtenir les dérivées d'ordre supérieur de l'application f en a en fonction des dérivées partielles de l'application F en (a, b).

Si la condition $\partial_y F(a, b) \neq 0$ n'est pas réalisée, le théorème 11.2 ne s'applique pas. Toutefois, si $\partial_y F(a, b) = 0$ et $\partial_x F(a, b) \neq 0$ alors il est possible d'échanger le rôle des variables x et y dans la preuve du théorème 11.2 et ainsi d'expliciter x en fonction de y. Plus précisément, on a le résultat suivant.

THÉORÈME 11.3 (Théorème des fonctions implicites - 2e forme)
Soient U un ouvert de $\mathbb{R}^2$, F une application de classe C^k de U dans $\mathbb{R}$ ($k \in \mathbb{N}^* \cup \{\infty\}$) et (a, b) un élément de U.
Si $F(a, b) = 0$ et $\partial_x F(a, b) \neq 0$ alors il existe :
- un intervalle ouvert I de $\mathbb{R}$ contenant b,
- un intervalle ouvert J de $\mathbb{R}$ contenant a tel que $J \times I \subset U$,
- et une unique application $f \in C^k(I, J)$ tels que pour tout $(x, y) \in J \times I$

$$F(x, y) = 0 \iff x = f(y).$$

En particulier, on a $f(b) = a$ et $F(f(y), y) = 0$ pour tout $y \in I$.

Exemple L'application $F : (x, y) \in \mathbb{R}^2 \mapsto x^4 + x^2y^2 + 4x^2y - 2x^2 + 3y^2 - 4y + 1$ est de classe C^∞ sur $\mathbb{R}^2$ et

$$\partial_y F(x, y) = (2x^2 + 3)y + 4(x^2 - 1).$$

On a $F(0, 1) = 0$ et $\partial_y F(0, 1) = -1$, donc, d'après le théorème 11.2, il existe un intervalle ouvert I contenant 0, un intervalle ouvert J contenant 1 et une unique application $f \in C^\infty(I, J)$ tels que pour $(x, y) \in I \times J$

$$F(x, y) = 0 \iff y = f(x).$$

[88] On note $\partial_1 F$ et $\partial_2 F$ les deux dérivées partielles premières de F pour éviter toute confusion relativement au sens à donner à x.

Cherchons à déterminer l'application f. On a

$$F(x,y)=0 \iff (x^2+3)y^2+4(x^2-1)y+(x^2-1)^2=0.$$

On en déduit que si $(x,y)\in I\times J$ vérifie $F(x,y)=0$ alors y est racine du polynôme du second degré $P=aX^2+bX+c$ où $a=x^2+3$, $b=4(x^2-1)$ et $c=(x^2-1)^2$. Ce polynôme a pour discriminant

$$\Delta=b^2-4ac=4(1-x^2)^3$$

et cette quantité est toujours strictement positive pour $x\in]-1,1[$. On en déduit que pour tout $x\in]-1,1[$, le polynôme P admet deux racines réelles distinctes qui sont

$$\frac{(1-x^2)(2+\sqrt{1-x^2})}{x^2+3} \quad \text{et} \quad \frac{(1-x^2)(2-\sqrt{1-x^2})}{x^2+3}.$$

La première racine prend la valeur 1 en 0 alors que la seconde prend la valeur $1/3$ en 0. C'est donc la première qui représente l'expression de $f(x)$. On en déduit qu'au voisinage du point $(0,1)$,

$$F(x,y)=0 \iff y=\frac{(1-x^2)(2+\sqrt{1-x^2})}{x^2+3}.$$

Notons que puisque $\partial_x F(0,1)=0$, le théorème des fonctions implicites dans sa seconde version ne s'applique pas. Enfin, on peut également noter que $F(1,0)=F(-1,0)=0$ mais que $\partial_x F(1,0)=\partial_y F(1,0)=0$ et que $\partial_x F(-1,0)=\partial_y F(-1,0)=0$, donc le théorème des fonctions implicites ne s'applique dans aucune de ses deux versions en $(1,0)$ et $(-1,0)$.

L'ensemble des points du plan de coordonnées (x,y) tel que $F(x,y)=0$ est représenté sur la figure 21. La courbe implicite d'équation $F(x,y)=0$ est connue sous le nom de *courbe du bicorne*[89].

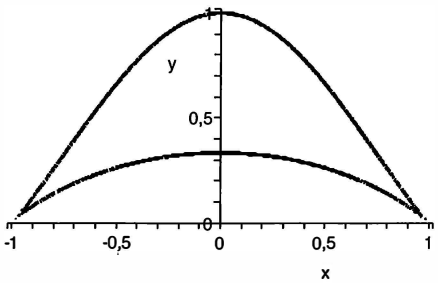


Fig. 21 La courbe du bicorne.

[89] En référence à sa forme qui rappelle celle du chapeau bicorne, couvre-chef masculin du 18e siècle porté encore de nos jours par les membres de l'Académie Française et les élèves de l'École Polytechnique.

11.6.2 Étude des courbes implicites du plan

Le théorème des fonctions implicites permet d'obtenir une condition pour qu'une courbe implicite du plan admette une représentation cartésienne au voisinage d'un de ses points. Avant de l'énoncer, nous devons définir la notion de point régulier pour une courbe implicite.

DÉFINITION 11.22 (Point régulier d'une courbe implicite)
Soit F une application de classe $\mathcal{C}^1$ définie sur un ouvert U de $\mathbb{R}^2$ et à valeurs dans $\mathbb{R}$. Un point (x_0, y_0) de la courbe implicite du plan d'équation $F(x,y) = 0$ est dit régulier *si $\nabla F(x_0, y_0) \neq \mathbf{0}$. Dans le cas contraire, on dit que le point est* singulier.

Exemples

1. Le *Folium de Descartes*[90] est la courbe implicite du plan d'équation $x^3 + y^3 - 3cxy = 0$ où $c \in \mathbb{R}^*$. L'application $F : (x,y) \in \mathbb{R}^2 \mapsto x^3 + y^3 - 3cxy$ est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}^2$ et pour tout $(x,y) \in \mathbb{R}^2$ on a

$$\nabla F(x,y) = \begin{pmatrix} \partial_x F(x,y) \\ \partial_y F(x,y) \end{pmatrix} = \begin{pmatrix} 3(x^2 - c) \\ 3(y^2 - c) \end{pmatrix}.$$

Si $c < 0$, le gradient de F ne s'annule jamais et si $c > 0$ le gradient de F s'annule pour $(x,y) \in \{-\sqrt{c}, \sqrt{c}\}^2$ mais ces points n'appartiennent pas à la courbe implicite d'équation $F(x,y) = 0$. On en déduit que tous les points du Folium de Descartes sont des points réguliers.

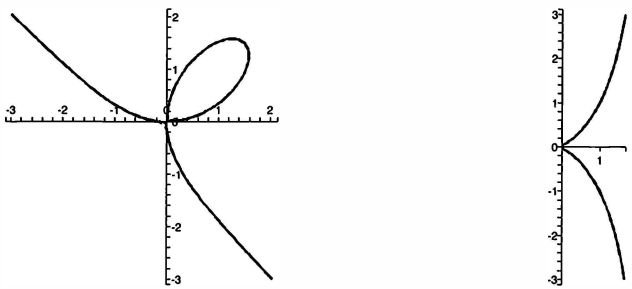


Fig. 22 Folium de Descartes à gauche et Cissoïde de Dioclès à droite.

2. La *Cissoïde de Dioclès*[91] est la courbe implicite du plan d'équation $x^3 + xy^2 - cy = 0$ où $c \in \mathbb{R}^*$. L'application $F : (x,y) \in \mathbb{R}^2 \mapsto x^3 + xy^2 - cy$ est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}^2$ et pour tout $(x,y) \in \mathbb{R}^2$ on a

---

[90] René DESCARTES (1596-1650), philosophe, mathématicien et physicien français. Le nom attribué à cette courbe provient du latin *folium* qui signifie *feuille*.

[91] DIOCLÈS, mathématicien grec du 2e siècle avant J.-C. Le nom attribué à cette courbe provient du grec *kissos* qui signifie *lierre* car la courbe rappelle la forme des nervures d'une feuille de lierre.

$$\nabla F(x,y) = \begin{pmatrix} \partial_x F(x,y) \\ \partial_y F(x,y) \end{pmatrix} = \begin{pmatrix} 3x^2 + y^2 \\ 3(x-c)y \end{pmatrix}.$$

La première composante de ∇F ne s'annule que pour $x = y = 0$. Pour ces valeurs, la seconde composante est nulle et comme $F(0,0) = 0$, le point de coordonnées $(0,0)$ appartient à la courbe. On en déduit que l'origine est l'unique point singulier de la Cissoïde de Dioclès.

Remarques

1. Nous disposons de deux notions de points réguliers pour une courbe. L'une pour les courbes paramétrées (voir la définition 11.11 p. 745), l'autre pour les courbes implicites. On prendra garde au fait qu'un point d'une courbe peut être régulier pour l'une de ces deux définitions et singulier pour l'autre. Par exemple, dans le plan la droite d'équation $y = x$ peut être vue comme une courbe implicite d'équation $F(x,y) = 0$ où $F : (x,y) \in \mathbb{R}^2 \mapsto x - y$ et tous les points de cette courbe implicite sont réguliers. Cette même droite peut être décrite comme une courbe paramétrée dont un paramétrage est $\varphi : t \in \mathbb{R} \mapsto (t^3, t^3)$ et l'origine est un point singulier pour l'arc paramétré φ.

2. Dans le plan, une courbe cartésienne d'équation $y = f(x)$ peut être décrite comme une courbe paramétrée dont un paramétrage est $\varphi : t \mapsto (t, f(t))$. Tous les points de cette courbe paramétrée sont réguliers (voir la remarque p. 799). Cette même courbe cartésienne peut être également décrite comme une courbe implicite d'équation $F(x,y) = 0$ où $F : (x,y) \mapsto y - f(x)$. Le gradient de F qui est l'application $(x,y) \mapsto (-f'(x), 1)$ ne s'annulant jamais, tous les points de cette courbe implicite sont réguliers. Ainsi, tous les points d'une courbe cartésienne sont réguliers, aussi bien au sens de points réguliers d'une courbe paramétrée qu'au sens de points réguliers d'une courbe implicite. ◇

Exercice 11.10 Montrer que le point de coordonnées $(2,1)$ est un point régulier de la courbe implicite du plan d'équation $x^2 + y^4 - 3xy + x = 1$.

Le résultat suivant indique qu'au voisinage d'un point régulier, une courbe implicite du plan admet une représentation cartésienne.

PROPOSITION 11.16 *Soit F une application de classe $\mathcal{C}^1$ définie sur un ouvert U de $\mathbb{R}^2$ et à valeurs dans $\mathbb{R}$. Au voisinage de tout point régulier, la courbe implicite de $\mathbb{R}^2$ d'équation $F(x,y) = 0$ admet une représentation cartésienne de la forme $y = f(x)$ ou de la forme $x = f(y)$ où f désigne une fonction de $\mathbb{R}$ dans $\mathbb{R}$.*

Démonstration Soit $M_0 = (a, b)$ un point régulier de la courbe implicite C d'équation $F(x, y) = 0$. On a $F(a, b) = 0$ et l'une des dérivées partielles de F en (a, b) est non nulle. Supposons dans un premier temps qu'il s'agisse de la deuxième dérivée partielle et appliquons le théorème des fonctions implicites (théorème 11.2) : il existe un intervalle ouvert I de $\mathbb{R}$ contenant a, un intervalle ouvert J de $\mathbb{R}$ contenant b avec $I \times J \subset U$ et une unique application f de I dans J de classe $\mathcal{C}^k$ tels que

$$\forall (x, y) \in I \times J \qquad \big(F(x, y) = 0 \iff y = f(x)\big).$$

Cette équivalence signifie qu'au voisinage de M_0, la courbe implicite d'équation $F(x, y) = 0$ coïncide avec la courbe cartésienne d'équation $y = f(x)$.

Si $\partial_y F(a, b) = 0$ alors, comme le point M_0 est régulier, on a nécessairement $\partial_x F(a, b) \neq 0$ et on peut appliquer le théorème 11.3. Un raisonnement similaire au précédent, où le rôle des variables x et y est échangé, permet d'établir qu'au voisinage du point M_0, la courbe implicite d'équation $F(x, y) = 0$ coïncide avec une courbe cartésienne dont l'équation est de la forme $x = f(y)$ où f désigne une fonction de $\mathbb{R}$ dans $\mathbb{R}$. □

Le théorème des fonctions implicites permet également de préciser l'allure d'une courbe implicite au voisinage d'un point régulier. Nous allons illustrer cela sur un exemple.

Exemple Considérons la courbe implicite C du plan d'équation $F(x, y) = 0$ où $F : (x, y) \in \mathbb{R}^2 \mapsto x^2 + y^4 - 3xy + x - 1$. Nous avons établi au cours de l'exercice 11.10 que le point M_0 de coordonnées $(2, 1)$ est un point régulier de la courbe C et que pour tout $(x, y) \in \mathbb{R}^2$ on a $\partial_x F(x, y) = 2x - 3y + 1$ et $\partial_y F(x, y) = 4y^3 - 3x$. Étudions l'allure de la courbe C au voisinage de M_0.

L'application F est de classe $\mathcal{C}^\infty$ et $\partial_y F(2, 1) = -2 \neq 0$ donc, d'après le théorème des fonctions implicites, il existe un intervalle ouvert I contenant 2, un intervalle ouvert J contenant 1 et une unique application f de I dans J de classe $\mathcal{C}^\infty$ tels que au voisinage du point M_0, la courbe implicite C coïncide avec la courbe cartésienne d'équation $y = f(x)$. Pour préciser l'allure de la courbe C au voisinage du point M_0, déterminons un développement limité de l'application f au voisinage de $x = 2$.

D'après le théorème 11.2, pour tout $x \in I$ on a $F(x, f(x)) = 0$. En dérivant les deux membres de cette égalité par rapport à x, on obtient :

$$\partial_x F(x, f(x)) + f'(x)\, \partial_y F(x, f(x)) = 0.$$

On en déduit que

$$f'(2) = -\frac{\partial_x F(2, 1)}{\partial_y F(2, 1)} = 1.$$

En dérivant une nouvelle fois par rapport à x, on obtient :

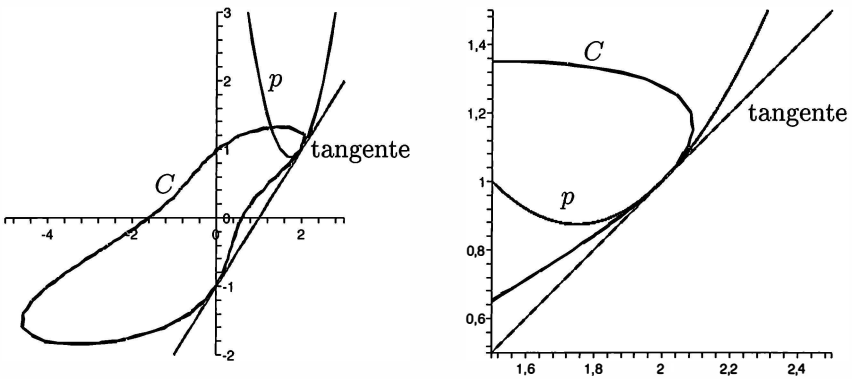
$$\partial^2_{xx} F(x, f(x)) + 2f'(x)\, \partial^2_{xy} F(x, f(x)) + f''(x)\, \partial_y F(x, f(x)) + f'(x)^2\, \partial^2_{yy} F(x, f(x)) = 0.$$

En prenant $x = 2$ dans cette égalité, on en déduit que $f''(2) = 4$. L'application f étant de classe $\mathcal{C}^\infty$ sur I, elle admet un développement limité à l'ordre 2 donné par la formule de Taylor-Young[92] :

$$\begin{aligned} f(x) &= f(2) + (x-2)\,f'(2) + \tfrac{1}{2}(x-2)^2\,f''(2) + o_2((x-2)^2) \\ &= 1 + (x-2) + 2(x-2)^2 + o_2((x-2)^2). \end{aligned}$$

On en déduit[93] que la tangente à la courbe C au point $(2,1)$ a pour équation $y = x - 1$ et que la courbe est au-dessus de sa tangente. On en déduit aussi qu'au voisinage du point $(2,1)$, la courbe est « proche » de la représentation graphique de la fonction polynomiale $p : x \mapsto 1 + (x-2) + 2(x-2)^2$. On peut visualiser la courbe C, sa tangente et le graphe de l'application p avec MAPLE.

```
> fig1:=implicitplot(x^2+y^4-3*x*y+x-1, x=-5..3, y=-2..3):
> fig2:=plot([1+(x-2)+2*(x-2)^2,1+(x-2)], x=-5..3, y=-2..3):
> display({fig1,fig2});
> display({fig1,fig2},view=[1.5..2.5,0.5..1.5]);
```



Plus généralement, la proposition 11.16 et le théorème des fonctions implicites permettent d'exploiter les résultats introduits pour l'étude des courbes cartésiennes dans le plan afin d'étudier localement les courbes implicites.

Exercice 11.11 On s'intéresse dans cet exercice à la courbe implicite C d'équation $F(x,y) = 0$ où $F : (x,y) \in \mathbb{R}^2 \mapsto \sin(x) - \sin(y) + 2x - y$.
1 - Montrer que l'on peut appliquer le théorème des fonctions implicites à F en $(0,0)$ et en déduire qu'il existe une fonction f telle qu'au voisinage de l'origine la courbe C est le graphe de f.
2 - Calculer $f(0)$, $f'(0)$ et $f''(0)$ et déterminer le développement limité à l'ordre 2 de f au point $x = 0$.
3 - Dessiner l'allure de la courbe C au voisinage de l'origine.

[92] Voir le théorème 17.1 p. 778 du *Cours de première année*.
[93] Voir le *Cours de première année* p. 755.

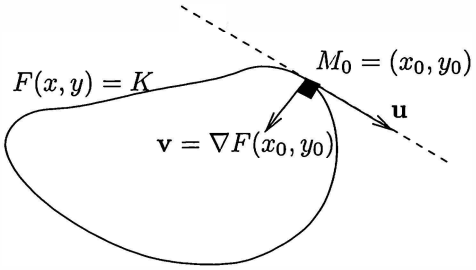


Fig. 23 Le vecteur gradient est orthogonal au vecteur tangent.

Vecteur gradient et courbes de niveau

Nous avons défini p. 728 les courbes de niveau d'une surface cartésienne S d'équation $z = F(x, y)$ où F désigne une application de classe $\mathcal{C}^1$ définie sur un ouvert U de $\mathbb{R}^2$ et à valeurs dans $\mathbb{R}$, comme étant les courbes implicites du plan d'équation $F(x, y) = K$ où K désigne un réel fixé. La proposition suivante, qui est une conséquence du théorème des fonctions implicites, exprime le fait que le vecteur gradient $\nabla F(x_0, y_0) \in \mathbb{R}^2$ en un point (x_0, y_0) d'une courbe de niveau de la surface S est orthogonal au vecteur tangent à la courbe en ce point, voir la figure 23.

PROPOSITION 11.17 *Soit F une application de classe $\mathcal{C}^1$ définie sur un ouvert U de $\mathbb{R}^2$ et à valeurs dans $\mathbb{R}$. Le vecteur gradient $\nabla F(x_0, y_0)$ en un point régulier $M_0 = (x_0, y_0)$ de la courbe implicite d'équation $F(x, y) = K$ (où $K \in \mathbb{R}$) est orthogonal au vecteur tangent à la courbe en ce point.*

Démonstration Notons C la courbe implicite d'équation $F(x, y) = K$ où $K \in \mathbb{R}$. D'après la proposition 11.16, le point $M_0 = (x_0, y_0)$ étant un point régulier, la courbe implicite C admet au voisinage de M_0 une représentation cartésienne de la forme $y = f(x)$ ou de la forme $x = f(y)$ où f est une fonction de $\mathbb{R}$ dans $\mathbb{R}$ de classe $\mathcal{C}^1$. Supposons que l'on ait $y = f(x)$ (le raisonnement est analogue dans l'autre cas, il suffit d'échanger les rôles de x et de y) ; d'après la relation (17) p. 812 on a

$$f'(x_0) = -\frac{\partial_x F(x_0, y_0)}{\partial_y F(x_0, y_0)}. \tag{18}$$

La tangente à la courbe cartésienne d'équation $y = f(x)$ en (x_0, y_0) est dirigée par le vecteur[(94)]

$$\mathbf{u} = \begin{pmatrix} 1 \\ f'(x_0) \end{pmatrix},$$

[(94)] Voir le *Cours de première année* p. 714.

et le gradient de F en (x_0, y_0) est

$$\mathbf{v} = \begin{pmatrix} \partial_x F(x_0, y_0) \\ \partial_y F(x_0, y_0) \end{pmatrix}.$$

Pour établir le résultat énoncé, il faut montrer que le produit scalaire $\mathbf{u} \cdot \mathbf{v}$ est nul. On a

$$\mathbf{u} \cdot \mathbf{v} = \partial_x F(x_0, y_0) + f'(x_0)\, \partial_y F(x_0, y_0)$$

et cette quantité est nulle d'après la relation (18). □

Remarques

1. Selon les valeurs du réel K, la courbe implicite d'équation $F(x, y) = K$ peut être l'ensemble vide. C'est par exemple le cas pour $F : (x, y) \in \mathbb{R}^2 \mapsto x^2 + y^2$ si $K < 0$.

2. La proposition 11.17 indique que « les gradients sont orthogonaux aux courbes de niveau d'une fonction ». Ainsi, en observant une carte topographique contenant des courbes de niveau, on peut en déduire que les lignes de plus grandes pentes (dirigées par les gradients) sont orthogonales à ces lignes de niveau. ◇

11.6.3 Théorème des fonctions implicites pour des fonctions de $\mathbb{R}^{n+q}$ dans $\mathbb{R}^q$

Le théorème des fonctions implicites énoncé p. 810 pour les fonctions F de $\mathbb{R}^2$ dans $\mathbb{R}$, qui indique à quelles conditions on peut exprimer les solutions de l'équation $F(x, y) = 0$ sous la forme $y = f(x)$, admet une généralisation aux fonctions de $\mathbb{R}^m$ dans $\mathbb{R}^q$ lorsque $m > q$. Cette généralisation du théorème des fonctions implicites donne les conditions permettant d'expliciter q variables $y_1, \ldots, y_q$ en fonction de n autres $x_1, \ldots, x_n$ à partir de q relations de la forme :

$$\begin{cases} F_1(x_1, \ldots, x_n, y_1, \ldots, y_q) = 0 \\ F_2(x_1, \ldots, x_n, y_1, \ldots, y_q) = 0 \\ \qquad\qquad \vdots \\ F_q(x_1, \ldots, x_n, y_1, \ldots, y_q) = 0 \end{cases} \tag{19}$$

Le théorème des fonctions implicites dans le cas d'une application linéaire F

Avant de considérer le cas général, commençons par nous intéresser au cas où le système (19) est un système linéaire, c'est-à-dire un système de la forme

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1,n+q}x_{n+q} = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2,n+q}x_{n+q} = 0 \\ \qquad\qquad\qquad\qquad\qquad\qquad \\ a_{q1}x_1 + a_{q2}x_2 + \cdots + a_{q,n+q}x_{n+q} = 0 \end{cases} \tag{20}$$

où a_{ij}, $(i, j) \in \{1, \ldots, q\} \times \{1, \ldots, n+q\}$ désignent $q(n+q)$ réels.

En désignant par F l'application *linéaire* de $\mathbb{R}^{n+q}$ dans $\mathbb{R}^q$ où $(n,q) \in \mathbb{N}^* \times \mathbb{N}^*$ dont les q applications composantes $F_1,\ldots,F_q$ sont les applications linéaires

$$F_i : \mathbf{x} \in \mathbb{R}^{n+q} \longmapsto a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{i,n+q}x_{n+q} \in \mathbb{R},$$

le système (20) est bien de la forme (19).

Ce système est un système *sous-abondant* c'est-à-dire qu'il possède plus d'inconnues $(n+q)$ que d'équations (q). D'après le théorème de Rouché-Fontené[95], on sait que si le rang de ce système linéaire est égal à q alors il admet une infinité de solutions et grâce à la méthode d'élimination de Gauss[96], on peut exprimer q des inconnues en fonction des n autres. Plus précisément, si on suppose[97] que ce sont les q dernières colonnes de la matrice $\mathrm{A} = (a_{ij})_{i\in\{1,\ldots,q\},j\in\{1,\ldots,n+q\}}$ qui ne sont pas liées alors on peut utiliser les inconnues $x_{n+1},\ldots,x_{n+q}$ comme « inconnues principales » et écrire le système (20) sous la forme :

$$\begin{cases} a_{1,n+1}x_{n+1} + a_{1,n+2}x_{n+2} + \cdots + a_{1,n+q}x_{n+q} = -a_{11}x_1 - a_{12}x_2 - \ldots + a_{1n}x_n \\ a_{2,n+1}x_{n+1} + a_{2,n+2}x_{n+2} + \cdots + a_{2,n+q}x_{n+q} = -a_{21}x_1 - a_{22}x_2 - \ldots + a_{2n}x_n \\ \\ a_{q,n+1}x_{n+1} + a_{q,n+2}x_{n+2} + \cdots + a_{q,n+q}x_{n+q} = -a_{q1}x_1 - a_{q2}x_2 - \ldots + a_{qn}x_n \end{cases}$$

Ce système est, sous les hypothèses mentionnées, un système de Cramer[98] et possède donc pour chaque valeur de $\overline{\mathbf{x}} = (x_1,\ldots,x_n) \in \mathbb{R}^n$, une unique solution $\overline{\mathbf{y}} = (x_{n+1},\ldots,x_{n+q}) \in \mathbb{R}^q$ que la méthode d'élimination de Gauss permet d'exprimer en fonction des valeurs de $(x_1,\ldots,x_n) \in \mathbb{R}^n$. En d'autres termes, on vient d'établir que si F est une application linéaire de $\mathbb{R}^n \times \mathbb{R}^q = \mathbb{R}^{n+q}$ dans $\mathbb{R}^q$ telle que $F(x_1,\ldots,x_{n+q}) = 0$ et

$$\det \begin{pmatrix} a_{1,n+1} & a_{1,n+2} & \cdots & a_{1,n+q} \\ a_{2,n+1} & a_{2,n+2} & \cdots & a_{2,n+q} \\ \vdots & \vdots & & \vdots \\ a_{q,n+1} & a_{q,n+2} & \cdots & a_{q,n+q} \end{pmatrix} \neq 0$$

alors il existe une unique application linéaire f de $\mathbb{R}^n$ dans $\mathbb{R}^q$ tel que

$$\forall(\overline{\mathbf{x}},\overline{\mathbf{y}}) \in \mathbb{R}^n \times \mathbb{R}^q \qquad \big(F(\overline{\mathbf{x}},\overline{\mathbf{y}}) = 0 \iff \overline{\mathbf{y}} = f(\overline{\mathbf{x}})\big).$$

Exemple Intéressons-nous au système linéaire

$$\begin{cases} 2x_1 + 3x_2 - x_3 = 0 \\ x_1 + x_2 + x_3 = 0 \end{cases} \tag{21}$$

---

[95] Voir le *Cours de première année* p. 499.

[96] Voir le *Cours de première année* p. 500.

[97] Si cette condition n'est pas vérifiée, une renumérotation des inconnues permet de se ramener à cette situation.

[98] Voir la définition 11.13 p.493 du *Cours de première année.*

qui s'écrit $F(x_1, x_2, x_3) = 0$ où

$$F : (x_1, x_2, x_3) \in \mathbb{R}^3 \longmapsto (2x_1 + 3x_2 - x_3, x_1 + x_2 + x_3) \in \mathbb{R}^2$$

et sous la forme matricielle $\mathrm{A}\mathbf{x} = \mathbf{0}$ où

$$\mathrm{A} = \begin{pmatrix} 2 & 3 & -1 \\ 1 & 1 & 1 \end{pmatrix} \quad \text{et} \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

La matrice A est de rang 2 et on peut récrire le système (21) sous la forme

$$\begin{cases} 3x_2 - x_3 = -2x_1 \\ x_2 + x_3 = -x_1 \end{cases}.$$

Pour chaque réel x_1 fixé, ce système est un système de Cramer dont la solution est $x_2 = -\frac{3}{4}x_1$, $x_3 = -\frac{1}{3}x_1$. Ainsi, pour tout $(x_1, x_2, x_3) \in \mathbb{R}^3$,

$$\big(\mathrm{A}\mathbf{x} = \mathbf{0} \iff F(x_1, x_2, x_3) = 0 \iff (x_2, x_3) = f(x_1)\big)$$

où $f : x_1 \in \mathbb{R} \mapsto \big(-\frac{3}{4}x_1, -\frac{1}{4}x_1\big) \in \mathbb{R}^2$.

Le théorème des fonctions implicites dans le cas général

Le théorème suivant étend au cas d'une fonction F quelconque le résultat obtenu dans le cas où F est une application linéaire. On notera toutefois une différence notable : la relation $\mathbf{y} = f(\mathbf{x})$ n'est plus valable que localement.

THÉORÈME 11.4 (Théorème des fonctions implicites)
Soient F une application de classe $\mathcal{C}^k$ ($k \in \mathbb{N}^* \cup \{\infty\}$) définie sur un ouvert U de $\mathbb{R}^n \times \mathbb{R}^q$, à valeurs dans $\mathbb{R}^q$ et $(\mathbf{a}, \mathbf{b})$ un élément de U.
Si $F(\mathbf{a}, \mathbf{b}) = \mathbf{0}$ et si

$$\det \begin{pmatrix} \dfrac{\partial F_1}{\partial y_1}(\mathbf{a}, \mathbf{b}) & & \dfrac{\partial F_1}{\partial y_q}(\mathbf{a}, \mathbf{b}) \\ \vdots & & \vdots \\ \dfrac{\partial F_q}{\partial y_1}(\mathbf{a}, \mathbf{b}) & & \dfrac{\partial F_q}{\partial y_q}(\mathbf{a}, \mathbf{b}) \end{pmatrix} \neq 0,$$

alors il existe :
- un ouvert V de $\mathbb{R}^n$ contenant $\mathbf{a}$,
- un ouvert W de $\mathbb{R}^q$ contenant $\mathbf{b}$ avec $V \times W \subset U$,
- et une unique application $f \in \mathcal{C}^k(V, W)$ tels que pour tout $(\mathbf{x}, \mathbf{y}) \in V \times W$

$$F(\mathbf{x}, \mathbf{y}) = 0 \iff \mathbf{y} = f(\mathbf{x}).$$

En particulier, on a $f(\mathbf{a}) = \mathbf{b}$ et $F(\mathbf{x}, f(\mathbf{x})) = 0$ pour tout $\mathbf{x} \in V$.

Démonstration Nous allons généraliser le raisonnement qui a permis de démontrer le théorème 11.2 et qui correspondait au cas $n = 1$ et $q = 1$. Considérons l'application G de U dans $\mathbb{R}^n \times \mathbb{R}^q$ définie pour $(\mathbf{x}, \mathbf{y}) \in U$ par $G(\mathbf{x}, \mathbf{y}) = (\mathbf{x}, F(\mathbf{x}, \mathbf{y}))$. D'après les hypothèses, cette application est de classe $\mathcal{C}^k$ et son jacobien en $(\mathbf{a}, \mathbf{b})$ vaut[(99)]

$$\mathcal{J}_G(\mathbf{a}, \mathbf{b}) = \det \begin{pmatrix} \mathrm{Id}_n & \mathrm{M}_1 \\ 0 & \mathrm{M}_2 \end{pmatrix} = \det(\mathrm{M}_2),$$

où Id_n désigne la matrice identité de taille n et où les matrices $\mathrm{M}_1 \in \mathcal{M}_{qn}$ et $\mathrm{M}_2 \in \mathcal{M}_q$ ont pour termes généraux respectifs $\partial_{x_j} F_i(\mathbf{a}, \mathbf{b})$ et $\partial_{y_j} F_i(\mathbf{a}, \mathbf{b})$.

Par hypothèse, $\det(\mathrm{M}_2)$ est non nul et par conséquent il en est de même de $\mathcal{J}_G(\mathbf{a}, \mathbf{b})$. Le théorème d'inversion locale[(100)] assure donc l'existence d'un voisinage U_1 de $(\mathbf{a}, \mathbf{b})$ et d'un voisinage V_1 de $G(\mathbf{a}, \mathbf{b}) = (\mathbf{a}, \mathbf{0})$ tel que l'application G soit un $\mathcal{C}^k$-difféomorphisme de U_1 sur V_1. On peut supposer[(101)] que $U_1 = V \times W$ et que $V_1 = W \times Z$ où V, W et Z désignent respectivement des voisinages de $\mathbf{a} \in \mathbb{R}^n$, $\mathbf{b} \in \mathbb{R}^q$ et $0 \in \mathbb{R}$.

Notons $G^{-1} : (\chi, \zeta) \in V \times Z \mapsto (G_1^{-1}(\chi, \zeta), G_2^{-1}(\chi, \zeta)) \in V \times W$ la bijection réciproque de G. Puisque $G(\mathbf{a}, \mathbf{b}) = (\mathbf{a}, \mathbf{0})$, on a

$$G_1^{-1}(\mathbf{a}, \mathbf{0}) = \mathbf{a} \quad \text{et} \quad G_2^{-1}(\mathbf{a}, \mathbf{0}) = \mathbf{b}.$$

De plus, pour tout $(\chi, \zeta) \in V \times Z$, on a

$$\begin{aligned}(\chi, \zeta) &= G\left(G^{-1}(\chi, \zeta)\right) = G\left(G_1^{-1}(\chi, \zeta), G_2^{-1}(\chi, \zeta)\right) \\ &= \left(G_1^{-1}(\chi, \zeta), F\big(G_1^{-1}(\chi, \zeta), G_2^{-1}(\chi, \zeta)\big)\right)\end{aligned}$$

ce qui indique que $G_1^{-1} : (\chi, \zeta) \in V \times Z \mapsto \chi \in V$ et que

$$\forall (\chi, \zeta) \in V \times Z \qquad F\big(\chi, G_2^{-1}(\chi, \zeta)\big) = \zeta. \tag{22}$$

On déduit de ce qui précède que pour tout $(\mathbf{x}, \mathbf{y}) \in V \times W$, on a les équivalences

$$\begin{aligned}F(\mathbf{x}, \mathbf{y}) = \mathbf{0} &\iff G(\mathbf{x}, \mathbf{y}) = (\mathbf{x}, \mathbf{0}) &&\iff (\mathbf{x}, \mathbf{y}) = G^{-1}(\mathbf{x}, \mathbf{0}) \\ &\iff (\mathbf{x}, \mathbf{y}) = \left(G_1^{-1}(\mathbf{x}, \mathbf{0}), G_2^{-1}(\mathbf{x}, \mathbf{0})\right) \\ &\iff (\mathbf{x}, \mathbf{y}) = (\mathbf{x}, G_2^{-1}(\mathbf{x}, \mathbf{0})) &&\iff \mathbf{y} = G_2^{-1}(\mathbf{x}, \mathbf{0}).\end{aligned}$$

On peut donc en conclure qu'il existe un ouvert V de $\mathbb{R}^n$ contenant $\mathbf{a}$, un ouvert W de $\mathbb{R}^q$ contenant $\mathbf{b}$ avec $V \times W \subset U$, et une unique application $f : \mathbf{x} \in V \mapsto G_2^{-1}(\mathbf{x}, \mathbf{0}) \in W$ de classe $\mathcal{C}^k$ tels que pour tout $(\mathbf{x}, \mathbf{y}) \in V \times W$

$$F(\mathbf{x}, \mathbf{y}) = 0 \iff \mathbf{y} = f(\mathbf{x}).$$

---

(99) Le jacobien est ici symboliquement écrit de façon condensée « par blocs » ; on pourra écrire ce déterminant sous forme explicite et le développer successivement par rapport aux n premières colonnes pour établir le résultat indiqué : il est égal à $\det(M_2)$.

(100) Voir le théorème 10.4 p. 665.

(101) Quitte à considérer des voisinages plus « petits ».

De plus, on a $f(\mathbf{a}) = G_2^{-1}(\mathbf{a}, \mathbf{0}) = \mathbf{b}$ et d'après la relation (22), pour tout $\mathbf{x} \in V$ on a

$$F(\mathbf{x}, f(\mathbf{x})) = F(\mathbf{x}, G_2^{-1}(\mathbf{x}, \mathbf{0})) = \mathbf{0}.$$

Le théorème est démontré. □

Exemple Considérons l'équation $F(x_1, x_2, x_3) = \mathbf{0}$ où

$$F : (x_1, x_2, x_3) \in \mathbb{R}^3 \longmapsto \big(2x_1 + 3x_2 - x_3 \cos(x_3),\ x_1 + x_2 + x_3\big) \in \mathbb{R}^2$$

et montrons qu'il existe un voisinage V de 0 dans $\mathbb{R}$ et un voisinage W de $(0, 0)$ dans $\mathbb{R}^2$ tel que pour tout réel x_1 fixé dans V, l'équation admet une unique solution (x_2, x_3) dans W.

Utilisons le théorème 11.4 avec $n = 1$ et $q = 2$; l'application F est de classe $\mathcal{C}^\infty$ sur l'ouvert $U = \mathbb{R}^3$, on a $F(0, 0, 0) = \mathbf{0}$ et

$$\det \begin{pmatrix} \dfrac{\partial F_1}{\partial x_2}(0,0,0) & \dfrac{\partial F_1}{\partial x_3}(0,0,0) \\ \dfrac{\partial F_2}{\partial x_2}(0,0,0) & \dfrac{\partial F_2}{\partial x_3}(0,0,0) \end{pmatrix} = \det \begin{pmatrix} 3 & -1 \\ 1 & 1 \end{pmatrix} = 4 \neq 0.$$

Le théorème 11.4 assure l'existence d'un ouvert V dans $\mathbb{R}$ contenant 0, d'un ouvert W de $\mathbb{R}^2$ contenant $(0, 0)$ et d'une unique application f définie sur V et à valeurs dans W telle que pour tout $(x_1, (x_2, x_3)) \in V \times W$ on ait

$$F(x_1, x_2, x_3) = 0 \iff (x_2, x_3) = f(x_1).$$

Autrement dit, pour tout $x_1 \in V$ fixé, il existe un unique couple $(x_2, x_3) \in W$ (donné par $(x_2, x_3) = f(x_1)$) solution de l'équation $F(x_1, x_2, x_3) = \mathbf{0}$.

On remarquera avec cet exemple que le théorème des fonctions implicites est un théorème d'existence et qu'il est en pratique souvent impossible d'expliciter l'application f. Le théorème des fonctions implicites indique à quelles conditions une équation de la forme $F(\mathbf{x}, \mathbf{y}) = \mathbf{0}$ admet au voisinage de $(\mathbf{a}, \mathbf{b})$ une unique solution $\mathbf{y}$ de la forme $\mathbf{y} = f(\mathbf{x})$ mais ne fournit pas explicitement l'application f. Il ne fournit pas non plus d'information sur les ouverts V et W sur lesquels l'équivalence a lieu. On sait seulement que $f(\mathbf{a}) = \mathbf{b}$ et que $F(\mathbf{x}, f(\mathbf{x})) = \mathbf{0}$ dans un voisinage de $\mathbf{a}$. Ces deux informations permettent toutefois de connaître toutes les dérivées partielles de l'application f en $\mathbf{a}$, et par conséquent d'obtenir des développements limités pour f au voisinage de $\mathbf{a}$.

Exemple Considérons à nouveau l'équation $F(x_1, x_2, x_3) = \mathbf{0}$ où

$$F : (x_1, x_2, x_3) \in \mathbb{R}^3 \longmapsto \big(2x_1 + 3x_2 - x_3 \cos(x_3),\ x_1 + x_2 + x_3\big) \in \mathbb{R}^2.$$

Nous avons vu qu'il existe une unique application f définie sur un voisinage V de 0 et à valeurs dans un voisinage W de $(0, 0)$ telle que pour tout $(x_1, (x_2, x_3))$ dans $V \times W$, on ait

$$F(x_1, x_2, x_3) = \mathbf{0} \iff (x_2, x_3) = f(x_1).$$

D'après le théorème des fonctions implicites, on sait que $f(0) = (0,0)$ et $F(x, f(x)) = \mathbf{0}$ pour tout $x \in V$. En notant f_1 et f_2 (resp. F_1 et F_2) les deux applications coordonnées de f (resp. de F), on a $f_1(0) = 0$, $f_2(0) = 0$ et

$$\forall x_1 \in V \qquad \big(F_1(x_1, f_1(x_1), f_2(x_1)) = 0 \quad \text{et} \quad F_2(x_1, f_1(x_1), f_2(x_1)) = 0 \big).$$

En dérivant les deux membres de ces égalités par rapport à x_1, on obtient

$$\begin{array}{rlll} & \partial_1 F_1(x_1, f_1(x_1), f_2(x_1)) & + & f_1'(x_1)\, \partial_2 F_1(x_1, f_1(x_1), f_2(x_1)) \\ & & + & f_2'(x_1)\, \partial_3 F_1(x_1, f_1(x_1), f_2(x_1)) \quad = \quad 0 \\ \text{et} & \partial_1 F_2(x_1, f_1(x_1), f_2(x_1)) & + & f_1'(x_1)\, \partial_2 F_2(x_1, f_1(x_1), f_2(x_1)) \\ & & + & f_2'(x_1)\, \partial_3 F_2(x_1, f_1(x_1), f_2(x_1)) \quad = \quad 0. \end{array}$$

Or $\partial_1 F_1(x_1, x_2, x_3) = 2$, $\partial_2 F_1(x_1, x_2, x_3) = 3$,

$$\partial_3 F_1(x_1, x_2, x_3) = -\cos(x_3) + x_3 \sin(x_3)$$

et $\partial_1 F_2(x_1, x_2, x_3) = 1$, $\partial_2 F_2(x_1, x_2, x_3) = 1$, $\partial_3 F_3(x_1, x_2, x_3) = 1$. Comme $f_1(0) = f_2(0) = 0$, en $x_1 = 0$ les équations précédentes s'écrivent

$$\begin{cases} 2 + 3f_1'(0) - f_2'(0) = 0 \\ 1 + f_1'(0) + f_2'(0) = 0 \end{cases}.$$

Le système obtenu est un système linéaire de Cramer[102] dont les solutions sont $f_1'(0) = -3/4$ et $f_2'(0) = -1/4$. On en déduit que le développement limité à l'ordre 1 de la fonction f au voisinage de 0 est :

$$f(x_1) = \big(-\tfrac{3}{4}x_1 + o_0(x_1), -\tfrac{1}{4}x_1 + o_0(x_1)\big).$$

Une approximation de la solution de l'équation $F(x_1, x_2, x_3) = 0$ au voisinage de $(0,0,0)$ est donc $(x_1, -\frac{3}{4}x_1, -\frac{1}{4}x_1)$ et l'erreur commise est de l'ordre de x_1.

Nous avons énoncé deux versions[103] du théorème des fonctions implicites pour les fonctions de $\mathbb{R}^2$ dans $\mathbb{R}$, la première indiquant à quelles conditions les solutions de l'équation $F(x,y) = 0$ s'expriment sous la forme $y = f(x)$ et la seconde à quelles conditions elles s'expriment sous la forme $x = f(y)$. Pour ce qui est des fonctions de $\mathbb{R}^{n+q}$ dans $\mathbb{R}^q$, il existe autant de versions du théorème des fonctions implicites que de possibilités de choisir q variables parmi $n+q$, c'est-à-dire C_{n+q}^q versions ! Ces différentes versions s'obtiennent à partir du théorème 11.4 en échangeant le rôle de certaines variables. Nous ne donnerons pas d'énoncé général mais nous allons illustrer sur un exemple les modifications à apporter dans la pratique au théorème 11.4 pour appliquer ces variantes.

---

[102] De façon générale, le système linéaire obtenu par cette approche est toujours un système de Cramer car son déterminant est celui donné dans le théorème 11.4 par hypothèse est non nul.

[103] Voir les théorèmes 11.2 et 11.3.

Exemple Considérons à nouveau l'équation $F(x_1, x_2, x_3) = \mathbf{0}$ où

$$F : (x_1, x_2, x_3) \in \mathbb{R}^3 \longmapsto \big(2x_1 + 3x_2 - x_3\cos(x_3)\ ,\ x_1 + x_2 + x_3\big) \in \mathbb{R}^2$$

et intéressons-nous à la possibilité, au voisinage de $(0,0,0)$, d'exprimer les solutions de l'équation sous la forme $x_3 = f(x_1, x_2)$.

L'application F est de classe $\mathcal{C}^\infty$ sur l'ouvert $U = \mathbb{R}^3$, on a $F(0,0,0) = \mathbf{0}$ et

$$\det \begin{pmatrix} \partial_1 F_1(0,0,0) & \partial_2 F_1(0,0,0) \\ \partial_1 F_2(0,0,0) & \partial_2 F_2(0,0,0) \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 1 & 1 \end{pmatrix} = -1 \neq 0.$$

On en déduit qu'il existe un ouvert V dans $\mathbb{R}$ contenant 0, un ouvert W de $\mathbb{R}^2$ contenant $(0,0)$ et une unique application f définie sur V et à valeurs dans W telle que pour tout $((x_1, x_2), x_3) \in W \times V$ on ait

$$F(x_1, x_2, x_3) = 0 \quad \Longleftrightarrow \quad (x_1, x_2) = f(x_3).$$

Autrement dit, pour tout $x_3 \in V$ fixé, il existe un unique couple $(x_1, x_2) \in W$ (donné par $(x_1, x_2) = f(x_2)$) solution de l'équation $F(x_1, x_2, x_3) = \mathbf{0}$.

Il est possible sur cet exemple d'expliciter la fonction f. On a

$$F(x_1, x_2, x_3) = \mathbf{0} \quad \Longleftrightarrow \quad \begin{cases} 2x_1 + 3x_2 - x_3\cos(x_3) = 0 \\ x_1 + x_2 + x_3 = 0 \end{cases}$$

On peut à partir de la seconde équation exprimer x_2 en fonction de x_1 et de x_3. En substituant la relation $x_2 = -x_1 - x_3$ dans la première équation, on obtient $x_1 = -x_3(3 + \cos(x_3))$ et on en déduit que $x_2 = x_3(2 + \cos(x_3))$. On a donc

$$f : x_3 \in \mathbb{R} \longmapsto \big(-x_3(3 + \cos(x_3)),\ x_3(2 + \cos(x_3))\big).$$

On peut vérifier que l'on a bien $f(0) = (0,0)$ et $F(f(x_3), x_3) = (0,0)$, comme le prévoit le théorème des fonctions implicites.

Exercice 11.12 Pour $(x_1, x_2) \in \mathbb{R}^2$ fixé, on s'intéresse à l'équation $F(x_1, x_2, x_3) = 0$ où

$$F : (x_1, x_2, x_3) \in \mathbb{R}^3 \mapsto x_1^3 + x_2^3 + x_3^3 - x_3x_1 - x_1 + x_2 - 2x_3 + 1 \in \mathbb{R}.$$

On note $(\mathcal{E}_{x_1,x_2})$ cette équation qui a pour inconnue le réel x_3 et pour paramètres les réels x_1 et x_2.

1 - Montrer que $x_3 = 1$ est une solution de l'équation $(\mathcal{E}_{0,0})$.

2 - Montrer que pour (x_1, x_2) « proche » de $(0,0)$, il existe une unique solution x_3 « proche » de 1 à l'équation $(\mathcal{E}_{x_1,x_2})$. On note par la suite $f(x_1, x_2)$ cette solution.

3 - Donner un développement limité au voisinage de $(0,0)$ à l'ordre 2 de la fonction $f : (x_1, x_2) \in \mathbb{R}^2 \mapsto f(x_1, x_2)$.

11.6.4 Application du théorème des fonctions implicites à l'étude des surfaces implicites dans l'espace

Le théorème des fonctions implicites permet d'obtenir une condition pour qu'une surface implicite de $\mathbb{R}^3$ admette une représentation cartésienne au voisinage d'un de ses points. Avant de l'énoncer, nous devons définir la notion de point régulier pour une surface implicite.

DÉFINITION 11.23 (Point régulier d'une surface implicite)
Soient U un ouvert de $\mathbb{R}^3$ et F une application de U dans $\mathbb{R}$ différentiable. Un point $M_0 = (x_0, y_0, z_0)$ de la surface implicite d'équation $F(x, y, z) = 0$ est dit régulier *si $\nabla F(x_0, y_0, z_0) \neq \mathbf{0}$. Dans le cas contraire, il est dit* singulier.

Remarque Comme pour les courbes de $\mathbb{R}^2$, nous avons deux notions de points réguliers pour une surface. L'une a trait aux surfaces paramétrées (voir la définition 11.19), l'autre aux surfaces implicites. On prendra garde au fait qu'un point d'une surface décrite à la fois comme surface implicite et comme surface paramétrée peut être régulier pour l'une des deux descriptions et singulier pour l'autre. On notera toutefois qu'une surface cartésienne est toujours une surface régulière qu'elle soit considérée comme surface paramétrée ou comme surface implicite. ◇

Considérons un point régulier $M_0 = (x_0, y_0, z_0)$ de la surface implicite d'équation $F(x, y, z) = 0$ où F désigne une fonction de classe $\mathcal{C}^1$ définie sur un ouvert U de $\mathbb{R}^3$ et à valeurs dans $\mathbb{R}$. On a $F(x_0, y_0, z_0) = 0$ et $\nabla F(x_0, y_0, z_0) \neq \mathbf{0}$ ce qui implique que l'une au moins des trois dérivées partielles de F est non nulle en (x_0, y_0, z_0). Supposons qu'il s'agisse de la troisième dérivée partielle. Nous sommes en mesure d'appliquer le théorème 11.4 avec $m = 3$ et $q = 1$: il existe un ouvert V de $\mathbb{R}^2$ contenant (x_0, y_0), un intervalle ouvert J de $\mathbb{R}$ contenant z_0 avec $V \times J \subset U$ et une unique application f de V dans J de classe $\mathcal{C}^1$ tels que pour tout $((x, y), z) \in V \times J$

$$F(x, y, z) = 0 \iff z = f(x, y).$$

Cela signifie que la surface implicite d'équation $F(x, y, z) = 0$ coïncide au voisinage de (x_0, y_0, z_0) avec la surface cartésienne d'équation $z = f(x, y)$. Dans le cas où c'est une autre dérivée partielle de F en (x_0, y_0, z_0) qui est non nulle, on peut adapter le raisonnement en échangeant les variables x, y et z. On a ainsi établi le résultat suivant.

PROPOSITION 11.18 *Soit F est une application de classe $\mathcal{C}^1$ définie sur un ouvert U de $\mathbb{R}^3$ et à valeurs dans $\mathbb{R}$. Au voisinage de tout point régulier, la surface implicite d'équation $F(x, y, z) = 0$ admet une représentation cartésienne de l'une des formes suivantes : $z = f(x, y)$ ou $y = f(x, z)$ ou $x = f(y, z)$, où f désigne une fonction de classe $\mathcal{C}^1$ de $\mathbb{R}^2$ dans $\mathbb{R}$.*

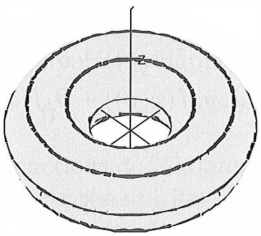
Pour une surface implicite S dont tous les points sont réguliers, la proposition 11.18 affirme que cette surface peut être localement représentée par une équation cartésienne d'un des types suivants : $z = f(x, y)$ ou $y = f(x, z)$ ou $x = f(y, z)$. Mais rien n'assure que ce soit le même type de représentation en tous les points de la surface. Ainsi, il est possible qu'au voisinage d'un point $M_0 \in S$ on sache représenter la surface sous la forme $z = f_1(x, y)$ et qu'en un autre point on ne puisse le faire que sous la forme $y = f_2(x, z)$ ou $x = f_3(y, z)$.

Exemple Soient F l'application de $\mathbb{R}^3$ dans $\mathbb{R}$ définie par

$$F : (x, y, z) \in \mathbb{R}^3 \longmapsto (x^2 + y^2 + z^2 + 3)^2 - 16(x^2 + y^2),$$

et S la surface implicite de $\mathbb{R}^3$ d'équation $F(x, y, z) = 0$. La surface S est un tore qui peut être visualisé avec MAPLE.

```
> implicitplot3d((x^2+y^2+z^2+3)^2=16*(x^2+y^2),
                 x=-3..3, y=-3..3, z=-3..3);
```



Montrons que tous les points de la surface S sont réguliers. Un point $M_0 = (x_0, y_0, z_0)$ de S est un point singulier si

$$\nabla F(x_0, y_0, z_0) = \begin{pmatrix} 4x_0(x_0^2 + y_0^2 + z_0^2 - 5) \\ 4y_0(x_0^2 + y_0^2 + z_0^2 - 5) \\ 4z_0(x_0^2 + y_0^2 + z_0^2 + 3) \end{pmatrix} = \mathbf{0}.$$

Ses coordonnées (x_0, y_0, z_0) vérifient donc

$$\begin{aligned} &(x_0^2 + y_0^2 + z_0^2 + 3)^2 = 16(x_0^2 + y_0^2), \\ &\text{et } x_0 = 0 \quad \text{ou} \quad x_0^2 + y_0^2 + z_0^2 = 5, \\ &\text{et } y_0 = 0 \quad \text{ou} \quad x_0^2 + y_0^2 + z_0^2 = 5, \\ &\text{et } z_0 = 0 \quad \text{car} \quad x_0^2 + y_0^2 + z_0^2 + 3 \neq 0. \end{aligned}$$

La première équation indique que x_0, y_0 et z_0 ne peuvent pas être tous nuls. On a $z_0 = 0$ (dernière équation) ce qui implique que $x_0^2 + y_0^2 = 5$. Cela est incompatible avec la première équation. On en conclut qu'il n'existe pas de point singulier, autrement dit que tous les points du tore sont des points réguliers. D'après la proposition 11.18, il s'ensuit que la surface S coïncide au voisinage de chacun de ses points avec une surface cartésienne.

Exercice 11.13 Soit S la surface implicite d'équation $F(x,y,z)=0$ où

$$F : (x,y,z) \in [-2,2]^3 \longmapsto x^3+y^3+z^3+1-(x+y+z+1)^3.$$

Montrer que tous les points de la surface S sont réguliers.

Reprenons l'étude générale d'une surface implicite dans l'espace ; considérons un point régulier $M_0=(x_0,y_0,z_0)$ de la surface implicite de classe $\mathcal{C}^1$ de $\mathbb{R}^3$ d'équation $F(x,y,z)=0$ et supposons[104] que $\partial_z F(x_0,y_0,z_0) \neq 0$. On a vu que le théorème des fonctions implicites assure l'existence d'un ouvert V de $\mathbb{R}^2$ contenant (x_0,y_0), d'un intervalle ouvert J de $\mathbb{R}$ contenant z_0 avec $V\times J\subset U$ et d'une unique application f de V dans J de classe $\mathcal{C}^1$ tels que pour tout $((x,y),z)\in V\times J$

$$F(x,y,z)=0 \iff z=f(x,y).$$

De plus, $z_0=f(x_0,y_0)$ et pour tout $(x,y)\in V$ on a $F(x,y,f(x,y))=0$. Dérivons[105] cette relation par rapport à x et à y ; on a pour tout $(x,y)\in V$

$$\begin{cases} \partial_x F(x,y,f(x,y)) = \partial_1 F(x,y,f(x,y)) + \partial_1 f(x,y)\,\partial_3 F(x,y,f(x,y)) = 0 \\ \partial_y F(x,y,f(x,y)) = \partial_2 F(x,y,f(x,y)) + \partial_2 f(x,y)\,\partial_3 F(x,y,f(x,y)) = 0 \end{cases}$$

ce qui indique que

$$\partial_1 f(x_0,y_0) = -\frac{\partial_1 F(x_0,y_0,z_0)}{\partial_3 F(x_0,y_0,z_0)} \quad \text{et} \quad \partial_2 f(x_0,y_0) = -\frac{\partial_2 F(x_0,y_0,z_0)}{\partial_3 F(x_0,y_0,z_0)}.$$

Au voisinage du point M_0, la surface implicite d'équation $F(x,y,z)=0$ coïncide avec la surface cartésienne d'équation $z=f(x,y)$. On a vu à la p. 731, qu'un vecteur normal à la surface cartésienne d'équation $z=f(x,y)$ au point (x_0,y_0,z_0) est $\mathbf{n}=\big(-\partial_1 f(x_0,y_0),\ -\partial_2 f(x_0,y_0),\ 1\big)$. On en déduit qu'un vecteur normal à la surface implicite d'équation $F(x,y,z)=0$ au point (x_0,y_0,z_0) est

$$\mathbf{n}=\begin{pmatrix} \dfrac{\partial_1 F(x_0,y_0,z_0)}{\partial_3 F(x_0,y_0,z_0)} \\ \dfrac{\partial_1 F(x_0,y_0,z_0)}{\partial_3 F(x_0,y_0,z_0)} \\ 1 \end{pmatrix} = \frac{1}{\partial_3 F(x_0,y_0,z_0)}\,\nabla F(x_0,y_0,z_0).$$

On a donc le résultat suivant.

---

[104] Dans le cas où $\partial_z F(x_0,y_0,z_0)=0$, on peut effectuer un raisonnement similaire avec une autre des dérivées partielles de F.

[105] On note dans ce qui suit $\partial_1 F$, $\partial_2 F$ et $\partial_3 F$ les trois dérivées partielles premières de F et on note $\partial_x[F(x,y,f(x,y))]$ la dérivée de la fonction $(x,y)\mapsto F(x,y,f(x,y))$ par rapport à la première variable x (même chose pour y).

Proposition 11.19 (Vecteur normal à une surface implicite)
Soient F est une application de classe $\mathcal{C}^1$ définie sur un ouvert U de $\mathbb{R}^3$, à valeurs dans $\mathbb{R}$ et M_0 un point régulier de la surface implicite S d'équation $F(x,y,z)=0$. Un vecteur normal à la surface S au point M_0 est le vecteur $\nabla F(x_0,y_0,z_0)$.

On a également vu p. 731 que le plan tangent à la surface cartésienne d'équation $z=f(x,y)$ au point $M_0=(x_0,y_0,z_0)$ est le plan passant par M_0 dont un vecteur normal est le vecteur $\mathbf{n}=\big(-\partial_1 f(x_0,y_0),\ -\partial_2 f(x_0,y_0),\ 1\big)$. On en déduit que la plan tangent à la surface implicite S d'équation $F(x,y,z)=0$ en un point régulier M_0 de coordonnées (x_0,y_0,z_0) a pour équation

$$\partial_1 F(x_0,y_0,z_0)\,(x-x_0)+\partial_2 F(x_0,y_0,z_0)\,(y-y_0)+\partial_3 F(x_0,y_0,z_0)\,(z-z_0)=0.$$

11.6.5 Application du théorème des fonctions implicites à l'étude des courbes implicites dans l'espace

Nous n'avons jusqu'à présent jamais étudié les courbes implicites dans $\mathbb{R}^3$. La raison en a été donnée à la p. 720 : il est difficile d'un point de vue mathématique de définir simplement la notion de courbe implicite dans l'espace. On peut vouloir définir une courbe implicite de $\mathbb{R}^3$ comme l'intersection de deux surfaces implicites : si S_1 et S_2 sont des surfaces implicites d'équation

$$F_1(x,y,z)=0 \qquad \text{et} \qquad F_2(x,y,z)=0$$

où F_1 et F_2 désignent deux fonctions de $\mathbb{R}^3$ dans $\mathbb{R}$ alors l'intersection des deux surfaces S_1 et S_2 admet pour représentation implicite $F(x,y,z)=\mathbf{0}$ où F est la fonction de $\mathbb{R}^3$ dans $\mathbb{R}^2$ définie par $F:\mathbf{x}\mapsto(F_1(\mathbf{x}),F_2(\mathbf{x}))$. Toutefois, on sait que l'intersection de deux surfaces n'est pas toujours une courbe (il suffit par exemple de considérer le cas où $S_1=S_2$) ce qui ne permet pas d'envisager cette façon de définir une courbe implicite dans l'espace en toute généralité. On peut introduire la notion de point régulier d'une « courbe implicite » dans l'espace pour éviter les cas dégénérés où l'intersection de deux surfaces n'est pas une courbe au sens usuel. On dit qu'un point $\mathbf{x}_0$ de la « courbe » C est *régulier* si $\operatorname{rg}(\mathrm{J}_F(\mathbf{x}_0))=2$.

Exemples

1. Considérons deux plans P_1 et P_2 dans $\mathbb{R}^3$ d'équations respectives

$$a_1x+b_1y+c_1z=d_1 \qquad \text{et} \qquad a_2x+b_2y+c_2z=d_2.$$

L'intersection de ces deux plans admet pour représentation implicite $F(x,y,z)=0$ où

$$F:(x,y,z)\in\mathbb{R}^3\longmapsto(a_1x+b_1y+c_1z-d_1,\ a_2x+b_2y+c_2z-d_2)\in\mathbb{R}^2.$$

La matrice jacobienne de F au point $\mathbf{x}_0\in\mathbb{R}^3$ est

$$\mathrm{J}_F(\mathbf{x}_0)=\begin{pmatrix}a_1 & b_1 & c_1\\ a_2 & b_2 & c_2\end{pmatrix}.$$

Cette matrice est de rang 2 si et seulement si les vecteurs (a_1, b_1, c_1) et (a_2, b_2, c_2) ne sont pas colinéaires. Ces vecteurs étant aussi des vecteurs normaux aux plans P_1 et P_2, on en déduit que l'intersection des deux plans est une courbe régulière si et seulement si les deux plans ne sont pas parallèles.

2. Considérons les deux surfaces implicites S_1 et S_2 de $\mathbb{R}^3$ d'équations[(106)] $F_1(x, y, z) = 0$ et $F_2(x, y, z) = 0$ où

$$F_1 : (x, y, z) \in \mathbb{R}^3 \longmapsto (x^2 + y^2 + z^2 + 3)^2 - 16(x^2 + y^2),$$
$$\text{et} \quad F_2 : (x, y, z) \in \mathbb{R}^3 \longmapsto x^3 + y^3 + z^3 + 1 - (x + y + z + 1)^3.$$

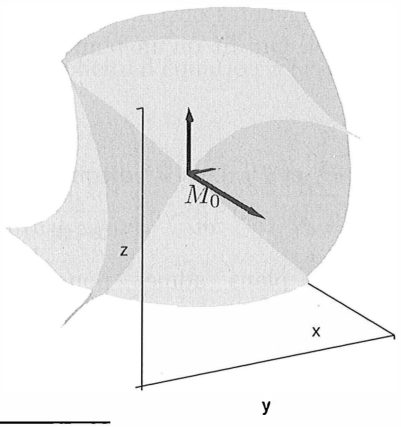
Le point de coordonnées $(-1, 0, 0) \in \mathbb{R}^3$ appartient aux deux surfaces S_1 et S_2 car $F_1(-1, 0, 0) = F_2(-1, 0, 0) = 0$. Pour déterminer si ce point est un point régulier ou non de la courbe $S_1 \cap S_2$, il faut déterminer le rang de la jacobienne de l'application $F : \mathbf{x} \in \mathbb{R}^3 \mapsto (F_1(\mathbf{x}), F_2(\mathbf{x}))$ en $\mathbf{x}_0 = (-1, 0, 0)$. On a

$$\mathrm{J}_F(\mathbf{x}_0) = \begin{pmatrix} \partial_1 F_1(\mathbf{x}_0) & \partial_2 F_1(\mathbf{x}_0) & \partial_3 F_1(\mathbf{x}_0) \\ \partial_1 F_2(\mathbf{x}_0) & \partial_2 F_2(\mathbf{x}_0) & \partial_3 F_2(\mathbf{x}_0) \end{pmatrix} = \begin{pmatrix} 16 & 0 & 0 \\ 3 & 0 & 0 \end{pmatrix}.$$

Le rang de cette matrice vaut 1 car tous les déterminants d'ordre 2 extraits de la matrice $\mathrm{J}_F(\mathbf{x}_0)$ sont nuls. On en déduit que le point $(-1, 0, 0)$ n'est pas un point régulier de la courbe $S_1 \cap S_2$.

Graphiquement, la figure ci-dessous montre qu'au voisinage du point $(-1, 0, 0)$ l'intersection des deux surfaces est la réunion de deux courbes qui s'intersectent en $(-1, 0, 0)$.

```
> fig1:=implicitplot3d((x^2+y^2+z^2+3)^2=16*(x^2+y^2),x=-1.2..-0.8,
     y=-0.5..0.5,z=-0.5..0.5,color=green,style=patchnogrid):
> fig2:=implicitplot3d(x^3+y^3+z^3+1=(x+y+z+1)^3, x=-1.2..-0.8,
     y=-0.5..0.5,z=-0.5..0.5,color=gray,style=patchnogrid):
> fig3:=arrow({[-1,0,0]},{[0,0,1],[0,1,0],[1,0,0]}):
> display3d({fig1,fig2,fig3},orientation=[-30,70],axes=frame);
```



(106) Ces deux surfaces ont déjà été étudiées à la p. 827 pour la surface S_1 et à l'exercice 11.13 pour la surface S_2. En particulier, on a montré que ces deux surfaces implicites sont régulières.

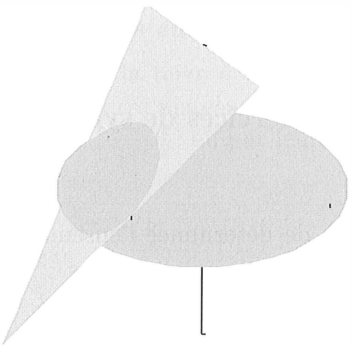
3. Considérons l'intersection des deux surfaces implicites d'équation $x^2 + y^2 + 4z^2 = 1$ et $x + y + z = 1$. En notant F l'application $(x, y, z) \in \mathbb{R}^3 \mapsto (x^2+y^2+4z^2-1, x+y+z-1)$, on a pour tout $(x, y, z) \in \mathbb{R}^3$,

$$\mathrm{J}_F(x, y, z) = \begin{pmatrix} 2x & 2y & 8z \\ 1 & 1 & 1 \end{pmatrix}.$$

Les 3 déterminants d'ordre 2 issus de $\mathrm{J}_F(x, y, z)$ valent $2(x - y)$, $2(x - 4z)$ et $2(y - 4z)$. Pour que le rang de $\mathrm{J}_F(x, y, z)$ ne soit pas égal à 2, il faut que ces 3 quantités soient nulles simultanément. Cela est réalisé si $(x, y, z) = (4\alpha, 4\alpha, \alpha)$ avec $\alpha \in \mathbb{R}$. Ces points n'appartiennent pas à l'intersection des deux surfaces car $F(4\alpha, 4\alpha, \alpha) = (9\alpha - 1, 36\alpha^2 - 1)$ ne peut jamais être nul. On en déduit que pour tous les points de l'intersection des deux surfaces, le rang de la matrice jacobienne J_F vaut 2. Tous les points de l'intersection des deux surfaces sont donc réguliers.

L'intersection de ces deux surfaces implicites, qui correspondent à un ellipsoïde et à un plan, est une ellipse[107] que l'on peut visualiser avec MAPLE.

```
> fig1:=implicitplot3d(x^2+y^2+4*z^2=1,x=-1..1,y=-1..1,z=-1..1):
> fig2:=implicitplot3d(x+y+z=1,x=-2..2,y=-2..2,z=-2..2):
> display(fig1,fig2);
```



Nous disposons maintenant de deux descriptions possibles pour une courbe dans l'espace : l'une sous forme paramétrée, l'autre sous forme implicite. Le théorème des fonctions implicites permet d'établir le lien suivant entre ces deux notions.

PROPOSITION 11.20 *Soient F une application de classe C^1 définie sur un ouvert U de $\mathbb{R}^3$ et à valeurs dans $\mathbb{R}^2$, et C la courbe implicite de $\mathbb{R}^3$ d'équation $F(x, y, z) = \mathbf{0}$. Au voisinage de tout point régulier, la courbe implicite C admet une représentation paramétrique.*

[107] Plus généralement, on montre que l'intersection d'une quadrique avec un plan est toujours une conique (qui peut être dégénérée).

Démonstration Considérons un point régulier $M_0 = (a, b, c)$ de la courbe implicite C d'équation $F(x, y, z) = \mathbf{0}$; on a $F(a, b, c) = \mathbf{0}$ et $\mathrm{rg}(\mathrm{J}_F(a, b, c)) = 2$. Puisque la matrice jacobienne de F en (a, b, c) est de rang 2, l'un au moins des déterminants suivant est non nul :

$$\begin{vmatrix} \partial_1 F_1(M_0) & \partial_2 F_1(M_0) \\ \partial_1 F_2(M_0) & \partial_2 F_2(M_0) \end{vmatrix}, \quad \begin{vmatrix} \partial_2 F_1(M_0) & \partial_3 F_1(M_0) \\ \partial_2 F_2(M_0) & \partial_3 F_2(M_0) \end{vmatrix}, \quad \begin{vmatrix} \partial_1 F_1(M_0) & \partial_3 F_1(M_0) \\ \partial_1 F_2(M_0) & \partial_3 F_2(M_0) \end{vmatrix}.$$

Supposons qu'il s'agisse du second[108]. Nous sommes en mesure d'appliquer le théorème des fonctions implicites avec $n = 1$ et $q = 2$: il existe un intervalle ouvert I contenant a, un ouvert W de $\mathbb{R}^2$ contenant (b, c) et une unique fonction f de I dans W telle que, pour $(x, y, z) \in I \times W$ on ait

$$F(x, y, z) = \mathbf{0} \iff (y, z) = f(x). \tag{23}$$

Désignons par f_1 et f_2 les deux composantes de l'application f et considérons la courbe paramétrée $\widetilde{C}$ de $\mathbb{R}^3$ dont un paramétrage est

$$\varphi : x \in I \longmapsto (x, f_1(x), f_2(x)) \in \mathbb{R}^3.$$

L'équivalence (23) exprime le fait que la courbe implicite C d'équation $F(x, y, z) = \mathbf{0}$ coïncide dans un voisinage de (a, b, c) avec la courbe paramétrée $\widetilde{C}$. □

11.7 Exercices de synthèse

Exercice 11.14

L'objectif de cet exercice est de déterminer l'ensemble des points de $\mathbb{R}^3$ par lesquels on peut mener trois plans perpendiculaires[109] entre eux et tangents à un même ellipsoïde.

On considère un ellipsoïde dont l'équation implicite est

$$\mathcal{E} \ : \quad \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1,$$

où a, b et c désignent trois réels non nuls.

1 - Soient A un point de $\mathbb{R}^3$ de coordonnées (x_0, y_0, z_0) et $\mathbf{e}_1$ un vecteur de norme 1 et de composantes (α, β, γ). On désigne par $\mathcal{P}$ le plan passant par le point A dont un vecteur normal est $\mathbf{e}_1$. Donner l'équation de $\mathcal{P}$.

2 - Écrire l'équation du plan tangent à l'ellipsoïde en un point M de coordonnées (x_1, y_1, z_1).

---

[108] Un raisonnement analogue, consistant à échanger le rôle des variables, conduit à un résultat semblable dans les autres cas.

[109] On dit que deux plans sont perpendiculaires si leurs vecteurs normaux sont orthogonaux entre eux.

3 - Soient $\mathcal{P}_1$ et $\mathcal{P}_2$ deux plans d'équation respective

$$a_1x + b_1y + c_1z = d_1 \quad \text{et} \quad a_2x + b_2y + c_2z = d_2$$

où les réels a_2, b_2, c_2 et d_2 sont supposés non nuls.
Montrer que $\mathcal{P}_1 = \mathcal{P}_2 \iff \dfrac{a_1}{a_2} = \dfrac{b_1}{b_2} = \dfrac{c_1}{c_2} = \dfrac{d_1}{d_2}$.

4 - En déduire que le plan $\mathcal{P}$ défini à la question 1 est un plan tangent à l'ellipsoïde si et seulement si $q(\mathbf{e}_1) = 0$ où q est une forme quadratique dépendant du point A et des coefficients a, b et c.

5 - Montrer que l'on peut faire passer trois plans tangents à $\mathcal{E}$ et perpendiculaires entre eux par le point A si et seulement si il existe une base orthonormée $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ de $\mathbb{R}^3$ pour laquelle

$$\forall i \in \{1, 2, 3\} \quad q(\mathbf{e}_i) = 0.$$

En utilisant l'exercice 8.11, déterminer l'ensemble des points qui satisfont cette condition.

Exercice 11.15 ***Loxodromies sur la terre***

Dans cet exercice, on assimile la terre à une sphère S de $\mathbb{R}^3$ de centre 0 et de rayon R. Les points de la terre peuvent être représentés par leur longitude $\theta \in [0, \pi[$ et leur latitude $\varphi \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ ou par leurs coordonnées cartésiennes x, y et z, voir la figure 24.

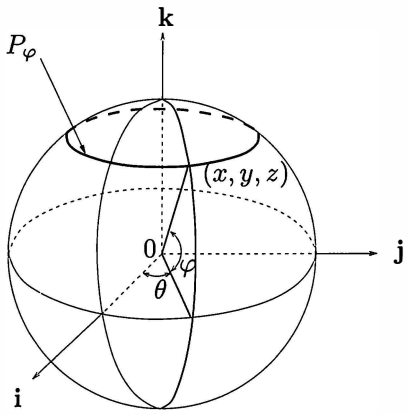


Fig. 24 Notations utilisées.

On a la correspondance suivante entre les coordonnées cartésiennes d'un point à la surface de la terre et sa longitude θ et sa latitude φ :

$$x = R\cos(\theta)\cos(\varphi), \quad y = R\sin(\theta)\cos(\varphi), \quad z = R\sin(\varphi).$$

Un parallèle P_{φ_0} est une courbe tracée sur la sphère S dont les points sont de même latitude φ_0 et dont un paramétrage est

$$\pi_{\varphi_0} : \theta \in \mathbb{R} \longmapsto (R\cos(\theta)\cos(\varphi_0),\ R\sin(\theta)\cos(\varphi_0),\ R\sin(\varphi_0)) \in S.$$

1 - a) Indiquer, selon les valeurs de $\varphi_0 \in [-\frac{\pi}{2}, \frac{\pi}{2}]$, les caractéristiques du parallèle P_{φ_0}.

b) Pour $\varphi_0 \in]-\frac{\pi}{2}, \frac{\pi}{2}[$, donner un vecteur tangent τ_P au parallèle P_{φ_0} en un point de longitude $\theta \in \mathbb{R}$ puis calculer la norme euclidienne de ce vecteur.

On considère des courbes C tracées sur la sphère S dont un paramétrage, donné en fonction de la longitude θ, est

$$\gamma : \theta \in I \longmapsto (R\cos(\theta)\cos(f(\theta)),\ R\sin(\theta)\cos(f(\theta)),\ R\sin(f(\theta))) \qquad (24)$$

où I désigne un intervalle de $\mathbb{R}$ et f une application de I dans $]-\frac{\pi}{2}, \frac{\pi}{2}[$ de classe $\mathcal{C}^2$.

2 - a) Indiquer à quelle application f correspond le parallèle P_{φ_0} ?

b) Dans le cas général, calculer en fonction de f et de f', le vecteur tangent τ_C à la courbe C en un point de longitude $\theta \in \mathbb{R}$. Donner la norme euclidienne de ce vecteur tangent.

Soit θ_0 un élément de I fixé et $\varphi_0 = f(\theta_0)$; les deux courbes P_{φ_0} et C associées aux arcs paramétrées π_{φ_0} et γ s'intersectent au point M_0 de coordonnées (θ_0, φ_0). On désigne par α avec $\alpha \in]-\frac{\pi}{2}, \frac{\pi}{2}[$, l'angle formé au point M_0 par le vecteur tangent τ_C à la courbe C et le vecteur tangent τ_P au parallèle P_{φ_0}.

3 - a) Calculer en fonction de α le produit scalaire et le produit vectoriel des deux vecteurs τ_C et τ_P.

b) En déduire que

$$|\tan(\alpha)| = \frac{|f'(\theta_0)|}{\cos(f(\theta_0))}.$$

On appelle *loxodromie* une courbe C qui est l'image d'un arc paramétré γ de la forme (24) et qui vérifie la propriété suivante : quel que soit le point de longitude $\theta \in I$ et de latitude $\varphi = f(\theta)$ de la courbe C, l'angle α formé par la courbe et le parallèle P_φ est constant[(110)]. L'image d'un arc paramétré γ est donc une loxodromie si

$$\exists \alpha \in]-\tfrac{\pi}{2}, \tfrac{\pi}{2}[\quad \forall \theta \in I \qquad \frac{|f'(\theta)|}{\cos(f(\theta))} = |\tan(\alpha)|.$$

(110) Sur une carte de Mercator, les parallèles étant représentés par des segments de droites parallèles, une loxodromie est elle-même représentée par un segment de droite. Pour un navire, une route maritime entre deux ports qui correspond à une loxodromie à l'avantage d'être simple à suivre car il s'agit d'une route à cap constant. Elle n'est toutefois pas la plus courte en distance.

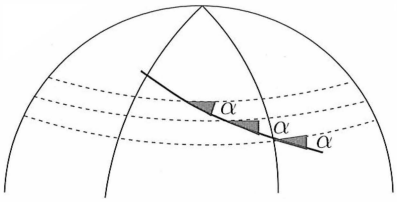


Fig. 25 Loxodromie : l'angle α formé par la courbe et les parallèles est constant.

4 - a) Montrer que l'application $g : x \in]-\frac{\pi}{2}, \frac{\pi}{2}[\mapsto \ln\left(\tan\left(\frac{x}{2} + \frac{\pi}{4}\right)\right)$ est dérivable puis montrer que $g'(x) = 1/\cos(x)$ pour tout $x \in]-\frac{\pi}{2}, \frac{\pi}{2}[$.

b) En déduire que si la courbe C est une loxodromie passant par le point de coordonnées (θ_1, φ_1) et faisant un angle constant $\alpha \in [0, \frac{\pi}{2}[$ avec les parallèles alors la longitude $\theta \in I$ et la latitude $\varphi = f(\theta)$ des points de cette courbe sont liées par la relation

$$\ln(\tan(\tfrac{1}{2}f(\theta) + \tfrac{\pi}{4})) - \ln(\tan(\tfrac{1}{2}\varphi_1 + \tfrac{\pi}{4})) = (\theta - \theta_1)\tan(\alpha).$$

On considère à présent une loxodromie pour laquelle l'application f est définie sur $\mathbb{R}$.

5 - a) Calculer la limite de f en $+\infty$ et $-\infty$. En déduire qu'une telle loxodromie traverse nécessairement l'équateur (c'est-à-dire le parallèle P_0) lorsque $\alpha > 0$.

b) Que se passe-t-il lorsque $\alpha = 0$?

11.8 Solutions des exercices

Solution de l'exercice 11.1

1 - Pour montrer que l'arc paramétré $\varphi_1 : t \in [0,1[\longmapsto (\cos(2\pi t^2), \sin(2\pi t^2))$ est simple, considérons deux valeurs t_1 et t_2 du paramètre dans $[0,1[$ telles que $\varphi_1(t_1) = \varphi_1(t_2)$. On a $\cos(2\pi t_1^2) = \cos(2\pi t_2^2)$ et $\sin(2\pi t_1^2) = \sin(2\pi t_2^2)$. Puisque $2\pi t_1^2 \in [0, 2\pi[$ et $2\pi t_2^2 \in [0, 2\pi[$ et que deux angles de $[0, 2\pi[$ ayant même sinus et cosinus sont égaux, on en déduit que $2\pi t_1^2 = 2\pi t_2^2$. Comme la fonction puissance 2 est injective sur $[0, 1[$, on en déduit que $t_1 = t_2$. L'arc paramétré φ_1 est donc un arc simple. On peut procéder de la même manière pour montrer que l'arc paramétré φ_2 est lui aussi un arc paramétré simple. L'image de l'arc paramétré φ_2 est le cercle de centre $(0,0)$ et de rayon 1. Lorsque le paramètre t varie dans l'intervalle $[0,1[$, le réel $2\pi t^2$ décrit l'intervalle $[0, 2\pi[$. L'image de l'arc paramétré φ_1 est donc elle aussi le cercle de centre $(0,0)$ et de rayon 1.

2 - On a $\varphi_1 = \varphi_2 \circ \theta_1$ où $\theta_1 : t \in [0,1[\mapsto 2\pi t^2 \in [0, 2\pi[$. L'application θ_1 est une bijection car la fonction puissance 2 est une bijection de $[0,1[$ dans $[0,1[$. On détermine aisément l'expression de la bijection réciproque en résolvant l'équation $s = 2\pi t^2$; on obtient $\theta_1^{-1} : t \in [0, 2\pi[\mapsto \sqrt{t/2\pi} \in [0,1[$. L'application θ_1 n'est pas un $\mathcal{C}^1$-difféomorphisme car θ_1^{-1} n'est pas dérivable en 0.

3 - On a $\varphi_3 = \varphi_4 \circ \theta_2$ où $\theta_2 : t \in]0,1] \mapsto 2\pi t^2 \in]0, 2\pi]$. L'application θ_2 est une bijection de classe $\mathcal{C}^\infty$ car la fonction puissance 2 est une bijection de $]0,1]$ dans $]0,1]$ et qu'elle est indéfiniment dérivable. Sa bijection réciproque qui est $\theta_2^{-1} : t \in]0, 2\pi] \mapsto \sqrt{t/2\pi} \in]0,1]$ est de classe $\mathcal{C}^\infty$ car la fonction racine carré est indéfiniment dérivable sur $]0,1]$. L'application θ_2 est donc un $\mathcal{C}^\infty$-difféomorphisme de $]0,1]$ dans $]0, 2\pi]$.

Solution de l'exercice 11.2

Pour montrer que la relation $\mathcal{R}_\mathcal{O}$ est une relation d'équivalence sur l'ensemble des arcs paramétrés, il faut montrer[111] qu'elle est réflexive, symétrique et transitive.

$\trianglerighteq$ La relation est réflexive car tout arc paramétré $\varphi \in \mathcal{C}^0(I, \mathbb{R}^n)$ est $\mathcal{C}^k$-équivalent à lui-même ; en effet, on a $\varphi = \varphi \circ \theta$ où $\theta : t \in I \mapsto t \in I$ est l'application identité et l'identité est un $\mathcal{C}^k$-difféomorphisme pour tout $k \in \mathbb{N}^* \cup \{\infty\}$ qui est strictement croissant.

$\trianglerighteq$ Montrons que la relation $\mathcal{R}_\mathcal{O}$ est symétrique. Si $\varphi \in \mathcal{C}^0(I, \mathbb{R}^n)$ est un arc paramétré $\mathcal{O}^k$-équivalent à l'arc paramétré $\psi \in \mathcal{C}^0(J, \mathbb{R}^n)$ alors il existe un $\mathcal{C}^k$-difféomorphisme θ de I sur J, strictement croissant, tel que $\varphi = \psi \circ \theta$. L'application θ étant bijective, on a $\psi = \varphi \circ \theta^{-1}$ où θ^{-1} désigne la bijection réciproque de θ.

---

[111] Voir la définition 2.28 p. 55 du *Cours de première année*.

Par définition d'un difféomorphisme[112], comme θ est un $\mathcal{C}^k$-difféomorphisme, l'application θ^{-1} est aussi un $\mathcal{C}^k$-difféomorphisme. Par ailleurs, une bijection et sa bijection réciproque ont même sens de variation[113] donc θ^{-1} est également strictement croissante. La relation $\psi = \varphi \circ \theta^{-1}$ permet alors de conclure que l'arc paramétré $\psi \in \mathcal{C}^0(J, \mathbb{R}^n)$ est $\mathcal{O}^k$-équivalent à l'arc paramétré $\varphi \in \mathcal{C}^0(I, \mathbb{R}^n)$.

$\trianglerighteq$ Montrons que la relation $\mathcal{R}_{\mathcal{O}}$ est transitive. Considérons trois arcs paramétrés $\varphi \in \mathcal{C}^0(I, \mathbb{R}^n)$, $\psi \in \mathcal{C}^0(J, \mathbb{R}^n)$ et $\zeta \in \mathcal{C}^0(K, \mathbb{R}^n)$ où I, J et K désignent trois intervalles de $\mathbb{R}$ et supposons d'une part que ψ est $\mathcal{O}^k$-équivalent à φ et d'autre part que ζ est $\mathcal{O}^k$-équivalent à ψ. Nous devons montrer que ζ est $\mathcal{O}^k$-équivalent à φ. D'après les hypothèses, il existe un $\mathcal{C}^k$-difféomorphisme θ_1 de I sur J, strictement croissant, tel que $\varphi = \psi \circ \theta_1$ et il existe un $\mathcal{C}^k$-difféomorphisme θ_2 de J sur K, strictement croissant, tel que $\psi = \zeta \circ \theta_2$. On en déduit que $\varphi = \zeta \circ \theta_2 \circ \theta_1$. L'application $\theta = \theta_2 \circ \theta_1$ qui est la composée de deux bijections de classe $\mathcal{C}^k$ est une bijection de classe $\mathcal{C}^k$ sur I. De plus, θ étant la composée de deux applications strictement croissantes, il s'agit d'une application strictement croissante. Sa bijection réciproque $\theta^{-1} = \theta_1^{-1} \circ \theta_2^{-1}$ qui est la composée de deux bijections de classe $\mathcal{C}^k$ est une bijection de classe $\mathcal{C}^k$ sur K. L'application θ est donc un $\mathcal{C}^k$-difféomorphisme de I sur K et la relation $\varphi = \zeta \circ \theta$ indique que l'arc paramétré ζ est $\mathcal{C}^k$-équivalent à l'arc paramétré φ.

Solution de l'exercice 11.3

1 - D'après la définition 11.7, les deux arcs paramétrés $\varphi \in \mathcal{C}^0(I, \mathbb{R}^n)$ et $\psi \in \mathcal{C}^0(J, \mathbb{R}^n)$ étant $\mathcal{C}^k$-équivalents, il existe un $\mathcal{C}^k$-difféomorphisme θ de I sur J tel que $\varphi = \psi \circ \theta$. Supposons que $M_0 = \varphi(t_0)$ est un point simple de l'arc paramétré φ et qu'il existe deux paramètres s_1 et s_2 dans J tel que $M_0 = \psi(s_1) = \psi(s_2)$. Puisque $\psi = \varphi \circ \theta^{-1}$, on a alors $\varphi(\theta^{-1}(s_1)) = \varphi(\theta^{-1}(s_2)) = M_0$. Le point M_0 étant un point simple de l'arc φ on a nécessairement $\theta^{-1}(s_1) = \theta^{-1}(s_2)$. Cela implique que $s_1 = s_2$ car θ^{-1} est une bijection. On peut donc affirmer qu'il existe un unique paramètre $s_1 \in J$ tel que $M_0 = \psi(s_1)$. Cela signifie que M_0 est un point simple de l'arc paramétré ψ.

2 - Un arc paramétré φ est simple si tout les points de la courbe paramétrée qui est son image C sont des points simples. On vient d'établir que si un point de C est un point simple de l'arc paramétré φ, il est aussi un point simple de tous les arcs paramétrés ψ qui sont $\mathcal{C}^k$-équivalents à φ. Comme deux arcs paramétrés $\mathcal{C}^k$-équivalents ont la même image C, les points de C sont également des points simples de ψ.

On peut également avoir recours à la définition 11.2 pour établir que si un arc paramétré est simple alors tous les arcs paramétrés qui lui sont $\mathcal{C}^k$-équivalents sont simples en remarquant que si φ est injective alors $\psi = \varphi \circ \theta^{-1}$ est également injective car θ^{-1} est une bijection.

---

[112] Voir la définition 10.15 p. 664.

[113] Voir la proposition 14.1 p. 623 du *Cours de première année*.

Solution de l'exercice 11.4

1 - Les dérivées successives de φ en 0 sont :

$$\varphi'(0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \varphi''(0) = \begin{pmatrix} -1 \\ 0 \end{pmatrix} \quad \text{et} \quad \varphi'''(0) = \begin{pmatrix} 0 \\ 6 \end{pmatrix}.$$

Avec les notations de la définition 11.13, on a donc $p = 2$. Le point $M_0 = \varphi(0) = (1,0)$ est donc un point singulier et la tangente à la courbe en ce point est la droite passant par ce point et dirigée par le vecteur de composantes $(-1,0)$. Pour déterminer la valeur de l'entier q, déterminons si les vecteurs $\varphi''(0)$ et $\varphi'''(0)$ forment un système libre ou non. Comme[114] le déterminant

$$\det(\varphi''(0), \varphi'''(0)) = \begin{vmatrix} -1 & 0 \\ 0 & 6 \end{vmatrix} = -6$$

est non nul, les vecteurs $\varphi''(0)$ et $\varphi'''(0)$ forment un système libre. On a donc $q = 3$. Le repère local de la courbe paramétrée C au point $M_0 = (1,0)$ est par conséquent $(M_0; \varphi''(0), \varphi'''(0))$.

2 - Intéressons-nous maintenant au point $M_2 = \varphi(2)$ correspondant au paramètre $t_0 = 2$. Les dérivées successives de φ en 2 sont :

$$\varphi'(2) = \begin{pmatrix} -\sin(2) \\ 12 \end{pmatrix}, \quad \varphi''(2) = \begin{pmatrix} -\cos(2) \\ 12 \end{pmatrix}.$$

On a donc $p = 1$ et $q = 2$ et le point $M_2 = (\cos(2), 8)$ est un point birégulier. Le repère local de la courbe paramétrée C au point $M_2 = (\cos(2), 8)$ est $(M_2; \varphi'(2), \varphi''(2))$.

Solution de l'exercice 11.5

1 - a) Puisque $M_0 = \varphi(t_0)$ où $t_0 \in I$ un point régulier de la courbe paramétrée C associée à φ, on a $\varphi'(t_0) \neq 0$ et $p = 1$. Par ailleurs, comme M_0 est un point d'inflexion de la courbe, q est impair. On en déduit que, s'il existe, l'entier q ne peut être égal à 2 (on a $q \geqslant 3$). Dire que $q \neq 2$, revient à dire, d'après la définition de q, que le système $\{\varphi'(t_0), \varphi''(t_0)\}$ est lié.

b) Puisque le système $\{\varphi'(t_0), \varphi''(t_0)\}$ est lié, il existe[115] un réel λ non nul tel que $\varphi''(t_0) = \lambda \varphi''(t_0)$, *i.e.* tel que

$$\begin{cases} \varphi_1''(t_0) &= \lambda \varphi_1'(t_0) \\ \varphi_2''(t_0) &= \lambda \varphi_2'(t_0) \end{cases}.$$

Le membre de droite de la première équation multiplié par le membre de gauche de la seconde est égal au membre de gauche de la première équation multiplié par le membre de droite de la seconde. Comme $\lambda \neq 0$, cette égalité s'écrit

$$\varphi_1'(t_0)\, \varphi_2''(t_0) = \varphi_1''(t_0)\, \varphi_2'(t_0).$$

---

[114] Voir le *Cours de première année* p. 465.

[115] Voir la définition 8.8 p. 317 du *Cours de première année*.

L'égalité demandée est établie.

Si $M_0 = \varphi(t_0)$ où $t_0 \in I$ un point qui est à la fois point régulier et point d'inflexion de la courbe paramétrée C associée à φ, alors t_0 vérifie $\varphi_1'(t_0)\,\varphi_2''(t_0) - \varphi_1''(t_0)\,\varphi_2'(t_0) = 0$. Mais tous les réels t solution de l'équation

$$\varphi_1'(t)\,\varphi_2''(t) - \varphi_1''(t)\,\varphi_2'(t) = 0$$

ne fournissent pas des points réguliers qui sont points d'inflexion de la courbe paramétrée. On a en effet obtenu cette relation par une succession d'implications, non d'équivalences. Par contre, un point régulier $M = \varphi(t)$ qui est point d'inflexion de la courbe paramétrée correspond à une valeur t du paramètre qui vérifie cette relation.

2 - Un point régulier $M = \varphi(t)$ qui est point d'inflexion de la courbe paramétrée correspond à une valeur t du paramètre qui vérifie

$$\varphi_1'(t)\,\varphi_2''(t) - \varphi_1''(t)\,\varphi_2'(t) = 0. \tag{25}$$

La dérivée de φ est $\varphi' : t \in [0,\pi] \mapsto (-\sin(2t),\ \cos^2(t)(3 + 4\cos^2(t)))$ et sa dérivée seconde est $\varphi'' : t \in [0,\pi] \mapsto (-2\cos(2t),\ \sin(2t)\,(3 - 8\cos^2(t)))$. On en déduit que :

$$\varphi_1'(t)\,\varphi_2''(t) - \varphi_1''(t)\,\varphi_2'(t) = -\sin^2(2t)\,(3 - 8\cos^2(t)) + 2\cos(2t)\cos^2(t)(3 + 4\cos^2(t)).$$

En utilisant les égalités trigonométriques $\cos(2t) = 2\cos^2(t) - 1$ et $\sin^2(2t) = 2\cos^2(t)\sin^2(t) = 2(\cos^2(t) - \cos^4(t))$, on obtient après simplification :

$$\varphi_1'(t)\,\varphi_2''(t) - \varphi_1''(t)\,\varphi_2'(t) = -2\cos^2(t)\,(8\cos^4(t) - 12\cos^2(t) + 3).$$

Les réels t solutions de l'équation (25) sont les réels t pour lesquels $\cos^2(t) = 0$ ou $8\cos^4(t) - 12\cos^2(t) + 3 = 0$. Dans l'intervalle $[0,\pi]$ qui est l'ensemble de définition de l'arc paramétré φ, les solutions sont $t = \frac{\pi}{2}$ ou t tel que $\cos^2(t)$ est racine du polynôme $8X^2 - 12X + 3$. Ce polynôme admet deux racines réelles qui sont $\lambda_1 = \frac{3-\sqrt{3}}{4}$ et $\lambda_2 = \frac{3+\sqrt{3}}{4}$. Seule λ_1 appartient à $[0,1]$ (on a $\lambda_2 > 1$). Les valeurs de $t \in [0,\pi]$ pour lesquelles $\cos^2(t) = \lambda_1$ (i.e. $\cos(t) = \sqrt{\lambda_1}$ ou $\cos(t) = -\sqrt{\lambda_1}$) sont

$$t = \arccos\left(\sqrt{\lambda_1}\right) \quad \text{et} \quad t = \arccos\left(-\sqrt{\lambda_1}\right) = \pi - \arccos\left(\sqrt{\lambda_1}\right).$$

Les points qui sont susceptibles d'être des points réguliers correspondant à des points d'inflexion correspondent donc aux valeurs suivantes du paramètre t :

$$\frac{\pi}{2},\ \pi - \arccos\left(\frac{\sqrt{3-\sqrt{3}}}{2}\right),\ \arccos\left(\frac{\sqrt{3-\sqrt{3}}}{2}\right).$$

Comme φ' s'annule en $\frac{\pi}{2}$, ce point est à rejeter (nous avons établi p. 763 qu'il s'agit d'un point de rebroussement de première espèce).

Pour être complet, il reste à vérifier que les points correspondant aux valeurs

$$t_1 = \pi - \arccos\left(\frac{\sqrt{3-\sqrt{3}}}{2}\right) \quad \text{et} \quad t_2 = \arccos\left(\frac{\sqrt{3-\sqrt{3}}}{2}\right)$$

du paramètre, sont effectivement des points d'inflexion (ce sont des points réguliers). On peut remarquer que

$$\begin{aligned} & \varphi_1(\pi - t) = \cos^2(\pi - t) = \cos^2(t) = \varphi_1(t) \\ \text{et} \quad & \varphi_2(\pi - t) = \cos^3(\pi - t)\sin(\pi - t) = -\cos^3(t)\sin(t) = -\varphi_2(t) \end{aligned}$$

ce qui indique que la courbe admet l'axe des abscisses comme axe de symétrie et que les points $M_1 = \varphi(t_1)$ et $M_2 = \varphi(t_2)$ sont symétriques par rapport à cet axe. On peut donc se contenter de vérifier que $M_2 = \varphi(t_2)$ est un point d'inflexion.

La vérification par un calcul « à la main » que le système $\{\varphi'(t_2), \varphi'''(t_2)\}$ est libre étant un peu fastidieuse, on peut avoir recours à MAPLE[116].

```
> phi:= t-> [cos(t)^2,cos(t)^3*sin(t)]:
> t2:=arccos(1/2*(3-3^(1/2))^(1/2)):
> with(linalg):
> B := transpose(matrix(2,2,[Dn(phi,1,t2),Dn(phi,3,t2)]));
> det(B); evalf(det(B))
```

$$-3\sqrt{3-\sqrt{3}}\sqrt{1+\sqrt{3}} + 3\sqrt{3-\sqrt{3}}\sqrt{1+\sqrt{3}}\sqrt{3}$$

$$-4.087500231$$

Comme le déterminant $\det(\varphi'(t_2), \varphi'''(t_2))$ est non nul, les deux vecteurs $\varphi'(t_2)$ et $\varphi'''(t_2)$ forment une famille libre. On a donc $p = 1$ et $q = 3$, ce qui indique que le point $M_2 = \varphi(t_2)$ est un point d'inflexion (et il en est de même du point $M_1 = \varphi(t_1)$).

La quartique piriforme, ainsi que les deux points d'inflexion M_1 et M_2, sont représentés p. 749.

Solution de l'exercice 11.6

L'application φ est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$. Notons $x : t \in \mathbb{R} \mapsto 2\sin(t)$ et $y : t \in \mathbb{R} \mapsto \cos\left(\frac{3}{2}t\right)$ les deux fonctions coordonnées de $\varphi(t)$.

[116] La procédure `Dn` utilisée ici est donnée p. 760.

▷ *Réduction de l'intervalle d'étude.*
▷ Puisque la première fonction coordonnée de φ est 2π-périodique et que la seconde fonction coordonnée de φ est $4\pi/3$-périodique, l'application φ est 4π-périodique[117] et la courbe C est donc la même si on restreint l'arc paramétré φ à l'intervalle $[-2\pi, 2\pi[$. Par ailleurs, puisque pour tout $t \in [-2\pi, 2\pi[$ on a $x(t) \in [-2, 2]$ et $y(t) \in [-1, 1]$, on en déduit que la courbe C est comprise dans le rectangle $[-2, 2] \times [-1, 1]$.

▷ Notons que pour tout $t \in [-2\pi, 2\pi]$ on a

$$-t \in [-2\pi, 2\pi] \quad \text{et} \quad x(-t) = -x(t), \qquad y(-t) = y(t).$$

Cela signifie que si un point M de coordonnées (x, y) appartient à la courbe alors le point M' de coordonnées $(-x, y)$ appartient aussi à la courbe. La courbe C est symétrique par rapport à l'axe des ordonnées et il suffit donc d'étudier l'arc paramétré φ en restriction à $[0, 2\pi]$. Pour obtenir la représentation graphique de C, on effectuera la symétrie de la courbe correspondant à $[0, 2\pi]$ par rapport à l'axe des ordonnées.

▷ De la même façon, on remarque que pour tout $t \in [0, 2\pi]$ on a

$$2\pi - t \in [0, 2\pi] \quad \text{et} \quad x(2\pi - t) = -x(t), \qquad y(2\pi - t) = -y(t).$$

Cela signifie que si un point M de coordonnées (x, y) appartient à la courbe alors le point M'' de coordonnées $(-x, -y)$ appartient aussi à la courbe. La courbe C est donc symétrique par rapport à l'origine et il suffit d'étudier l'arc paramétré φ en restriction à $[0, \pi]$. Pour obtenir la représentation graphique de C, on effectuera la symétrie de la courbe C' correspondant à $[0, \pi]$ par rapport à l'origine puis par rapport à l'axe des ordonnées.

▷ *Étude des variations des applications coordonnées.*
Les variations des applications x et y en restriction à $[0, \pi]$ s'étudient de la manière habituelle. On a pour tout $t \in [0, \pi]$,

$$x'(t) = 2\cos(t) \quad \text{et} \quad y(t) = -\tfrac{3}{2}\sin(\tfrac{3}{2}t).$$

Sur l'intervalle $[0, \pi]$, l'application x' ne s'annule que pour $t = \frac{1}{2}\pi$ et l'application y' ne s'annule que pour $t = \frac{2}{3}\pi$ On obtient donc le tableau de variation suivant pour x et y :

[117] La plus petite période de φ est la plus petite période commune aux deux fonctions coordonnées de φ.

t	0		$\frac{1}{2}\pi$		$\frac{2}{3}\pi$		π
$x'(t)$	2	$+$	0	$-$	-1	$-$	-2
$x(t)$	0	↗	2	↘	$\sqrt{3}$	↘	0
$y'(t)$	0	$-$	$-\frac{3}{4}\sqrt{2}$	$-$	0	$+$	$\frac{3}{2}$
$y(t)$	1	↘	$-1/\sqrt{2}$	↘	-1	↗	0

▷ *Recherche des points singuliers et des points multiples.*
Le tableau de variation nous indique qu'il n'existe aucun point singulier.

D'après la définition 11.10, un point $M \in \mathbb{R}^2$ est un point multiple de la courbe C' s'il existe deux valeurs différentes t_1 et t_2 du paramètre dans l'intervalle $[0, \pi]$ telles que $M = \varphi(t_1) = \varphi(t_2)$. Pour obtenir tous les points multiples de la courbe, résolvons le système (formé de deux équations non linéaires) suivant pour $(t_1, t_2) \in [0, \pi]^2$:

$$\begin{cases} 2\sin(t_1) &= 2\sin(t_2) \\ \cos(\frac{3}{2}t_1) &= \cos(\frac{3}{2}t_2) \end{cases}.$$

La première équation indique que l'on a $t_1 = t_2$ ou $t_1 = \pi - t_2$. Intéressons-nous au cas où $t_1 = \pi - t_2$, seul cas susceptible de fournir des points multiples. En combinant la relation $t_1 = \pi - t_2$ avec la seconde équation, on obtient que t_2 vérifie $\cos(\frac{3}{2}t_2) = \cos(\frac{3}{2}\pi - \frac{3}{2}t_2)$. Or, deux angles ont le même cosinus si, modulo 2π, ils sont égaux ou opposés. On a donc

$$\tfrac{3}{2}\pi - \tfrac{3}{2}t_2 = \tfrac{3}{2}t_2 + 2k\pi \quad \text{ou} \quad \tfrac{3}{2}\pi - \tfrac{3}{2}t_2 = -\tfrac{3}{2}t_2 + 2k\pi \qquad (k \in \mathbb{Z}).$$

La deuxième possibilité ne fournit pas de solution. La première indique que $t_2 = \frac{\pi}{2} - \frac{2}{3}k\pi$. La seule solution de cette forme appartenant à $[0, \pi]$ est $t_2 = \frac{\pi}{2}$. Dans ce cas, on a $t_1 = \pi - t_2 = \frac{\pi}{2} = t_2$, donc il ne s'agit pas d'un point double. En conclusion, la courbe C' ne possède pas de point double.

▷ *Étude des points d'arrêt ou des branches infinies.*
L'arc paramétré φ étant périodique, la courbe paramétrée C ne possède ni point d'arrêt, ni branche infinie. Toutefois, la courbe C' possède pour points d'arrêt le point M_1 de coordonnées $\varphi(0) = (0, 1)$ en lequel la courbe possède une tangente horizontale et l'origine correspondant à $\varphi(\pi)$ en lequel la tangente est dirigée par $\varphi'(\pi) = (-2, 3/2)$.

$\trianglerighteq$ *Construction de la courbe.*
Pour dessiner la courbe paramétrée associée à l'arc paramétré φ on exploite les renseignements fournis par le tableau de variation. On peut commencer à tracer la courbe à partir du point $M_1 = \varphi(0)$ de coordonnées $(0,1)$. Comme $\varphi'(0) = (2,0)$, la courbe possède en ce point une tangente horizontale. Le tableau indique que la première coordonnée va croître et la seconde décroître : la courbe se dirige donc vers le bas à droite. Cette allure se poursuit jusqu'au point $M_2 = \varphi(\frac{\pi}{2})$ de coordonnées $(2, -1/\sqrt{2})$. En ce point, comme $\varphi'(\frac{1}{2}\pi) = (0, -\frac{3}{4}\sqrt{2})$ la tangente à la courbe est verticale. Le tableau indique que les deux coordonnées vont décroître : la courbe se dirige donc vers le bas à gauche. Cette allure se poursuit jusqu'au point $M_3 = \varphi(\frac{2}{3}\pi)$ de coordonnées $(\sqrt{3}, -1)$. En ce point, comme $\varphi'(\frac{2}{3}\pi) = (-1,0)$ la tangente à la courbe est horizontale. Le tableau indique que la première coordonnée va décroître et la seconde croître : la courbe se dirige donc vers le haut à gauche. On atteint finalement le point $\varphi(\pi)$ qui est l'origine du repère. En ce point le vecteur tangent est $(-2, 3/2)$.

Il est également utile de déterminer les points d'intersection de la courbe avec les axes de coordonnées. La courbe intersecte l'axe des abscisses en des points de coordonnées $\varphi(t)$ où $t \in [0,\pi]$ est solution de $y(t) = 0$. Les solutions sont $t = \frac{1}{3}\pi$ et $t = \pi$. Cela nous donne deux points d'intersection : le point de coordonnées $M_5 = (\sqrt{3}, 0)$ et l'origine. La courbe intersecte l'axe des ordonnées en des points de coordonnées $\varphi(t)$ où $t \in [0,\pi]$ est solution de $x(t) = 0$. Les solutions sont $t = 0$ et $t = \pi$. Cela nous donne deux points d'intersection M_1 et l'origine.

La courbe est présentée sur la figure 26.

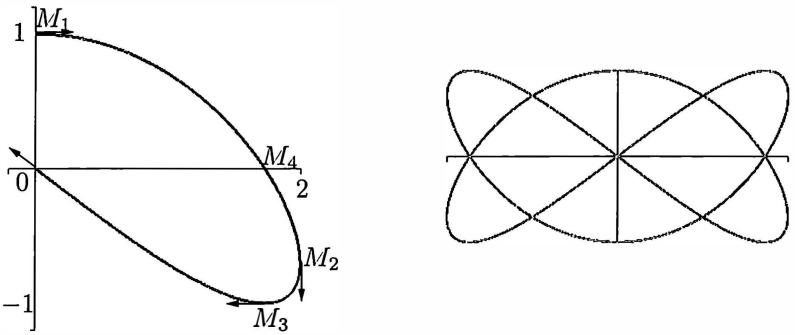


Fig. 26 La courbe paramétrée en restriction à l'intervalle $[0,\pi]$ (à gauche) et la courbe complète (à droite).

Solution de l'exercice 11.7

Un point $M(x,y,z)$ appartient à la surface S si et seulement si il est l'image par une rotation d'axe $(O;\mathbf{k})$ d'un point issu du segment $[AB]$. Désignons par r_θ la rotation d'angle $\theta \in [0, 2\pi[$ autour de l'axe $(O;\mathbf{k})$. Elle est représentée par la matrice

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

On a

$$\begin{aligned} M \in S &\iff \exists \theta \in [0, 2\pi[\quad \exists \lambda \in [0,1] \quad OM = r_\theta(\lambda OA + (1-\lambda)OB) \\ &\iff \exists \theta \in [0, 2\pi[\quad \exists \lambda \in [0,1] \quad \begin{cases} x = \lambda\cos(\theta) - (1-\lambda)\sin(\theta) \\ y = \lambda\sin(\theta) + (1-\lambda)\cos(\theta) \\ z = \lambda \end{cases} \end{aligned}$$

La surface S admet donc pour paramétrage

$$\varphi : (\theta, \lambda) \in [0, 2\pi[\times[0,1] \longmapsto \begin{pmatrix} \lambda\cos(\theta) - (1-\lambda)\sin(\theta) \\ \lambda\sin(\theta) + (1-\lambda)\cos(\theta) \\ \lambda \end{pmatrix}.$$

Solution de l'exercice 11.8

Considérons une nappe paramétrée simple $\varphi \in \mathcal{C}^0(U, \mathbb{R}^n)$ où U désigne un domaine (i.e. un ouvert connexe) de $\mathbb{R}^2$. Pour montrer que l'arc paramétré φ est orientable, il faut montrer qu'il existe exactement deux classes d'équivalence pour la relation $\mathcal{R}_{\mathcal{O}}$: « être $\mathcal{O}^1$-équivalent » parmi les nappes paramétrées $\mathcal{C}^1$-équivalentes à la nappe paramétrée φ.

Considérons la nappe paramétrée

$$\psi : (u, v) \in V \mapsto \varphi(-u, v) \quad \text{où} \quad V = \{(u, v) \in \mathbb{R}^2 \ ; \ (-u, v) \in U\}.$$

Cette nappe paramétrée est $\mathcal{C}^1$-équivalente à la nappe φ car on a $\varphi = \psi \circ \theta$ où $\theta : t \in U \mapsto (-u, v) \in V$ est clairement un $\mathcal{C}^1$-difféomorphisme. Le jacobien de θ en $(u, v) \in U$ vaut

$$\mathcal{J}_{\widetilde{\theta}}(u, v) = \begin{vmatrix} \partial_u\theta_1(u, v) & \partial_v\theta_1(u, v) \\ \partial_u\theta_2(u, v) & \partial_v\theta_2(u, v) \end{vmatrix} = \begin{vmatrix} -1 & 0 \\ 0 & 1 \end{vmatrix} = -1.$$

Pour montrer que la nappe paramétrée φ est orientable, nous allons montrer que la nappe paramétrée ψ n'appartient pas à la même classe d'équivalence que la nappe paramétrée φ pour la relation $\mathcal{R}_{\mathcal{O}}$. Cela établira l'existence de deux classes d'équivalence distinctes pour la relation $\mathcal{R}_{\mathcal{O}}$ et par conséquent le fait que la nappe paramétrée φ est orientable.

Raisonnons par l'absurde et supposons que les nappes paramétrées φ et ψ sont $\mathcal{O}^1$-équivalents : il existe un difféomorphisme $\widetilde{\theta}$ de V sur U à jacobien positif tel que $\psi = \varphi \circ \widetilde{\theta}$. Comme on a aussi $\varphi = \psi \circ \theta$, on en déduit que $\varphi = \varphi \circ \widetilde{\theta} \circ \theta$, c'est-à-dire que

$$\forall (u, v) \in U \quad \varphi(u, v) = \varphi(\widetilde{\theta}(\theta(u, v))).$$

Puisque l'application φ est injective, cette assertion implique que[118] pour tout $(u,v) \in U$ on a $(u,v) = \widetilde{\theta}(\theta(u,v))$, c'est-à-dire que $\widetilde{\theta} \circ \theta = \mathrm{id}$. Cette égalité implique que le jacobien en $(u,v) \in U$ de l'application $\widetilde{\theta} \circ \theta$ est égal à celui de l'identité (qui vaut $+1$). D'après la proposition 10.10 p. 638, la matrice jacobienne de $\widetilde{\theta} \circ \theta$ en $(u,v) \in U$ vérifie

$$\mathrm{J}_{\widetilde{\theta}\circ\theta}(u,v) = \mathrm{J}_{\widetilde{\theta}}(\theta(u,v)) \times \mathrm{J}_{\theta}(u,v)$$

ce qui implique que le jacobien en $(u,v) \in U$ de $\widetilde{\theta} \circ \theta$ vaut

$$\mathcal{J}_{\widetilde{\theta}}(-u,v) \times \underbrace{\mathcal{J}_{\theta}(u,v)}_{=-1} = -\mathcal{J}_{\widetilde{\theta}}(-u,v).$$

On aboutit à la contradiction suivante : d'une part $\widetilde{\theta}$ est à jacobien positif et d'autre part $-\mathcal{J}_{\widetilde{\theta}}(-u,v) = 1$ pour tout $(u,v) \in U$. On en conclut que les nappes paramétrées φ et ψ ne sont pas $\mathcal{O}^1$-équivalents, ce qui signifie qu'il existe deux classes d'équivalence distinctes pour la relation $\mathcal{R}_{\mathcal{O}}$ et par conséquent le fait que la nappe paramétrée φ est orientable.

Solution de l'exercice 11.9

1 - Pour montrer que la nappe φ est régulière, il suffit d'établir que quel que soit $(u,v) \in [0,2\pi[\times[0,1]$, le vecteur $\mathbf{N}(u,v) = \partial_u\varphi(u,v) \wedge \partial_v\varphi(u,v)$ n'est pas nul. On a

$$\begin{aligned}\mathbf{N}(u,v) &= \begin{pmatrix} -v\sin(u) - (1-v)\cos(u) \\ v\cos(u) - (1-v)\sin(u) \\ 0 \end{pmatrix} \wedge \begin{pmatrix} \cos(u) + \sin(u) \\ \sin(u) - \cos(u) \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} v\cos(u) - (1-v)\sin(u) \\ v\sin(u) + (1-v)\cos(u) \\ 1-2v \end{pmatrix}.\end{aligned}$$

On en déduit que $\|\mathbf{N}(u,v)\|^2 = v^2 + (1-v)^2 + (1-2v)^2$. Cette quantité est une somme de 3 carrés. Elle est nulle si et seulement les trois termes de la somme sont nuls simultanément, ce qui n'est pas possible. On en conclut que tous les points de S sont réguliers.

2 - Un point $P(x,y,z)$ appartient au plan tangent à la surface S en un point $M_0 = \varphi(u_0,v_0)$ si et seulement si le produit scalaire $\mathbf{N}(u_0,v_0) \cdot \overrightarrow{M_0P}$ est nul. Cette condition s'écrit

$$\begin{pmatrix} v_0\cos(u_0) - (1-v_0)\sin(u_0) \\ v_0\sin(u_0) + (1-v_0)\cos(u_0) \\ 1-2v_0 \end{pmatrix} \cdot \begin{pmatrix} x - v_0\cos(u_0) + (1-v_0)\sin(u_0) \\ y - v_0\sin(u_0) - (1-v_0)\cos(u_0) \\ z - v_0 \end{pmatrix} = 0.$$

[118] Rappelons, voir la définition 2.9 p. 41 du *Cours de première année*, que φ est injective signifie que $\forall (x,x') \in U^2 \quad (\varphi(x) = \varphi(x') \Rightarrow x = x')$.

En calculant ce produit scalaire, on obtient après simplification que l'équation du plan tangent en $M_0 = \varphi(u_0, v_0)$ est :

$$\big(v_0\cos(u_0) - (1-v_0)\sin(u_0)\big)x + \big(v_0\sin(u_0) + (1-v_0)\cos(u_0)\big)y + \big(1-2v_0\big)z = v_0 - 1.$$

Solution de l'exercice 11.10

La courbe d'équation $x^2+y^4-3xy+x=1$ est une courbe implicite de la forme $F(x,y)=0$ où $F:(x,y)\in\mathbb{R}^2 \mapsto x^2+y^4-3xy+x-1$. Le point de coordonnées $(2,1)$ est bien un point de la courbe car $F(2,1)=0$. L'application F est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}^2$ et pour tout $(x,y)\in\mathbb{R}^2$ on a

$$\nabla F(x,y) = \begin{pmatrix}\partial_x F(x,y)\\ \partial_y F(x,y)\end{pmatrix} = \begin{pmatrix}2x-3y+1\\ 4y^3-3x\end{pmatrix}.$$

On en déduit que $\nabla F(2,1) = (2,-2) \neq 0$ et par conséquent que le point de coordonnées $(2,1)$ est un point régulier de la courbe implicite du plan d'équation $x^2+y^4-3xy+x=1$.

Solution de l'exercice 11.11

1 - L'application F est de classe $\mathcal{C}^\infty$ et puisque $\partial_y F(0,0) = -2 \neq 0$, d'après le théorème des fonctions implicites, il existe un intervalle ouvert I contenant 0, un intervalle ouvert J contenant 0 et une unique application $f: I \to J$ de classe $\mathcal{C}^\infty$ tels que au voisinage de l'origine on ait

$$F(x,\psi(x)) = 0 \quad\Longleftrightarrow\quad y = f(x).$$

2 - D'après le théorème des fonctions implicites on a $f(0)=0$ et en utilisant la relation (17) p. 812, on obtient $f'(0) = 3/2$. Pour déterminer $f''(0)$, dérivons deux fois les deux membres de l'égalité $F(x,f(x))=0$:

$$\frac{\partial^2 F}{\partial x^2}(x,f(x)) + 2f'(x)\frac{\partial^2 F}{\partial x\partial y}(x,f(x)) + f''(x)\frac{\partial F}{\partial y}(x,f(x)) + f'(x)^2\frac{\partial^2 F}{\partial y^2}(x,f(x)) = 0. \tag{26}$$

En prenant $x=0$ dans cette égalité, on obtient $f''(0)=0$. Le développement limité à l'ordre 2 de la fonction f au voisinage de 0 est donc : $f(x) = \frac{3}{2}x + o_0(x^2)$.

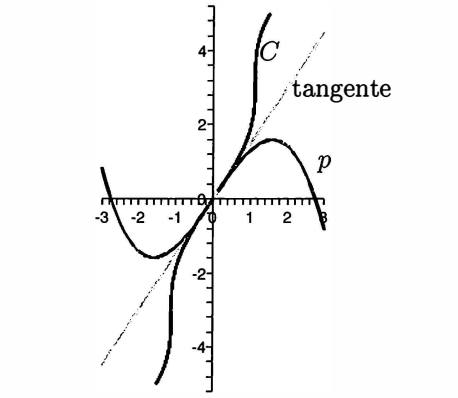
3 - On en déduit que la tangente à la courbe au point $(0,0)$ a pour équation $y = 3x/2$. Si on souhaite connaître la position de la courbe par rapport à sa tangente, on peut obtenir $f'''(0)$ en dérivant à nouveau les deux membres de la relation (26). On obtient $f'''(0) = -19/16$ ce qui montre qu'au voisinage de 0, l'application f admet pour développement limité

$$f(x) = \frac{3}{2}x - \frac{19}{96}x^3 + o_0(x^3).$$

Ainsi, la courbe traverse sa tangente à l'origine. De plus, au voisinage de l'origine, elle est « proche » du graphe de la fonction polynomiale $p : x \mapsto \frac{3}{2}x - \frac{19}{96}x^3$.

Visualisons la courbe et sa tangente à l'aide de MAPLE au voisinage de l'origine :

```
> fig1:=implicitplot(sin(x)-sin(y)+2*x-y, x=-3..3,y=-5..5):
> fig2:=plot([3*x/2 - 19*x^3/96, 3*x/2], x=-3..3):
> display({fig1,fig2},scaling=constrained);
```



SOLUTION DE L'EXERCICE 11.12

1 - Puisque $F(0,0,1) = 0$, une solution de $(\mathcal{E}_{0,0})$ est $x_3 = 1$. On pourra vérifier que l'équation $(\mathcal{E}_{0,0})$ admet deux autres solutions : $\frac{-1+\sqrt{5}}{2}$ et $\frac{-1-\sqrt{5}}{2}$.

2 - Appliquons le théorème des fonctions implicites (théorème 11.4 p. 821) ; l'application F est de classe $\mathcal{C}^\infty$ sur l'ouvert $\mathbb{R}^3$, $F(0,0,1) = 0$ et on a

$$\det\left(\frac{\partial F}{\partial x_3}(0,0,1)\right) = -2 \neq 0.$$

Par conséquent il existe une unique application f définie sur un voisinage V de $(0,0)$ à valeurs dans un intervalle I de $\mathbb{R}$ contenant 1 telle que pour tout $((x_1,x_2),x_3) \in V \times I$ on ait

$$F(x_1,x_2,x_3) = 0 \iff x_3 = f(x_1,x_2).$$

Autrement dit, pour tout (x_1,x_2) proche de $(0,0)$ (c'est-à-dire dans V) il existe un unique réel $x_3 = f(x_1,x_2)$ proche de 1 (c'est-à-dire dans I) solution de l'équation $(\mathcal{E}_{x_1,x_2})$.

3 - Le théorème des fonctions implicites affirme d'une part que f est de classe $\mathcal{C}^\infty$ et satisfait $f(0,0) = 1$ et d'autre part que pour tout $(x_1,x_2) \in V$ on a $F(x_1,x_2,f(x_1,x_2)) = 0$. En dérivant par rapport aux variables x_1 et x_2 cette dernière relation, on obtient

$$\begin{cases} \partial_1 F(x_1,x_2,f(x_1,x_2)) + \dfrac{\partial f}{\partial x_1}(x_1,x_2)\,\partial_3 F(x_1,x_2,f(x_1,x_2)) = 0 \\ \partial_2 F(x_1,x_2,f(x_1,x_2)) + \dfrac{\partial f}{\partial x_2}(x_1,x_2)\,\partial_3 F(x_1,x_2,f(x_1,x_2)) = 0 \end{cases} \tag{27}$$

En prenant $(x_1, x_2) = (0,0)$ et en utilisant le fait que $f(0,0) = 1$ on obtient

$$\frac{\partial f}{\partial x_1}(0,0) = -\frac{\partial_1 F(0,0,1)}{\partial_3 F(0,0,1)} = -1 \quad \text{et} \quad \frac{\partial f}{\partial x_2}(0,0) = -\frac{\partial_2 F(0,0,1)}{\partial_3 F(0,0,1)} = \frac{1}{2}.$$

Pour obtenir les dérivées secondes de f en $(0,0)$ permettant d'établir le développement limité à l'ordre 2 de l'application f, on dérive à nouveau les équations (27) par rapport aux variables x_1 et x_2, puis on évalue le résultat au point $(0,0)$. Par exemple, dérivons la première relation du système (27) par rapport à la variable x_1 ; on obtient

$$\partial^2_{11} F(x_1, x_2, f(x_1,x_2)) + 2\frac{\partial f}{\partial x_1}(x_1,x_2)\ \partial^2_{13} F(x_1,x_2,f(x_1,x_2))$$
$$+ \frac{\partial^2 f}{\partial x_1^2}(x_1,x_2)\ \partial_3 F(x_1,x_2,f(x_1,x_2)) + \left(\frac{\partial f}{\partial x_1}(x_1,x_2)\right)^2 \partial^2_{33} F(x_1,x_2,f(x_1,x_2)) = 0.$$

En prenant $(x_1, x_2) = (0,0)$, on établit à partir des valeurs des dérivées partielles premières de f en $(0,0)$ obtenues précédemment, que

$$\frac{\partial^2 f}{\partial x_1^2}(0,0) = 4.$$

De même, en dérivant la première équation de (27) par rapport à x_2 puis la seconde équation de (27) par rapport à x_2, on obtient les valeurs des autres dérivées partielles secondes de f en $(0,0)$:

$$\frac{\partial^2 f}{\partial x_1 \partial x_2}(0,0) = -\frac{7}{4} \quad \text{et} \quad \frac{\partial^2 f}{\partial x_2^2}(0,0) = \frac{3}{4}.$$

On en conclut que le développement limité à l'ordre 2 de f en $(0,0)$ est :

$$f(x_1,x_2) = 1 - x_1 + \frac{1}{2}x_2 + 2x_1^2 - \frac{7}{8}x_1x_2 + \frac{3}{8}x_2^2 + o_{(0,0)}(\|(x_1,x_2)\|^2).$$

Solution de l'exercice 11.13

Pour montrer que tous les points de la surface implicite S sont réguliers, il faut montrer que le gradient de F en un point (x,y,z) de S n'est jamais nul. Ce gradient vaut

$$\nabla F(x,y,z) = \begin{pmatrix} 3x^2 + (x+y+z+1)^2 \\ 3y^2 + (x+y+z+1)^2 \\ 3z^2 + (x+y+z+1)^2 \end{pmatrix}.$$

Chacune de ses composantes est une somme de carrés. Elle est nulle si et seulement si les deux termes sont nuls. On en déduit que le gradient de F est nul si et seulement si

$$x^2 = 0, \quad y^2 = 0, \quad z^2 = 0 \quad \text{et} \quad (x+y+z+1)^2 = 0.$$

Ces quatre conditions ne pouvant être réalisées simultanément, on en déduit que le gradient de F n'est jamais nul. Tous les points de la surface S sont donc réguliers.

Solution de l'exercice 11.14

1 - Le plan $\mathcal{P}$ passant par le point A de coordonnées (x_0, y_0, z_0) et dont un vecteur normal est le vecteur $\mathbf{e}_1$ de composantes (α, β, γ) a pour équation cartésienne $\alpha(x - x_0) + \beta(y - y_0) + \gamma(z - z_0) = 0$. Elle se réécrit

$$\alpha x + \beta y + \gamma z = \alpha x_0 + \beta y_0 + \gamma z_0.$$

2 - D'après la proposition 11.19 p. 829, un vecteur normal à l'ellipsoïde au point M de coordonnées (x_1, y_1, z_1) est $\mathbf{n} = \nabla F(x_1, y_1, z_1)$ où F est l'application

$$F : (x, y, z) \in \mathbb{R}^3 \longmapsto \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} - 1.$$

On obtient qu'un vecteur normal a pour composantes $(2x_1/a^2,\ 2y_1/b^2\ , 2z_1/c^2)$. Le plan tangent à l'ellipsoïde en M a donc pour équation

$$\frac{2x_1}{a^2}(x - x_1) + \frac{2y_1}{b^2}(y - y_1) + \frac{2z_1}{c^2}(z - z_1) = 0.$$

Étant donné que le point M est un point de l'ellipse $\mathcal{E}$, on a la relation $x_1^2/a^2 + y_1^2/a^2 + z_1^2/a^2 = 1$; l'équation du plan tangent se réécrit donc

$$\frac{x_1}{a^2}x + \frac{y_1}{b^2}y + \frac{z_1}{c^2}z = 1.$$

3 - Notons tout d'abord que deux plans sont égaux si et seulement si ils ont des vecteurs normaux colinéaires et possèdent un point en commun. Un vecteur normal au plan $\mathcal{P}_1$ a pour composantes (a_1, b_1, c_1) et un vecteur normal au plan $\mathcal{P}_2$ a pour composantes (a_2, b_2, c_2). Ces deux vecteurs sont colinéaires si et seulement si[119] $a_1/a_2 = b_1/b_2 = c_1/c_2$. D'autre part, le point de coordonnées $(d_2/a_2, 0, 0)$ appartient au plan $\mathcal{P}_2$. Ce point appartient au plan $\mathcal{P}_1$ si et seulement si $d_2\frac{a_1}{a_2} = d_1$. Autrement dit on a établi que

$$\mathcal{P}_1 = \mathcal{P}_2 \iff \left(\frac{a_1}{a_2} = \frac{b_1}{b_2} = \frac{c_1}{c_2} \quad \text{et} \quad \frac{a_1}{a_2} = \frac{d_1}{d_2}\right),$$

ce qui répond à la question.

4 - Le plan $\mathcal{P}$ est tangent à l'ellipse $\mathcal{E}$ si et seulement si il existe $M \in \mathcal{E}$ tel que $\mathcal{P}$ coïncide avec le plan tangent à $\mathcal{E}$ en M. On déduit des questions précédentes que le plan $\mathcal{P}$ est tangent à l'ellipse $\mathcal{E}$ si et seulement si il existe $(x_1, y_1, z_1) \in \mathbb{R}^3$ tel que

$$\frac{x_1^2}{a^2} + \frac{y_1^2}{b^2} + \frac{z_1^2}{c^2} = 1 \quad \text{et} \quad \frac{a^2\alpha}{x_1} = \frac{b^2\beta}{y_1} = \frac{c^2\gamma}{z_1} = \alpha x_0 + \beta y_0 + \gamma z_0.$$

[119] On utilise ici que les trois réels a_2, b_2 et c_2 sont non nuls.

Ces relations permettent d'exprimer x_1, y_1 et z_1 en fonction des réels a, b, c, x_0, y_0 et z_0 ; on a

$$x_1 = \frac{a^2\alpha}{\alpha x_0 + \beta y_0 + \gamma z_0},\ y_1 = \frac{b^2\beta}{\alpha x_0 + \beta y_0 + \gamma z_0} \text{ et } z_1 = \frac{c^2\gamma}{\alpha x_0 + \beta y_0 + \gamma z_0}.$$

Comme $\frac{x_1^2}{a^2} + \frac{y_1^2}{b^2} + \frac{z_1^2}{c^2} = 1$, on en déduit que le plan $\mathcal{P}$ est tangent à l'ellipse $\mathcal{E}$ si et seulement si

$$a^2\alpha^2 + b^2\beta^2 + c^2\gamma^2 = (\alpha x_0 + \beta y_0 + \gamma z_0)^2.$$

Le vecteur $\mathbf{e}_1$, normal au plan $\mathcal{P}$, a pour coordonnées (α, β, γ). On peut donc réécrire cette condition sous la forme $q(\mathbf{e}_1) = 0$ où q est la forme quadratique sur $\mathbb{R}^3$ définie par

$$q : (X, Y, Z) \in \mathbb{R}^3 \longmapsto a^2X^2 + b^2Y^2 + c^2Z^2 - (Xx_0 + Yy_0 + Zz_0)^2.$$

5 - Si on fait passer par le point A trois plans tangents à l'ellipsoïde $\mathcal{E}$ qui de plus sont perpendiculaires deux à deux alors leur trois vecteurs normaux $\mathbf{n}_1$, $\mathbf{n}_2$ et $\mathbf{n}_3$ forment une base orthogonale de $\mathbb{R}^3$ (car les plans sont au nombre de trois et sont perpendiculaires entre eux). De plus, ces vecteurs vérifient, d'après la question 4, la relation $q(\mathbf{n}_i) = 0$ pour tout $i \in \{1, 2, 3\}$. En considérant les vecteurs $\mathbf{e}_i = \mathbf{n}_i / \|\mathbf{n}_i\|$ pour $i \in \{1, 2, 3\}$, on obtient une base orthonormée $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ de $\mathbb{R}^3$ pour laquelle

$$\forall i \in \{1, 2, 3\} \quad q(\mathbf{e}_i) = 0.$$

Réciproquement, supposons que trois vecteurs $\mathbf{e}_1$, $\mathbf{e}_2$ et $\mathbf{e}_3$ forment une base orthonormée de $\mathbb{R}^3$ et satisfont $q(\mathbf{e}_i) = 0$, $i \in \{1, 2, 3\}$. D'après la question 4, chaque plan passant par le point A et de vecteur normal $\mathbf{e}_i$ sera tangent à l'ellipsoïde $\mathcal{E}$.

6 - D'après l'exercice 8.11 p. 520 on sait qu'il existe une base $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ de $\mathbb{R}^3$ telle que $q(\mathbf{e}_1) = q(\mathbf{e}_2) = q(\mathbf{e}_3) = 0$ si et seulement si la trace de la forme quadratique q est nulle. La trace de la forme quadratique q obtenue à la question 4 peut s'évaluer en écrivant la matrice de q dans la base canonique de $\mathbb{R}^3$: la trace correspond à la somme des éléments diagonaux, autrement dit la trace de q est la somme des coefficients devant les termes en α^2, β^2 et γ^2 dans l'expression « développée » de q. On a

$$\mathrm{Tr}(q) = (a^2 - x_0^2) + (b^2 - y_0^2) + (c^2 - z_0^2).$$

On en déduit que le point $A(x_0, y_0, z_0)$ est un point de $\mathbb{R}^3$ par lequel on peut faire passer trois plans tangents à $\mathcal{E}$ qui sont perpendiculaires entre eux si et seulement si

$$x_0^2 + y_0^2 + z_0^2 = a^2 + b^2 + c^2.$$

L'ensemble de ces points est une sphère centrée en l'origine et ayant pour rayon $\sqrt{a^2 + b^2 + c^2}$.

Solution de l'exercice 11.15

1 - a) Lorsque $\varphi_0 \in]-\frac{\pi}{2}, \frac{\pi}{2}[$ le parallèle P_{φ_0} est un cercle situé dans le plan d'équation $z = R\sin(\varphi_0)$, de centre $(0, 0, R\sin(\varphi_0))$ et de rayon $R\cos(\varphi_0)$. Lorsque $\varphi_0 = \frac{\pi}{2}$ (resp. $\varphi_0 = -\frac{\pi}{2}$), le parallèle P_{φ_0} est le point de coordonnées $(0, 0, 1)$ (resp. $(0, 0, -1)$).

b) La dérivée de l'application π_{φ_0} pour $\varphi_0 \in]-\frac{\pi}{2}, \frac{\pi}{2}[$ en $\theta \in \mathbb{R}$ est :

$$\pi'_{\varphi_0}(\theta) = \begin{pmatrix} -R\sin(\theta)\cos(\varphi_0) \\ R\cos(\theta)\cos(\varphi_0) \\ 0 \end{pmatrix}.$$

Comme d'une part pour $\varphi_0 \in]-\frac{\pi}{2}, \frac{\pi}{2}[$ on a $\cos(\varphi_0) \neq 0$ et comme d'autre part les fonction sinus et cosinus ne s'annulent jamais simultanément, on peut affirmer que le vecteur $\pi'_{\varphi_0}(\theta)$ n'est jamais nul, quel que soit $\theta \in \mathbb{R}$. D'après la définition 11.12 p. 754, un vecteur tangent au parallèle P_{φ_0} au point de longitude θ est $\pi'_{\varphi_0}(\theta)$. Sa norme euclidienne vaut

$$\|\pi'_{\varphi_0}(\theta)\| = \left(R^2\cos^2(\varphi_0)\left(\sin^2(\theta) + \cos^2(\theta)\right)\right)^{\frac{1}{2}} = R|\cos(\varphi_0)| = R\cos(\varphi_0).$$

2 - a) Le parallèle P_{φ_0} est une courbe qui coïncide avec la courbe C qui est l'image de l'arc paramétré γ associée à l'application f constante égale à φ_0.

b) La dérivée de l'application γ en $\theta \in \mathbb{R}$ est :

$$\gamma'(\theta) = \begin{pmatrix} -R\sin(\theta)\cos(f(\varphi)) - R\,f'(\theta)\,\cos(\theta)\sin(f(\varphi)) \\ R\cos(\theta)\cos(f(\varphi)) - R\,f'(\theta)\,\sin(\theta)\sin(f(\varphi)) \\ R\,f'(\varphi)\,\cos(f(\varphi)) \end{pmatrix}.$$

La norme euclidienne du vecteur $\gamma'(\theta)$ est

$$\|\gamma'(\theta)\| = R\sqrt{\cos(f(\theta))^2 + (f'(\theta))^2}.$$

Comme f est une application à valeurs dans $]-\frac{\pi}{2}, \frac{\pi}{2}[$, quel que soit $\theta \in I$ on a $\cos(f(\theta)) \neq 0$, ce qui implique que $\|\gamma'(\theta)\| \neq 0$ (le terme sous le radical est la somme de deux termes positifs dont l'un au moins est non nul) et donc que $\gamma'(\theta)$ est non nul. D'après la définition 11.12 p. 754, $\gamma'(\theta)$ est un vecteur tangent à la courbe C au point de longitude θ.

3 - a) Les composantes des vecteurs $\tau_C = \gamma'(\theta_0)$ et $\tau_P = \pi'_{\varphi_0}(\theta_0)$ ont été déterminées aux questions précédentes. Le produit scalaire entre le vecteur τ_C et le vecteur τ_P vaut[(120)] :

$$\gamma'(\theta_0) \cdot \pi'_{\varphi_0}(\theta_0) = R^2\cos^2(\varphi_0).$$

Le produit vectoriel entre ces deux vecteurs vaut[120] :

$$\begin{aligned}\gamma'(\theta_0) \wedge \pi'_{\varphi_0}(\theta_0) &= \begin{pmatrix} -R^2 f'(\theta_0) \cos(\theta_0) \cos^2(f(\theta_0)) \\ -R^2 f'(\theta_0) \sin(\theta_0) \cos^2(f(\theta_0)) \\ -R^2 f'(\theta_0) \cos(f(\theta_0)) \sin(f(\theta_0)) \end{pmatrix} \\ &= -R^2 f'(\theta_0) \cos(f(\theta_0)) \begin{pmatrix} \cos(\theta_0) \cos(f(\theta_0)) \\ \sin(\theta_0) \cos(f(\theta_0)) \\ \sin(f(\theta_0)) \end{pmatrix}.\end{aligned}$$

b) En utilisant les relations de calcul vectoriel :

$$\begin{aligned}\tau_C \cdot \tau_P &= \|\tau_C\| \, \|\tau_P\| \cos(\alpha) \\ \|\tau_C \wedge \tau_P\| &= \|\tau_C\| \, \|\tau_P\| \, |\sin(\alpha)|\end{aligned}$$

où $\alpha \in]-\frac{\pi}{2}, \frac{\pi}{2}[$ désigne l'angle formé par le vecteur τ_C et le vecteur τ_P, on en déduit que

$$|\tan(\alpha)| = \frac{|\sin(\alpha)|}{|\cos(\alpha)|} = \frac{\|\tau_C \wedge \tau_P\|}{|\tau_C \cdot \tau_P|}.$$

On a $|\tau_C \cdot \tau_P| = R^2 \cos^2(\varphi_0)$ et $\|\tau_C \wedge \tau_P\| = R^2 |f'(\theta_0)| \, |\cos(f(\theta_0))|$. De plus, comme $f(\theta_0) \in]-\frac{\pi}{2}, \frac{\pi}{2}[$, on a $\cos(f(\theta_0)) > 0$. On en déduit que

$$|\tan(\alpha)| = \frac{|f'(\theta_0)|}{\cos(\varphi(\theta_0))}.$$

4 - a) La fonction tangente est dérivable sur $]-\frac{\pi}{2}, \frac{\pi}{2}[$. Pour $x \in]-\frac{\pi}{2}, \frac{\pi}{2}[$, on a $\frac{x}{2} + \frac{\pi}{4} \in]0, \frac{\pi}{2}[$ et la fonction tangente est dérivable sur $]0, \frac{\pi}{2}[$ à valeurs dans $]0, +\infty[$. Comme la fonction logarithme est dérivable sur $]0, +\infty[$, on en déduit que l'application g est dérivable sur $]-\frac{\pi}{2}, \frac{\pi}{2}[$. Pour $x \in]-\frac{\pi}{2}, \frac{\pi}{2}[$, on obtient en utilisant la formule de dérivation d'une fonction composée :

$$\begin{aligned}g'(x) &= \left(\tan\left(\tfrac{1}{2}x + \tfrac{\pi}{4}\right)\right)' \times \left(\tan\left(\tfrac{1}{2}x + \tfrac{\pi}{4}\right)\right)^{-1} \\ &= \frac{1}{2} \frac{1}{\cos^2\left(\frac{1}{2}x + \frac{\pi}{4}\right)} \times \frac{1}{\tan\left(\frac{1}{2}x + \frac{\pi}{4}\right)} = \frac{1}{2 \sin\left(\frac{1}{2}x + \frac{\pi}{4}\right) \cos\left(\frac{1}{2}x + \frac{\pi}{4}\right)} \\ &= \frac{1}{\sin(x + \frac{\pi}{2})} = \frac{1}{\cos(x)}.\end{aligned}$$

b) Pour $\alpha \in [0, \frac{\pi}{2}[$, on a $\tan(\alpha) \geqslant 0$ et par conséquent $\tan(\alpha) = \frac{f'(\theta)}{\cos(f(\theta))}$, pour tout $\theta \in I$. On en déduit que pour tout $(\theta, \theta_1) \in I^2$

$$\int_{\theta_1}^{\theta} \tan(\alpha) \, \mathrm{d}t = \int_{\theta_1}^{\theta} \frac{f'(t)}{\cos(f(t))} \, \mathrm{d}t.$$

---

[120] On rappelle que le produit scalaire des deux vecteurs $\mathbf{x} = (x_1, x_2, x_3) \in \mathbb{R}^3$ et $\mathbf{y} = (y_1, y_2, y_3) \in \mathbb{R}^3$ a pour expression $\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_2 y_2 + x_2 y_2$ et que le produit vectoriel vaut :

$$\mathbf{x} \wedge \mathbf{y} = \begin{pmatrix} x_2 y_3 - x_3 y_2 \\ x_3 y_1 - x_1 y_3 \\ x_1 y_2 - x_2 y_1 \end{pmatrix}.$$

On a $\int_{\theta_1}^{\theta} \tan(\alpha)\, dt = (\theta - \theta_1)\tan(\alpha)$ et en effectuant le changement de variable $y = f(t)$, on obtient

$$\begin{aligned}\int_{\theta_1}^{\theta} \frac{f'(t)}{\cos(f(t))}\, dt &= \int_{f(\theta_1)}^{f(\theta)} \frac{1}{\cos(y)}\, dy = \Big[\ln(\tan(\tfrac{1}{2}y + \tfrac{\pi}{4}))\Big]_{f(\theta_1)}^{f(\theta)} \\ &= \ln(\tan(\tfrac{1}{2}f(\theta) + \tfrac{\pi}{4})) - \ln(\tan(\tfrac{1}{2}f(\theta_1) + \tfrac{\pi}{4})).\end{aligned}$$

On en conclut que pour tout $\theta \in I$

$$(\theta - \theta_1)\tan(\alpha) = \ln(\tan(\tfrac{1}{2}f(\theta) + \tfrac{\pi}{4})) - \ln(\tan(\tfrac{1}{2}\varphi_1 + \tfrac{\pi}{4})) \qquad (28)$$

où $\varphi_1 = f(\theta_1)$.

5 - a) Si $\alpha > 0$ alors $\lim_{\theta\to+\infty}(\theta - \theta_1)\tan(\alpha) = +\infty$ et la relation (28) indique que $\lim_{\theta\to+\infty} \ln(\tan(\frac{1}{2}f(\theta)+\frac{\pi}{4})) = +\infty$. Comme l'application logarithme est strictement croissante et admet pour limite $+\infty$ en $+\infty$, cela signifie que $\lim_{\theta\to+\infty} \tan(\frac{1}{2}f(\theta) + \frac{\pi}{4}) = +\infty$. Pour tout $\theta \in \mathbb{R}$, on a $f(\theta) \in]-\frac{\pi}{2}, \frac{\pi}{2}[$ et par conséquent $\frac{1}{2}f(\theta) + \frac{\pi}{4} \in]0, \frac{\pi}{2}[$. La fonction tangente étant une application strictement croissante sur $]0, \frac{\pi}{2}[$ de limite $+\infty$ en $\frac{\pi}{2}$, on en déduit que $\lim_{\theta\to+\infty} \varphi(\theta) = \frac{\pi}{2}$.

Par ailleurs, on a $\lim_{\theta\to-\infty}(\theta - \theta_1)\tan(\alpha) = -\infty$ et un raisonnement analogue permet d'établir que $\lim_{\theta\to+\infty} \varphi(\theta) = -\frac{\pi}{2}$.

Comme l'application f est continue, le théorème des valeurs intermédiaires[121] indique qu'il existe $\theta_2 \in \mathbb{R}$ tel que $\varphi(\theta_2) = 0$. Cela signifie que la loxodromie traverse l'équateur au point de longitude θ_2.

b) Lorsque $\alpha = 0$, l'application f est constante (elle vaut φ_1) et la courbe γ correspond au parallèle P_{φ_1}.

---

[121] Voir le *Cours de première année* p. 594.

CHAPITRE 12

Intégrale multiple

La construction de « l'intégrale multiple », *i.e.* de l'intégrale des fonctions de plusieurs variables réelles, présentée ici reprend les étapes de la construction qui a été menée pour l'intégrale d'une fonction d'une variable réelle dans le *Cours de première année*, chap. 18, qui est habituellement qualifiée d'« intégrale simple » par opposition à « intégrale multiple ». Dans un premier temps, nous allons définir l'intégrale multiple d'une fonction en escalier sur un pavé P de $\mathbb{R}^n, n \geqslant 2$ (un pavé étant la généralisation naturelle à $\mathbb{R}^n$ de la notion d'intervalle dans $\mathbb{R}$). Puis nous étendrons la définition d'intégrale multiple sur un pavé à un ensemble de fonctions beaucoup plus grand, dites fonctions *intégrables au sens de Riemann.* Il restera ensuite à définir la notion d'intégrale multiple pour des fonctions définies sur un ensemble borné quelconque de $\mathbb{R}^n$. Outre le fait que le principe de la construction de l'intégrale multiple reprend le cheminement de la construction de l'intégrale simple, nous verrons dans la seconde partie du chapitre que les méthodes de calcul d'une intégrale multiple consistent principalement à écrire une telle intégrale à l'aide d'intégrales simples et à recourir aux techniques de calcul des intégrales simples présentées dans le *Cours de première année.*

12.1 Définition et premières propriétés

12.1.1 Fonctions en escalier sur un pavé de $\mathbb{R}^n$

DÉFINITION 12.1 (Pavé de $\mathbb{R}^n$)

✗ *On appelle pavé de $\mathbb{R}^n$ tout sous-ensemble* P *de $\mathbb{R}^n$ de la forme*

$$\mathrm{P} = [a_1, b_1] \times \cdots \times [a_n, b_n]$$

où pour tout $i \in \{1, ..., n\}$ les réels a_i et b_i vérifient $a_i < b_i$.

✗ *On appelle* mesure du pavé $\mathrm{P} = \prod_{i=1}^{n} [a_i, b_i]$ *le réel strictement positif*

$$m(\mathrm{P}) = \prod_{i=1}^{n} (b_i - a_i).$$

Les pavés de $\mathbb{R}$ sont les intervalles fermés et bornés ; la mesure d'un pavé est égale à la longueur de l'intervalle. Dans $\mathbb{R}^2$, les pavés correspondent à des rectangles dont les côtés sont parallèles aux axes de coordonnées et la mesure d'un pavé est égale à l'aire du rectangle.

Rappelons[(1)] qu'une subdivision d'un intervalle $[a, b]$ de $\mathbb{R}$ est une famille finie $\sigma = (c_j)_{j\in\{0,\dots,p\}}$ $(p \in \mathbb{N}^*)$ de réels vérifiant :

1. $\forall j \in \{0, \dots, p\} \quad c_j \in [a, b]$,
2. $c_0 = a$ et $c_p = b$,
3. $\forall j \in \{1, \dots, p\} \quad c_{j-1} < c_j$.

Les réels c_j sont appelés les *nœuds de la subdivision* σ.

Dans $\mathbb{R}^n$, une subdivision d'un pavé est définie de la manière suivante.

DÉFINITION 12.2 (Subdivision d'un pavé de $\mathbb{R}^n$)

✗ *Dans* $\mathbb{R}^n$, *on appelle* subdivision du pavé $\mathrm{P} = \prod_{i=1}^n [a_i, b_i]$ *une famille* $\sigma = \{\sigma_1, \dots, \sigma_n\}$ *où pour chaque* $i \in \{1, \dots, n\}$, σ_i *désigne une subdivision de l'intervalle* $[a_i, b_i]$.

✗ *Pour tout* $i \in \{1, \dots, n\}$, *notons* $c_{i,0}$, …, c_{i,p_i} *les nœuds de la subdivision* σ_i. *Les pavés* $\prod_{i=1}^n [c_{i,k_i}, c_{i,k_i+1}]$ *pour* $k_i \in \{0, \dots, p_i - 1\}$ *sont appelés les* cellules de la subdivision σ *de* $\mathrm{P} = \prod_{i=1}^n [a_i, b_i]$.

Dans $\mathbb{R}^2$, une subdivision d'un pavé correspond à un « découpage » du pavé en un nombre fini de rectangles de côtés parallèles aux axes de coordonnées, voir la figure 1. Chacun de ces rectangles, qui lui-même est un pavé, est appelé *cellule de la subdivision.*

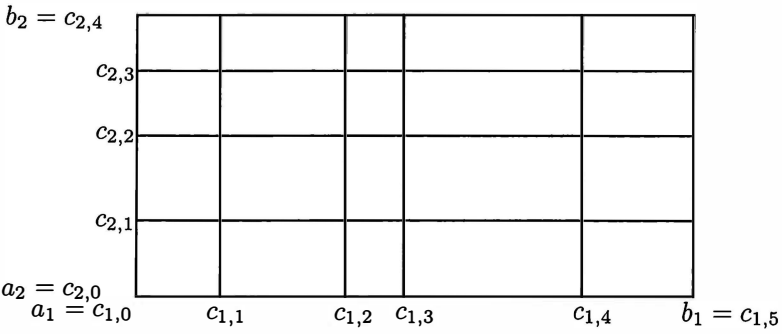


Fig. 1 Une subdivision d'un pavé du plan où la cellule $[c_{1,3}, c_{1,4}] \times [c_{2,1}, c_{2,2}]$ a été grisée.

(1) Voir le *Cours de première année*, p. 831.

De la même manière, dans $\mathbb{R}^n$ une subdivision d'un pavé correspond à un « découpage » du pavé en un nombre fini de pavés (les cellules). Pour toute subdivision d'un pavé P, la mesure du pavé $m(\mathrm{P})$ est égale à la somme des mesures des cellules $(\mathrm{P}_k)_{k\in\{1,\dots,N\}}$ formant la subdivision : $m(\mathrm{P}) = \sum_{k=1}^{N} m(\mathrm{P}_k)$. Si toutes les cellules $(\mathrm{P}_k)_{k\in\{1,\dots,N\}}$ de la subdivision du pavé P ont la même mesure, on dit que la subdivision du pavé P est *uniforme.*

DÉFINITION 12.3 (Fonction en escalier)
Soit f une application définie sur un pavé P *de $\mathbb{R}^n$, à valeurs dans $\mathbb{R}$.*

✗ *L'application f est dite* en escalier *sur le pavé* P *si f est bornée sur* P *et s'il existe une subdivision du pavé* P *telle que f soit constante sur l'intérieur*[(2)] *de chaque cellule de la subdivision.*

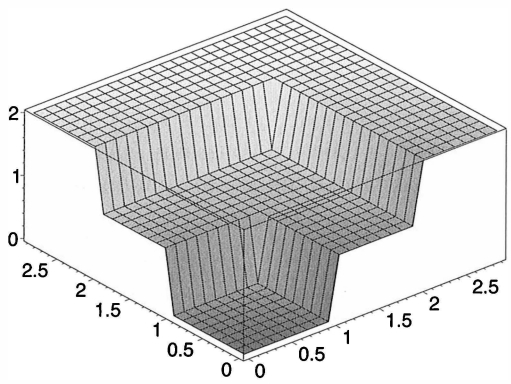
✗ *Une subdivision du pavé* P *est dite* adaptée *à l'application en escalier f si f est constante sur l'intérieur de chaque cellule de la subdivision.*

Exemple Soit $\mathrm{P} = [0,3] \times [0,3]$ un pavé de $\mathbb{R}^2$. L'application

$$(x_1, x_2) \in \mathrm{P} \longmapsto \max\{\mathrm{E}(x_1), \mathrm{E}(x_2)\} \in \mathbb{R},$$

où E désigne la fonction partie entière[(3)], est une application en escalier. Une subdivision adaptée est définie par les neuf cellules $\{[i, i+1] \times [j, j+1]\}_{0\leqslant i,j\leqslant 2}$. On peut visualiser graphiquement cette application avec MAPLE.

```
> with(plots):
> f:=(x,y)-> max(floor(x),floor(y));
> plot3d(f(x,y),x=0..3,y=0..3);
```



(2) Voir la définition 7.8 p. 377.

(3) Voir la proposition 3.9 p. 104 du *Cours de première année.*

Contrairement à ce qui a été vu pour les fonctions en escalier[4] sur $\mathbb{R}$, une fonction en escalier sur un pavé P de $\mathbb{R}^n$ avec $n \geqslant 2$ peut prendre un nombre infini de valeurs. En effet, à l'intérieur de chaque pavé la fonction ne prend qu'une seule valeur et il y a un nombre fini de pavés dans chaque subdivision mais la valeur de la fonction en escalier sur la frontière des pavés n'est pas imposée et peut prendre un nombre infini de valeurs.

Remarque L'ensemble des applications en escalier sur un pavé P de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}$ est un sous-espace vectoriel de l'espace vectoriel $\mathcal{A}(\mathrm{P}, \mathbb{R})$ des applications de P dans $\mathbb{R}$. ◇

12.1.2 Intégrale d'une fonction en escalier

DÉFINITION 12.4 (Intégrale d'une fonction en escalier)
Soient f une fonction en escalier sur un pavé P *et σ une subdivision adaptée à f. Désignons par* $\mathrm{P}_1,\ldots,\mathrm{P}_N$ *les cellules de cette subdivision et par* $\lambda_1,\ldots,\lambda_N$ *les valeurs constantes prises par f sur l'intérieur de chaque cellule.*
On appelle intégrale de f sur le pavé P *le réel*

$$I_\sigma(f) = \sum_{k=1}^{N} m(\mathrm{P}_k)\ \lambda_k$$

où $m(\mathrm{P}_k)$ désigne la mesure de la cellule P_k. Cette intégrale est notée $\int_{\mathrm{P}} f$ ou encore $\int_{\mathrm{P}} f(x)\ \mathrm{d}x$.

Remarques

1. On peut montrer que la valeur de $I_\sigma(f)$ ne dépend pas du choix de la subdivision σ adaptée à f.

2. Si f est la fonction constante égale à 1 alors $I_\sigma(f) = \sum_{k=1}^{N} m(\mathrm{P}_k) = m(\mathrm{P})$, autrement dit, l'intégrale de la fonction constante égale à 1 sur un pavé fournit la mesure du pavé (l'aire du pavé dans $\mathbb{R}^2$, le volume du pavé dans $\mathbb{R}^3$).

3. On peut donner une interprétation géométrique à la notion d'intégrale d'une fonction en escalier sur un pavé. Si f est une fonction positive en escalier sur un pavé P de $\mathbb{R}^2$ alors $m(\mathrm{P}_k)\ \lambda_k$ s'interprète comme le volume du parallélépipède rectangle de base le rectangle P_k et de hauteur λ_k, voir la figure 2. Ainsi, $I_\sigma(f) = \sum_{k=1}^{N} m(\mathrm{P}_k)\ \lambda_k$ peut s'interpréter comme le volume compris entre la surface d'équation $z = f(x, y)$ pour $(x, y) \in \mathrm{P}$ et le plan d'équation $x_3 = 0$. ◇

[4] Voir le *Cours de première année* p. 832.

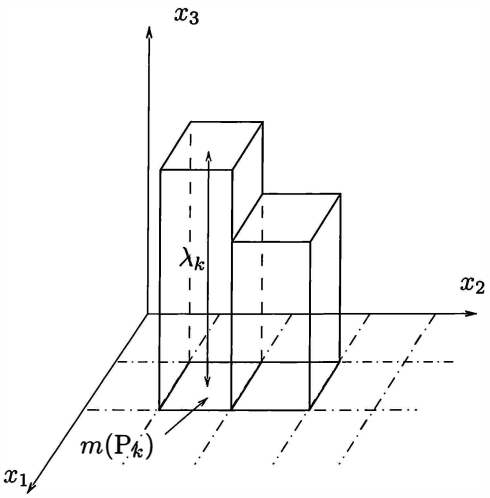


Fig. 2 Interprétation géométrique de l'intégrale d'une fonction en escalier.

Exemples

1. Considérons l'application f définie sur le pavé $\mathrm{P} = [0,3] \times [0,3]$ par

$$f : (x_1, x_2) \in \mathrm{P} \longmapsto \max\{\mathrm{E}(x_1), \mathrm{E}(x_2)\} \in \mathbb{R}.$$

L'application f est en escalier sur P et est constante sur l'intérieur de chacune des neufs cellules $\{[i, i+1] \times [j, j+1]\}_{(i,j)\in\{0,1,2\}}$ qui forment une subdivision du pavé P. Pour cette subdivision, la mesure des cellules est constante (on a une subdivision uniforme) et est égale à 1. Les valeurs prises par l'application f sur l'intérieur de chacune de ces neuf cellules sont données sur la figure 3.

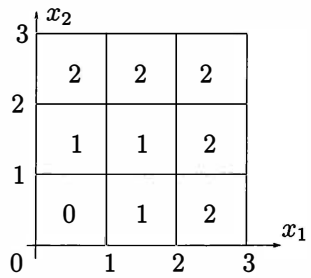


Fig. 3 Valeurs prises sur l'intérieur de chaque cellule par l'application $f : (x_1, x_2) \in [0,3]^2 \mapsto \max\{\mathrm{E}(x_1), \mathrm{E}(x_2)\}$.

D'après la définition 12.4, on a :

$$\int_{[0,3]\times[0,3]} f = 0 + 1 + 2 + 1 + 1 + 2 + 2 + 2 + 2 = 13.$$

2. Dans $\mathbb{R}^2$, considérons la subdivision uniforme du pavé $[0,1] \times [0,1]$ en n^2 cellules $(C_{ij})_{(i,j)\in\{1,\dots,n\}^2}$ $(n \in \mathbb{N}^*)$ définies par

$$C_{ij} = \left[\tfrac{i-1}{n}, \tfrac{i}{n}\right] \times \left[\tfrac{j-1}{n}, \tfrac{j}{n}\right]$$

et l'application en escalier ϕ_n définie sur $[0,1]^2$ par la valeur constante ij/n^2 qu'elle prend sur l'intérieur de chaque cellule[(5)] C_{ij}, voir la figure 4.

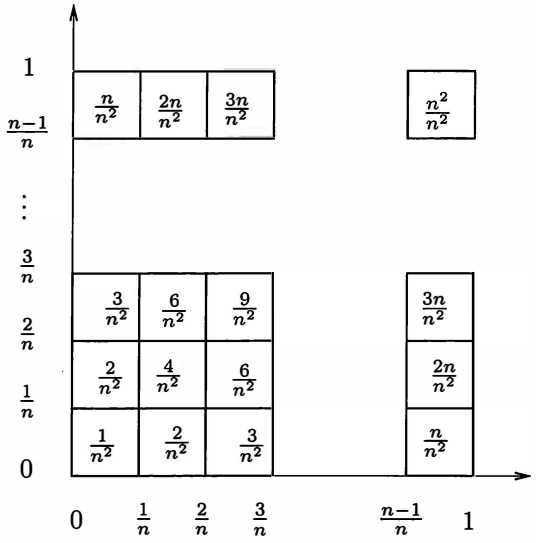


Fig. 4 Valeurs prises sur chaque cellule du pavé $[0,1]^2$ par l'application ϕ_n.

D'après la définition 12.4, l'intégrale de ϕ_n sur le pavé $[0,1]^2$ vaut

$$\int_{[0,1]^2} \phi_n = \sum_{i=1}^{n}\sum_{j=1}^{n} m(C_{ij})\,\frac{ij}{n^2}.$$

Étant donné que la mesure de toutes les cellules C_{ij} vaut $1/n^2$, on obtient

$$\int_{[0,1]^2} \phi_n = \frac{1}{n^4}\sum_{i=1}^{n}\sum_{j=1}^{n} ij = \frac{1}{n^4}\Big(\sum_{i=1}^{n} i\Big) \times \Big(\sum_{j=1}^{n} j\Big) = \frac{(n+1)^2}{4n^2}$$

car $\sum_{i=1}^{n} i = \frac{1}{2}n(n+1)$.

Exercice 12.1 Montrer que l'intégrale sur le pavé $[0,1]^2$ de l'application en escalier ψ_n $(n \in \mathbb{N}^*)$ définie sur l'intérieur de chaque cellule $C_{ij} = \left[\frac{i-1}{n}, \frac{i}{n}\right] \times \left[\frac{j-1}{n}, \frac{j}{n}\right]$ où $(i,j) \in \{1,\dots,n\}^2$ par $\psi_n(x_1,x_2) = \frac{(i-1)(j-1)}{n^2}$ vaut $(n-1)^2/(4n^2)$.

(5) La valeur de l'application f sur le bord de chaque cellule n'a aucune importance puisque cette valeur n'est pas prise en compte dans le calcul de l'intégrale d'une fonction en escalier.

12.1.3 Fonctions Riemann-intégrables sur un pavé

On étend la notion d'intégrale des fonctions en escalier sur un pavé de $\mathbb{R}^n$ aux fonctions de $\mathbb{R}^n$ dans $\mathbb{R}$ d'une manière similaire[6] à celle qui a permis d'étendre la notion d'intégrale des fonctions en escalier sur un intervalle $[a, b]$ aux fonctions de $\mathbb{R}$ dans $\mathbb{R}$.

DÉFINITION 12.5 (Riemann-intégrabilité sur un pavé)
Une application f définie sur un pavé P *de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}$ est dite* intégrable au sens de Riemann *(ou encore* Riemann-intégrable *ou plus simplement* intégrable*) sur* P *si pour tout réel ε strictement positif, il existe deux applications ψ_ε et ϕ_ε de* P *dans $\mathbb{R}$, en escalier sur* P*, telles que*

1. $\forall x \in \mathrm{P} \quad \psi_\varepsilon(x) \leqslant f(x) \leqslant \phi_\varepsilon(x)$,
2. $\displaystyle\int_{\mathrm{P}} (\phi_\varepsilon - \psi_\varepsilon) \leqslant \varepsilon$.

Remarques

1. Une fonction f en escalier sur un pavé P est intégrable au sens de Riemann sur P puisque les conditions de la définition 12.5 sont satisfaites en considérant $\phi_\varepsilon = \psi_\varepsilon = f$ pour tout $\varepsilon \in \mathbb{R}_+^*$.

2. Une fonction f en escalier sur un pavé P ayant été supposée bornée (voir la définition 12.3), une application intégrable au sens de Riemann sur un pavé P est bornée sur P d'après la condition 1 de la définition 12.5. ◇

Pour une fonction de $\mathbb{R}^n$ dans $\mathbb{R}$ intégrable au sens de Riemann sur un pavé P, on peut[7] d'après la définition 12.5 construire deux suites de fonctions en escalier $(\psi_k)_{k\in\mathbb{N}}$ et $(\phi_k)_{k\in\mathbb{N}}$ telles que :

1. $\forall k \in \mathbb{N} \quad \forall x \in \mathrm{P} \quad \psi_k(x) \leqslant f(x) \leqslant \phi_k(x)$,
2. $\displaystyle\lim_{k\to\infty} \int_{\mathrm{P}} (\phi_k - \psi_k) = 0$.

On peut montrer que les suites $(\int_{\mathrm{P}} \psi_k)_{k\in\mathbb{N}}$ et $(\int_{\mathrm{P}} \phi_k)_{k\in\mathbb{N}}$ sont des suites de Cauchy dans $\mathbb{R}$ et comme $\mathbb{R}$ est complet[8], cela implique qu'elles sont convergentes. On peut par ailleurs montrer qu'elles ont une même limite qui dépend uniquement de f et de P. Cela justifie que l'on introduise la définition suivante.

DÉFINITION 12.6 *Soit f une fonction de $\mathbb{R}^n$ dans $\mathbb{R}$ intégrable au sens de Riemann sur un pavé* P *de $\mathbb{R}^n$. On appelle* intégrale *de f sur* P *le réel* $I = \displaystyle\lim_{k\to\infty} \int_{\mathrm{P}} \psi_k = \lim_{k\to\infty} \int_{\mathrm{P}} \phi_k$ *que l'on note* $\displaystyle\int_{\mathrm{P}} f$ *ou* $\displaystyle\int_{\mathrm{P}} f(x)\,\mathrm{d}x$.

[6] Voir le *Cours de première année* p. 836.

[7] Il suffit de considérer successivement des réels ε de la forme $\varepsilon = \frac{1}{k+1}$ pour $k \in \mathbb{N}$.

[8] Voir la définition 7.21 p. 401.

Inversement, pour une fonction f de $\mathbb{R}^n$ dans $\mathbb{R}$ donnée, si on dispose de deux suites $(\psi_k)_{k\in\mathbb{N}}$ et $(\phi_k)_{k\in\mathbb{N}}$ de fonctions en escalier sur un pavé P de $\mathbb{R}^n$ telles que :

1. $\forall k \in \mathbb{N} \quad \forall x \in \mathrm{P} \quad \psi_k(x) \leqslant f(x) \leqslant \phi_k(x)$,
2. $\lim\limits_{k\to\infty} \int_{\mathrm{P}} (\phi_k - \psi_k) = 0$,

alors on peut en déduire que f est intégrable au sens de Riemann sur le pavé P avec

$$\int_{\mathrm{P}} f(x)\,\mathrm{d}x = \lim_{k\to\infty} \int_{\mathrm{P}} \psi_k = \lim_{k\to\infty} \int_{\mathrm{P}} \phi_k.$$

En effet, la convergence de la suite $(I_k)_k$ de terme général $I_k = \int_{\mathrm{P}} (\phi_k - \psi_k)$ vers 0 donnée par la condition 2, se traduit par[9]

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N_\varepsilon \in \mathbb{N} \quad \forall k \in \mathbb{N} \quad (k \geqslant N_\varepsilon \implies I_k \leqslant \varepsilon).$$

On en déduit que pour tout $\varepsilon \in \mathbb{R}_+^*$, les fonctions en escalier ϕ_{N_ε} et ψ_{N_ε}, d'une part vérifient

$$\int_{\mathrm{P}} (\phi_{N_\varepsilon} - \psi_{N_\varepsilon}) \leqslant \varepsilon$$

et d'autre part, d'après la condition 1, vérifient

$$\forall x \in \mathrm{P} \quad \psi_{N_\varepsilon}(x) \leqslant f(x) \leqslant \phi_{N_\varepsilon}(x).$$

Les deux conditions de la définition 12.6 étant satisfaites, on en conclut que l'application f est Riemann-intégrable sur P.

Remarques

1. Dans le cas où $n \geqslant 2$, l'intégrale sur un pavé P d'une fonction de $\mathbb{R}^n$ dans $\mathbb{R}$ Riemann-intégrable sur P est qualifiée *d'intégrale multiple*. Cette dénomination est opposée à l'intégrale sur un intervalle fermé et borné $[a, b]$ d'une fonction de $\mathbb{R}$ dans $\mathbb{R}$ Riemann-intégrable sur $[a, b]$ qui est qualifiée d'*intégrale simple*.

2. De nombreuses notations sont employées pour désigner l'intégrale d'une fonction de $\mathbb{R}^n$ dans $\mathbb{R}$ Riemann-intégrable sur un pavé P. On rencontrera par exemple à la place de $\int_{\mathrm{P}} f(x)\,\mathrm{d}x$ les notations suivantes :

$$\int_{\mathrm{P}} f \qquad \text{ou} \qquad \int_{\mathrm{P}} f(x_1, ..., x_n)\,\mathrm{d}x_1 ...\,\mathrm{d}x_n.$$

Dans le cas où $n = 2$, on parle d'*intégrale double* et on note le plus souvent l'intégrale avec deux symboles de sommation : $\iint_P f(x_1, x_2)\,\mathrm{d}x_1\mathrm{d}x_2$; dans le cas $n = 3$, on parle d'*intégrale triple* et on note souvent l'intégrale avec trois symboles de sommation : $\iiint_P f(x_1, x_2, x_3)\,\mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3$. ◇

[9] Remarquer que la suite $(I_k)_k$ est à termes positifs compte tenu de la condition 1.

Exemple Considérons l'application $f : (x_1, x_2) \in [0,1]^2 \mapsto x_1 x_2 \in \mathbb{R}$ et les deux suites fonctions en escalier $(\phi_{k\in\mathbb{N}})_{k\in\mathbb{N}}$ et $(\psi_{k\in\mathbb{N}})_{k\in\mathbb{N}}$ définies sur le pavé[10] $P = [0,1]^2$ par : $\forall (i,j) \in \{1,\dots,k\}^2 \quad \forall (x_1,x_2) \in \left]\frac{i-1}{k}, \frac{i}{k}\right[\times \left]\frac{j-1}{k}, \frac{j}{k}\right[$

$$\psi_k(x_1,x_2) = \frac{(i-1)(j-1)}{k^2} \quad \text{et} \quad \phi_k(x_1,x_2) = \frac{ij}{k^2}.$$

Remarquons que pour tout $(x_1,x_2) \in \left]\frac{i-1}{k}, \frac{i}{k}\right[\times \left]\frac{j-1}{k}, \frac{j}{k}\right[$ on a

$$\frac{(i-1)(j-1)}{k^2} \leqslant x_1\, x_2 \leqslant \frac{ij}{k^2}$$

ce qui implique, voir la figure 5, que[10]

$$\forall (x_1,x_2) \in [0,1]^2 \quad \psi_k(x_1,x_2) \leqslant f(x_1,x_2) \leqslant \phi_k(x_1,x_2).$$

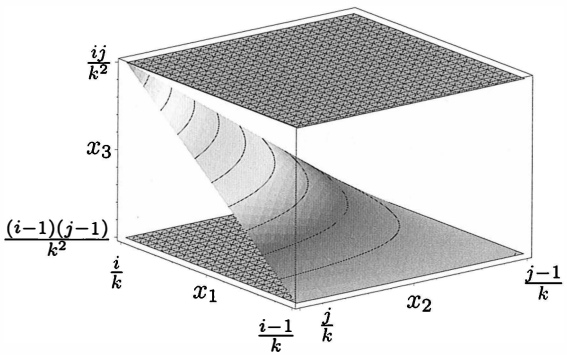


Fig. 5 Représentation graphique des fonctions f, ϕ_k et ψ_k sur la cellule $C_{ij} = \left[\frac{i-1}{k}, \frac{i}{k}\right] \times \left[\frac{j-1}{k}, \frac{j}{k}\right]$.

Pour tout $k \in \mathbb{N}^*$, on a[11] $\int_{[0,1]^2} \psi_k = \frac{(k-1)^2}{4k^2}$ et $\int_{[0,1]^2} \phi_k = \frac{(k+1)^2}{4k^2}$; on en déduit que[12] :

$$\int_{[0,1]^2} (\phi_k - \psi_k) = \frac{(k+1)^2}{4k^2} - \frac{(k-1)^2}{4k^2} = \frac{1}{k}$$

et par conséquent que $\lim_{k\to\infty} \int_P (\phi_k - \psi_k) = 0$. On en conclut que l'application f est Riemann-intégrable sur le pavé $[0,1]^2$ et que l'on a

$$\int_{[0,1]^2} f(x)\, \mathrm{d}x = \lim_{k\to+\infty} \int_{[0,1]^2} \psi_k = \lim_{k\to+\infty} \int_{[0,1]^2} \phi_k = \frac{1}{4}.$$

---

[10] Nous ne définissons les deux fonctions ψ_k et ϕ_k qu'à l'intérieur des cellules et non pas sur le pavé $[0,1]^2$. Comme nous l'avons mentionné, les valeurs de ces fonctions sur les bords des cellules n'entrent pas en compte pour le calcul de leur intégrale sur le pavé. On considère, sans le mentionner explicitement, que les deux fonctions ψ_k et ϕ_k prennent des valeurs sur le bord des cellules qui sont compatibles avec la relation 1 de la définition 12.5.

[11] Ces relations ont été établies dans l'exemple p. 860 et dans l'exercice 12.1.

[12] On utilise ici une propriété de linéarité de l'intégrale des fonctions en escalier sur un pavé : $\int_{[0,1]^2} \phi_k - \int_{[0,1]^2} \psi_k = \int_{[0,1]^2} (\phi_k - \psi_k)$. Cette propriété est directement issue de la définition 12.4 de l'intégrale d'une fonction en escalier sur un pavé.

Nous verrons ultérieurement[13] qu'il est possible de déterminer la valeur de l'intégrale $\int_{[0,1]^2} x_1 x_2 \, dx_1 dx_2$ de façon plus directe, sans avoir recours à la définition 12.6.

Comme pour l'intégrale des fonctions d'une variable réelle, un point important dans l'étude de l'intégrale des fonctions de plusieurs variables consiste à déterminer des critères, simples à vérifier, permettant d'établir qu'une fonction est Riemann-intégrable. Le plus utilisé d'entre eux est donné dans la proposition suivante.

PROPOSITION 12.1 *Toute application continue sur un pavé est intégrable au sens de Riemann sur ce pavé.*

Ce résultat est admis. Sa démonstration reprend l'approche utilisée pour démontrer qu'une fonction continue sur un intervalle $[a, b]$ de $\mathbb{R}$ est Riemann-intégrable sur cet intervalle, voir la proposition 18.2 p. 842 du *Cours de première année*.

Remarques

1. La condition d'intégrabilité donnée à la proposition 12.1 est une condition suffisante mais non nécessaire. Il existe des fonctions non continues qui sont Riemann-intégrables, comme par exemple les fonctions en escalier.

2. Si f est une fonction continue sur un pavé $P = [a_1, b_1] \times [a_2, b_2]$ de $\mathbb{R}^2$ et positive, l'intégrale double $\int_P f(x_1, x_2) \, dx_1 dx_2$ s'interprète comme le volume du domaine Ω limité par la surface d'équation $x_3 = f(x_1, x_2)$, le pavé P et les plans d'équation $x_1 = a_1$, $x_1 = b_1$, $x_2 = a_2$ et $x_2 = b_2$.

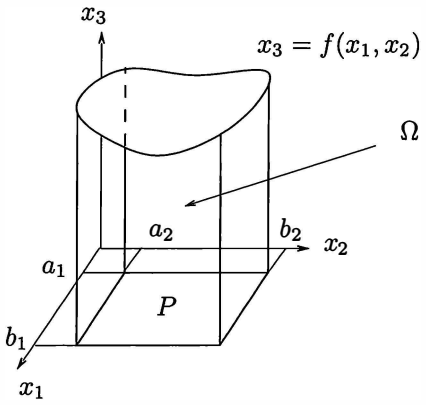


Fig. 6 Interprétation de l'intégrale double $\int_P f(x_1, x_2) \, dx_1 dx_2$.

[13] Voir p. 877.

Dans le cas où f n'est pas positive, l'intégrale double $\int_P f(x_1, x_2)\, \mathrm{d}x_1 \mathrm{d}x_2$ s'interprète comme le volume « algébrique » du domaine Ω, ce qui signifie que la partie du volume appartenant au demi-espace d'équation $x_3 \leqslant 0$ est comptabilisée comme étant négative. ◇

12.1.4 Fonctions Riemann-intégrables sur un ensemble borné

Nous nous sommes intéressé jusqu'à présent à définir l'intégrale d'une fonction de $\mathbb{R}^n$ dans $\mathbb{R}$ sur un pavé de $\mathbb{R}^n$. Dans la pratique du calcul intégral, on est souvent amené à devoir intégrer sur des ensembles de $\mathbb{R}^n$ plus généraux et il est indispensable de compléter la construction de l'intégrale multiple pour envisager cette situation. Nous allons décrire une manière de définir une notion d'intégrale sur un ensemble borné[14] quelconque. Une construction détaillée est donnée dans J. LELONG-FERRAND et J.M. ARNAUDIÈS, *Cours de mathématiques*, tome 4.

Considérons une application f définie sur un ensemble borné X de $\mathbb{R}^n$ et prolongeons l'application f à tout pavé[15] P de $\mathbb{R}^n$ contenant X de la manière suivante

$$\forall x \in \mathrm{P} \qquad f_{\mathrm{P}}(x) = \begin{cases} f(x) & \text{si } x \in X \\ 0 & \text{si } x \in \mathrm{P} \setminus X \end{cases}$$

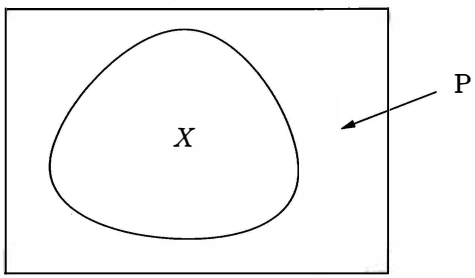


Fig. 7 Prolongement au pavé P de l'application f définie sur X.

On montre que si pour un choix du pavé P contenant l'ensemble X la fonction f_{P} est Riemann-intégrable sur P alors il en est de même pour tout autre choix de pavé contenant X ; de plus l'intégrale de f_{P} sur P ne dépend pas du choix du pavé. On pose par conséquent la définition suivante.

[14] Le cas des ensembles non bornés conduit, comme en dimension 1 (voir le *Cours de première année*, chap. 19) à la notion d'intégrale généralisée de plusieurs variables.

[15] Un tel pavé existe toujours puisque l'ensemble X est borné.

DÉFINITION 12.7 (Riemann-intégrabilité sur un ensemble borné)
Une application f définie sur un ensemble borné X de $\mathbb{R}^n$ est dite intégrable au sens de Riemann *(ou encore* Riemann-intégrable *ou plus simplement* intégrable*) sur X s'il existe un pavé* P *contenant X tel que l'application f_{P} qui est le prolongement par* 0 *de f sur ce pavé, soit intégrable au sens de Riemann sur* P*. On appelle alors* intégrale *de f sur X, et on note $\int_X f(x)\,\mathrm{d}x$ ou encore $\int_X f$, l'intégrale de f_{P} sur* P*.*

Remarque Une application intégrable au sens de Riemann sur un pavé P étant bornée sur P, il en résulte qu'une application intégrable sur un ensemble borné X de $\mathbb{R}^n$ est nécessairement bornée. ◇

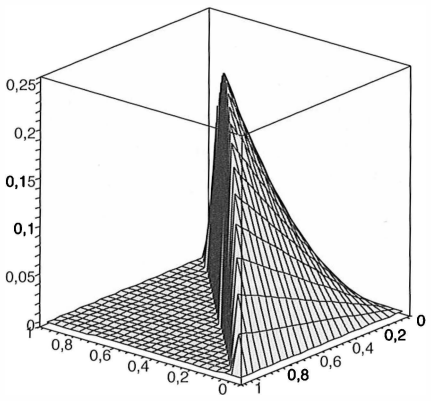
Exemple Notons T le sous-ensemble de $\mathbb{R}^2$ défini par[16]

$$T = \left\{(x_1, x_2) \in \mathbb{R}^2 \ ;\ 0 \leqslant x_1 \leqslant 1, \quad 0 \leqslant x_2 \leqslant 1, \quad x_1 + x_2 \leqslant 1\right\}$$

et considérons l'application $f : (x_1, x_2) \in T \mapsto x_1 x_2 \in \mathbb{R}$.

Notons f_{P} l'application qui prolonge f par 0 sur le pavé $\mathrm{P} = [0,1] \times [0,1]$ et montrons que f_{P} est Riemann-intégrable sur P.

```
> f:=proc(x,y)
    if (x>=0 and x<=1 and y>=0 and y<=1 and x+y<=1) then x*y;
    else 0; end if;
  end proc:
> plot3d(f,0..1,0..1,axes=boxed,orientation=[130,70],color=grey);
```



[16] Ce sous-ensemble du plan $\mathbb{R}^2$ est un triangle de sommet $(0,0)$, $(1,0)$ et $(0,1)$.

Considérons l'application en escalier ϕ_n ($n \in \mathbb{N}^*$) définie[17] sur les cellules $C_{ij} = \left[\frac{i-1}{n}, \frac{i}{n}\right] \times \left[\frac{j-1}{n}, \frac{j}{n}\right]$ par

$$\forall (x_1, x_2) \in \mathrm{Int}\,(C_{ij}) \quad \phi_n(x_1, x_2) = \begin{cases} \dfrac{i}{n} \times \dfrac{j}{n} & \text{si } \dfrac{i}{n} + \dfrac{j}{n} \leqslant 1 \\ 0 & \text{sinon} \end{cases}$$

et l'application en escalier ψ_n définie sur les cellules C_{ij} par

$$\forall (x_1, x_2) \in \mathrm{Int}\,(C_{ij}) \quad \psi_n(x_1, x_2) = \begin{cases} \dfrac{i-1}{n} \times \dfrac{j-1}{n} & \text{si } \dfrac{i}{n} + \dfrac{j}{n} \leqslant 1 \\ 0 & \text{sinon} \end{cases}.$$

Les valeurs prises par l'application ϕ_n sont indiquées sur la figure 8.

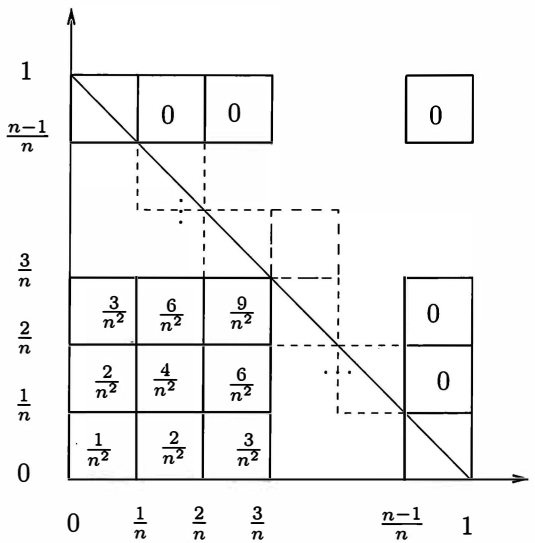


Fig. 8 Représentation de la fonction en escalier ϕ_n sur le pavé $[0,1]^2$.

Calculons l'intégrale de la fonction en escalier ϕ_n sur le pavé P. On a

$$\begin{aligned} \int_{\mathrm{P}} \phi_n &= \frac{1}{n^4} \sum_{i=1}^{n} \sum_{j=1}^{n-i} ij = \frac{1}{n^4} \sum_{i=1}^{n} i \frac{(n-i)(n-i+1)}{2} \\ &= \frac{1}{2n^4} \sum_{i=1}^{n} \left(i^3 - (2n+1)i^2 + n(n+1)i\right) \\ &= \frac{1}{2n^4} \left(\frac{n^2(n+1)^2}{4} - \frac{n(n+1)(2n+1)^2}{6} + \frac{n^2(n+1)^2}{2} \right) \end{aligned}$$

[17] Nous ne définissons en fait les deux fonctions ψ_n et ϕ_n qu'à l'intérieur des cellules. Comme nous l'avons mentionné, les valeurs de ces fonctions sur les bords des cellules n'entrent pas en compte pour le calcul de leur intégrale sur un pavé. On considère, sans le mentionner explicitement, que les deux fonctions ψ_n et ϕ_n prennent des valeurs sur le bord des cellules qui sont compatibles avec la relation 1 de la définition 12.5.

car[18] $\sum_{i=1}^{n} i^3 = \frac{1}{4}n^2(n+1)^2$, $\sum_{i=1}^{n} i^2 = \frac{1}{6}n(n+1)(2n+1)$ et $\sum_{i=1}^{n} i = \frac{1}{2}n(n+1)$. On en déduit que :

$$\lim_{n\to+\infty} \int_{\mathrm{P}} \phi_n = \frac{1}{2}\left(\frac{1}{4} - \frac{4}{6} + \frac{1}{2}\right) = \frac{1}{24}.$$

Par un calcul analogue, on évalue $\int_{\mathrm{P}} \psi_n$ pour chaque $n \in \mathbb{N}^*$ et on montre que sa limite lorsque n tend vers $+\infty$ vaut aussi $1/24$. On a donc $\lim_{n\to+\infty} \int_{\mathrm{P}} (\phi_n - \psi_n) = 0$.

On vérifie par ailleurs que pour tout $n \in \mathbb{N}^*$ et pour tout $(x_1, x_2) \in \mathrm{P}$, on a[18]

$$\psi_n(x_1, x_2) \leqslant f_{\mathrm{P}}(x_1, x_2) \leqslant \phi_n(x_1, x_2).$$

On déduit des définitions 12.5 et 12.6 que l'application f_{P} est Riemann-intégrable sur le pavé P et que $\int_{\mathrm{P}} f_{\mathrm{P}} = 1/24$. Cela implique d'après la définition 12.7 que l'application f est Riemann-intégrable sur le triangle T et que $\int_T f = 1/24$.

Nous verrons ultérieurement dans ce chapitre des méthodes de calcul d'une intégrale multiple beaucoup plus rapides que celle présentée dans l'exemple précédent qui utilise explicitement la définition de l'intégrale multiple. Parallèlement, nous devons disposer de critères simples à utiliser permettant d'établir, sans avoir recours à la définition, qu'une fonction définie sur un ensemble borné est Riemann-intégrable. Nous avons indiqué à la proposition 12.1 que toute application continue sur un pavé est intégrable au sens de Riemann sur ce pavé. Un tel critère ne permet pas en pratique d'établir la Riemann-intégrabilité d'une fonction sur un ensemble borné X. En effet, en prolongeant une application f continue sur un ensemble borné X par la valeur 0 à un pavé P qui contient X, on dispose d'une application f_{P} qui possède des discontinuités en chaque point de la frontière de X où f n'a pas une limite nulle.

Nous allons dans ce qui suit expliciter des conditions d'intégrabilité sur un ensemble borné, pour des fonctions de $\mathbb{R}^n$ dans $\mathbb{R}$.

12.1.5 Conditions d'intégrabilité

DÉFINITION 12.8 (Ensemble négligeable au sens de Riemann)
Un ensemble E de $\mathbb{R}^n$ est dit négligeable *au sens de Riemann si pour tout réel ε strictement positif, il existe une famille finie de pavés $(\mathrm{P}_k)_{k\in\{1,\dots,N\}}$ tel que :*

1. *la famille $(\mathrm{P}_k)_{k\in\{1,\dots,N\}}$ recouvre E, i.e. $E \subset \bigcup_{k=1}^{N} \mathrm{P}_k$;*
2. *la somme des mesures*[19] *est inférieure à ε, i.e. $\sum_{k=1}^{N} m(\mathrm{P}_k) \leqslant \varepsilon$.*

[18] Voir la proposition 3.5 p. 99 du *Cours de première année*.
[19] Voir la définition 12.1 p. 855.

Exemples

1. La frontière du pavé $P = [0,1] \times [0,1]$ dans $\mathbb{R}^2$ est négligeable. Pour l'établir, considérons les quatre pavés représentés à la figure 9 et définis par $P_1 = [-1/n, 1+1/n] \times [-1/n, 1/n]$, $P_2 = [1-1/n, 1+1/n] \times [-1/n, 1+1/n]$, $P_3 = [-1/n, 1+1/n] \times [1-1/n, 1+1/n]$ et $P_4 = [-1/n, 1/n] \times [-1/n, 1+1/n]$. Ils constituent un recouvrement de la frontière du pavé $[0,1]^2$.

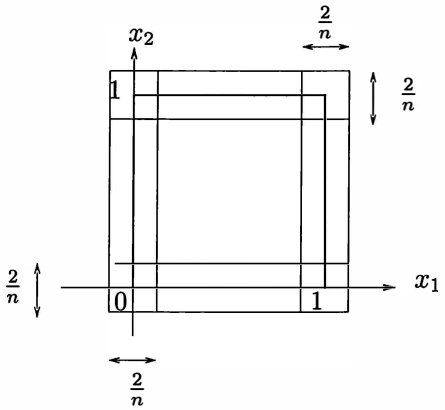


Fig. 9 Recouvrement de la frontière du pavé $[0,1]^2$ à l'aide de quatre pavés.

On établit facilement que la somme des mesures de ces 4 pavés vaut $s_n = \frac{8}{n}(1 + \frac{2}{n})$. Comme $\lim_{n \to +\infty} s_n = 0$, pour tout réel ε strictement positif, il existe un entier N tel que[20] $s_N \leqslant \varepsilon$. On en conclut que la frontière du pavé $P = [0,1]^2$ dans $\mathbb{R}^2$ est négligeable.

On montrerait sur le même principe que le bord de tout polygone du plan $\mathbb{R}^2$ est négligeable.

2. Montrons que dans $\mathbb{R}^2$, le quart de cercle de centre (0,0), de rayon 1 et d'extrémités $(0,1)$ et $(1,0)$ est négligeable. Pour $n \in \mathbb{N}^*$ fixé, considérons les pavés représentés à la figure 10 et définis pour $k \in \{1, \ldots, n\}$ par

$$P_k = \left[\cos(\tfrac{k\pi}{2n}), \cos(\tfrac{(k-1)\pi}{2n})\right] \times \left[\sin(\tfrac{(k-1)\pi}{2n}), \sin(\tfrac{k\pi}{2n})\right].$$

[20] La condition $\lim_{n \to +\infty} s_n = 0$ se traduit sous forme d'assertion quantifiée, voir la définition 5.1 p. 168 du *Cours de première année*, par

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists N \in \mathbb{N}^* \quad \forall n \in \mathbb{N} \quad (n \geqslant N \implies |s_n| \leqslant \varepsilon).$$

Puisque pour tout $n \in \mathbb{N}^*$ le réel s_n est strictement positif on a $0 < s_n \leqslant \varepsilon$ pour tout $n \geqslant N$; en particulier pour l'entier N, on a $0 < s_N \leqslant \varepsilon$.

Ils constituent un recouvrement du quart de cercle d'extrémités $(0,1)$ et $(1,0)$. La mesure du pave P_k vaut[21]

$$\begin{aligned} m(P_k) &= \left(\cos\left(\frac{(k-1)\pi}{2n}\right) - \cos\left(\frac{k\pi}{2n}\right)\right) \times \left(\sin\left(\frac{k\pi}{2n}\right) - \sin\left(\frac{(k-1)\pi}{2n}\right)\right) \\ &= 4\,\sin\left(\frac{(2k-1)\pi}{2n}\right)\,\sin^2\left(\frac{\pi}{2n}\right)\,\cos\left(\frac{(2k-1)\pi}{2n}\right) \\ &= 2\,\sin\left(\frac{(2k-1)\pi}{n}\right)\,\sin^2\left(\frac{\pi}{2n}\right). \end{aligned}$$

On en déduit que la somme s_n des mesures de ces n pavés vérifie[22]

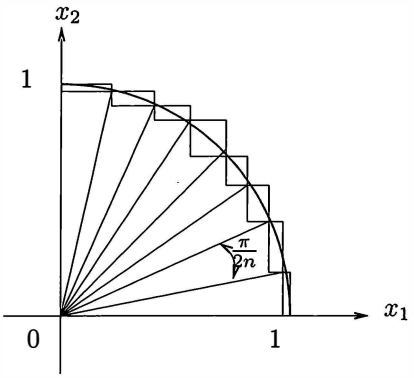


Fig. 10 Recouvrement du quart de cercle unité à l'aide de pavés.

$$0 \leqslant s_n = \sum_{k=1}^{n} m(P_k) = 2\,\sin^2\left(\frac{\pi}{2n}\right)\,\sum_{k=1}^{n} \underbrace{\sin\left((2k-1)\pi/n\right)}_{\leqslant 1} \leqslant 2n\,\sin^2\left(\frac{\pi}{2n}\right).$$

Comme $2n\sin^2\left(\frac{\pi}{2n}\right) \underset{n\to+\infty}{\sim} 2n\,\left(\frac{\pi}{2n}\right)^2 = \frac{\pi^2}{2n}$, on a $\lim_{n\to+\infty} s_n = 0$. Cela implique que pour tout réel ε strictement positif, il existe un entier N tel que $s_N \leqslant \varepsilon$, ce qui permet de conclure que le quart de cercle d'extrémités $(0,1)$ et $(1,0)$ est négligeable.

On peut également montrer en utilisant un raisonnement similaire que tout cercle du plan $\mathbb{R}^2$ est négligeable.

Remarque On montre qu'une réunion finie d'ensembles négligeables est un ensemble négligeable. ◇

[21] On utilise les formules trigonométrique de transformation de sommes en produits, voir le *Cours de première année* p. 156.

[22] On majore chaque terme en sinus de la somme par 1.

Nous avons mentionné le fait qu'en prolongeant une application f continue sur un ensemble borné X par la valeur 0 à un pavé P qui contient X (comme spécifié par la définition 12.7), on dispose d'une application f_P qui possède des discontinuités en chaque point de la frontière de X où f n'a pas une limite nulle et par conséquent que la proposition 12.1 ne permet pas d'établir l'intégrabilité de f_P sur P donc de f sur X. On démontre que si l'ensemble des points de discontinuité de f est un ensemble négligeable alors f_P est Riemann-intégrable sur P et par conséquent f est Riemann-intégrable sur X. De façon pratique, on utilise le résultat suivant que nous admettrons.

PROPOSITION 12.2 (Condition suffisante d'intégrabilité)
Soient X un sous-ensemble borné de $\mathbb{R}^n$ et f une application continue et bornée de X dans $\mathbb{R}$. Une condition suffisante pour que f soit Riemann-intégrable sur X est que la frontière de X soit négligeable.

La proposition 12.2 indique que si l'ensemble des points de discontinuité de la fonction intégrée f est négligeable alors il « n'intervient pas » au niveau de l'intégrabilité de f. La proposition suivante indique que par ailleurs les ensembles négligeables « n'interviennent pas » dans la valeur de l'intégrale.

PROPOSITION 12.3 *Soient X un sous-ensemble négligeable de $\mathbb{R}^n$ et f une application bornée sur X. L'application f est Riemann-intégrable sur X et son intégrale sur X est nulle.*

Démonstration Comme X est négligeable, pour tout $\varepsilon' \in \mathbb{R}_+^*$, il existe une famille finie de pavés $(P_k)_{k\in\{1,\dots,N\}}$ telle que

$$X \subset \bigcup_{k=1}^{N} P_k \quad \text{et} \quad \sum_{k=1}^{N} m(P_k) \leqslant \varepsilon'.$$

Notons $Q = \bigcup_{k=1}^{N} P_k$ et considérons un pavé P tel que $Q \subset P$ sur lequel on définit le prolongement f_P de f en posant $f_P(x) = f(x)$ si $x \in X$ et $f_P(x) = 0$ sinon. Pour montrer que f est Riemann-intégrable sur X, il faut montrer que f_P est Riemann-intégrable sur P. En d'autres termes, d'après la définition 12.5 il faut montrer que pour tout $\varepsilon \in \mathbb{R}_+^*$, il existe deux applications ψ et ϕ de P dans $\mathbb{R}$ en escalier sur P telles que

$$\forall x \in P \quad \psi(x) \leqslant f(x) \leqslant \phi(x) \quad \text{et} \quad \int_P (\phi - \psi) \leqslant \varepsilon.$$

L'application f étant bornée sur X, le réel $M = \sup_{x\in X} |f(x)|$ existe. Les fonctions ϕ et ψ définies sur P par

$$\psi(x) = \begin{cases} -M & \text{si } x \in Q \\ 0 & \text{sinon} \end{cases} \quad \text{et} \quad \phi(x) = \begin{cases} M & \text{si } x \in Q \\ 0 & \text{sinon} \end{cases}$$

sont des fonctions en escalier sur P qui vérifient $\psi(x) \leqslant f(x) \leqslant \phi(x)$ pour tout $x \in \mathrm{P}$. De plus,

$$\int_{\mathrm{P}} (\phi - \psi) \quad = \quad \int_{\mathrm{Q}} (\phi - \psi) = \sum_{k=1}^{N} \int_{\mathrm{P_k}} 2M = 2M \sum_{k=1}^{N} m(\mathrm{P_k}) = 2M\varepsilon'.$$

En reconsidérant le raisonnement pour la valeur $\varepsilon' = \frac{\varepsilon}{2M}$, on établit que $\int_{\mathrm{P}} (\phi - \psi) \leqslant \varepsilon$. On en conclut que f_{P} est Riemann-intégrable sur P et par conséquent que f est Riemann-intégrable sur X. Comme par ailleurs,

$$0 \leqslant \int_{\mathrm{P}} \phi = \int_{\mathrm{Q}} M = \sum_{k=1}^{N} \int_{\mathrm{P_k}} M = M \sum_{k=1}^{N} m(\mathrm{P_k}) = M\varepsilon' = \frac{\varepsilon}{2},$$

on en déduit que $\lim_{\varepsilon \to 0} \int_{\mathrm{P}} \phi = 0$ ce qui implique que $\int_{\mathrm{P}} f_P = 0$. On a donc également $\int_X f = 0$. □

La proposition 12.2 indique que les ensembles négligeables peuvent être « négligés » dans le calcul d'intégrales multiples. Si X est un ensemble borné tel que $X = X_1 \cup X_2$ avec $X_1 \cap X_2 = \varnothing$ et que X_2 est un ensemble négligeable, alors

$$\int_X f(x)\, \mathrm{d}x = \int_{X_1} f(x)\, \mathrm{d}x.$$

D'un point de vue pratique, on a fréquemment recours au résultat suivant.

Proposition 12.4 *Soit X un sous-ensemble borné de $\mathbb{R}^n$ de frontière négligeable. Toute fonction f de $\mathbb{R}^n$ dans $\mathbb{R}$ qui est Riemann-intégrable sur X est aussi Riemann-intégrable sur son intérieur* $\mathrm{Int}\,(X)$ *et sur son adhérence* $\overline{X}$ *et on a*

$$\int_X f(x)\, \mathrm{d}x = \int_{\mathrm{Int}(X)} f(x)\, \mathrm{d}x = \int_{\overline{X}} f(x)\, \mathrm{d}x.$$

Exemple Comme l'ensemble $X = [0,1]^2$ est de frontière négligeable, pour toute fonction f continue sur X et à valeurs réelles on a :

$$\int_{[0,1]^2} f(x)\, \mathrm{d}x = \int_{]0,1[^2} f(x)\, \mathrm{d}x.$$

De même, si f est une fonction continue sur le disque ouvert $D(0,1)$ dans $\mathbb{R}^2$ alors f est Riemann-intégrable sur le disque ouvert $D(0,1)$ ainsi que sur le disque fermé $\overline{D}(0,1)$ et on a

$$\int_{D(0,1)} f(x)\, \mathrm{d}x = \int_{\overline{D}(0,1)} f(x)\, \mathrm{d}x.$$

12.1.6 Propriétés des fonctions Riemann-intégrables

Les propriétés suivantes de l'intégrale de Riemann s'obtiennent, comme dans le cas des fonctions de $\mathbb{R}$ dans $\mathbb{R}$, à partir des propriétés analogues que l'on peut prouver dans un premier temps pour l'intégrale des fonctions en escalier. Elles sont admises. La première de ces propriétés est une propriété de linéarité.

PROPOSITION 12.5 *Soient X un ensemble borné de $\mathbb{R}^n$, f et g deux applications de X dans $\mathbb{R}$, Riemann-intégrables sur X. Pour tous réels λ et μ l'application $\lambda f + \mu g$ est Riemann-intégrable sur X et on a*

$$\int_X (\lambda f(x) + \mu g(x))\ \mathrm{d}x = \lambda \int_X f(x)\ \mathrm{d}x + \mu \int_X g(x)\ \mathrm{d}x.$$

La seconde propriété est une propriété dite de positivité.

PROPOSITION 12.6 *Soient X un ensemble borné de $\mathbb{R}^n$ et f une application de X dans $\mathbb{R}$, Riemann-intégrable sur X.*
Si l'application f est positive alors $\displaystyle\int_X f(x)\ \mathrm{d}x \geqslant 0.$

Ces deux premières propriétés permettent de montrer que si f et g sont deux applications d'un ensemble borné X de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}$, Riemann-intégrables sur X qui satisfont $f(x) \leqslant g(x)$ pour tout $x \in X$ alors $\int_X f(x)\ \mathrm{d}x \leqslant \int_X g(x)\ \mathrm{d}x$.

On montre aussi que si f est une application d'un ensemble borné X de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}$, Riemann-intégrable sur X, alors l'application $|f[$ est Riemann-intégrable sur X et

$$\left|\int_X f(x)\ \mathrm{d}x\right| \leqslant \int_X |f(x)|\ \mathrm{d}x.$$

La propriété suivante, connue sous le nom de *relation de Chasles*, nécessite l'introduction de la notion de partition d'un ensemble. Étant donné un sous-ensemble X de $\mathbb{R}^n$, on appelle *partition* de X toute famille $\{X_1, ..., X_N\}$ de sous-ensemble de X vérifiant les deux conditions suivantes :

1. $X = \bigcup_{k=1}^{N} X_k$,
2. $\forall (i,j) \in \{1, ..., N\}^2 \quad (\, i \neq j \implies \mathrm{Int}\,(X_i \cap X_j) = \varnothing\,)$.

PROPOSITION 12.7 (Relation de Chasles[23])
Soient X un ensemble borné de $\mathbb{R}^n$, $\{X_1, ..., X_N\}$ une partition de X et f une application de X dans $\mathbb{R}$, Riemann-intégrable sur X. On a :

$$\int_X f(x)\ \mathrm{d}x = \int_{X_1} f(x)\ \mathrm{d}x + \cdots + \int_{X_N} f(x)\ \mathrm{d}x.$$

On sera attentif au fait que la propriété suivante n'est vraie que pour des applications continues et positives.

Proposition 12.8 *Soient X un ensemble borné de $\mathbb{R}^n$ et f une application de X dans $\mathbb{R}$, positive et continue sur X. On a*

$$\int_X f(x)\ \mathrm{d}x = 0 \iff f = 0.$$

L'hypothèse de continuité dans la proposition 12.8 est essentielle. Par exemple la fonction en escalier

$$f : (x_1, x_2) \in [0,1] \times [0,1] \longmapsto \begin{cases} 1 & \text{si } x_1 = \frac{1}{2} \text{ ou } x_2 = \frac{1}{2} \\ 0 & \text{sinon.} \end{cases}$$

est non nulle mais son intégrale sur $[0,1] \times [0,1]$ est nulle.

La proposition 12.9 concernant l'intégrabilité du produit de deux fonctions Riemann-intégrables est une extension à l'intégrale multiple d'un résultat établi dans le *Cours de première année* pour l'intégrale simple[24]. Elle se démontre exactement de la même manière.

Proposition 12.9 (Inégalité de Cauchy-Schwarz)
Soient X un ensemble borné de $\mathbb{R}^n$, f et g deux applications de X dans $\mathbb{R}$ Riemann-intégrables sur X. L'application $f \times g$ est Riemann-intégrable sur X et

$$\left(\int_X f(x)\ g(x)\ \mathrm{d}x\right)^2 \leqslant \left(\int_X f(x)^2\ \mathrm{d}x\right) \times \left(\int_X g(x)^2\ \mathrm{d}x\right).$$

Exemples d'intégrales multiples dans les sciences de l'ingénieur

Terminons cette partie consacrée à la présentation de l'intégrale multiple en donnant quelques situations où intervient la notion d'intégrale multiple.

Le volume d'un solide Ω dans l'espace (de même que l'aire d'un solide dans le plan) est obtenu en intégrant la fonction constante égale à 1 sur Ω ; en d'autres termes

$$V = \int_\Omega 1\ \mathrm{d}x.$$

La masse M d'un solide Ω de densité volumique de masse ρ est donnée par

$$M = \int_\Omega \rho(x)\ \mathrm{d}x.$$

[23] Michel Chasles, mathématicien français né à Epernon en 1793 et mort à Paris en 1880.
[24] Voir la proposition 18.6 p. 846 du *Cours de première année*.

Dans un repère orthonormé direct $(O;\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3)$ les coordonnées (x_1^G,x_2^G,x_3^G) du centre d'inertie (ou centre de masse) G d'un solide Ω de densité volumique de masse ρ sont données pour $i\in\{1,2,3\}$ par

$$x_i^G=\frac{1}{M}\int_\Omega f_i(x)\,\mathrm{d}x \quad \text{où} \quad f_i : x=(x_1,x_2,x_3)\in\mathbb{R}^3 \longmapsto x_i\ \rho(x_1,x_2,x_3).$$

Dans le cas d'un solide de densité volumique de masse constante (« la matière est uniformément répartie ») on a plus simplement $x_i^G=\int_\Omega p_i(x)\,\mathrm{d}x$ où $p_i : x=(x_1,x_2,x_3)\in\mathbb{R}^3 \longmapsto x_i$ désigne la i–ième projection canonique.

Le moment d'inertie d'un solide par rapport à un axe quantifie la résistance de ce solide lorsqu'il est soumis à une mise en rotation par rapport à cet axe : plus le moment d'inertie est grand, plus il sera difficile de mettre en rotation l'objet. Le moment d'inertie d'un solide Ω de masse volumique ρ par rapport à un axe Δ est donné par l'intégrale

$$J_\Delta=\int_\Omega \rho(x)\ d(x,\Delta)^2\ \mathrm{d}x$$

où $d(x,\Delta)$ désigne la distance du point de coordonnées $x=(x_1,x_2,x_3)$ de Ω à l'axe Δ.

Il est courant lorsque le solide est de faible épaisseur de faire abstraction de la troisième dimension d'espace et de considérer que l'objet est situé dans un plan ; on est alors amené à calculer des intégrales doubles pour déterminer le centre de gravité de l'objet ou un de ses moments d'inertie.

12.2 Calcul des intégrales multiples

12.2.1 Le théorème de Fubini

Le calcul d'intégrales multiples en utilisant la définition 12.6 est souvent long et fastidieux. Nous allons présenter dans cette partie un certain nombre de méthodes permettant de calculer la valeur de l'intégrale d'une fonction Riemann-intégrable sans avoir à utiliser la définition 12.6. Le principe dans ces différentes méthodes est d'exprimer l'intégrale multiple à l'aide d'intégrales simples (à comprendre au sens d'intégrales de fonctions d'une variable réelle) afin d'utiliser les différentes techniques de calcul de l'intégrale d'une fonction d'une variable réelle présentées dans le chap. 18 du *Cours de première année*.

Les méthodes de calcul d'intégrales multiples que nous allons présenter ici sont toutes basées sur le théorème suivant que nous admettrons.

FUBINI, Guido (1879, Venise - 1943, New York).



Fubini entra à l'école Normale Supérieure de Pise en 1896 où il soutint en 1900 une thèse sur *le parallélisme de Clifford dans les espaces elliptiques.* À partir de 1908, il enseigna à l'Université de Turin où ses travaux de recherche concernèrent l'analyse complexe et l'analyse fonctionnelle. Lors de la Première Guerre mondiale, il s'intéressa à des sujets plus militaires notamment à la précision des tirs d'artillerie. En 1939, l'Italie de Mussolini adopta des lois anti-juives et Fubini craignit pour sa famille. Il accepta alors un poste d'enseignant à l'Université de Princeton où il mourut quatre ans plus tard.

THÉORÈME 12.1 (Théorème de Fubini)
Soit $\mathrm{P} = \prod_{i=1}^{n}[a_i, b_i]$ *un pavé de* $\mathbb{R}^n$ *et* $i_0 \in \{1, ..., n\}$ *un entier. Pour toute application* f *de* P *dans* $\mathbb{R}$, continue *sur* P *on a*

$$\int_{\mathrm{P}} f = \int_{a_{i_0}}^{b_{i_0}} \left(\int_{\mathrm{Q}} f \right) = \int_{\mathrm{Q}} \left(\int_{a_{i_0}}^{b_{i_0}} f \right)$$

où Q *désigne le pavé de* $\mathbb{R}^{n-1}$ *défini par* $\mathrm{Q} = \prod_{i=1, i\neq i_0}^{n}[a_i, b_i]$.

Considérons une application f de $\mathrm{P} = \prod_{i=1}^{n}[a_i, b_i]$ dans $\mathbb{R}$, continue sur P. En utilisant le théorème de Fubini successivement avec l'indice $i_0 = 1$, puis $i_0 = 2$, en poursuivant jusqu'à l'indice $i_0 = n$, on obtient :

$$\int_{\mathrm{P}} f = \int_{a_1}^{b_1} \left(\int_{a_2}^{b_2} \left(\cdots \left(\int_{a_n}^{b_n} f \right) \cdots \right) \right);$$

de plus, l'ordre dans lequel sont retenus les indices peut être choisi arbitrairement. Ainsi, le calcul d'une intégrale multiple sur un pavé de $\mathbb{R}^n$ se ramène au calcul de n intégrales simples.

Dans le cas où $n = 2$, l'intégrale sur un pavé $\mathrm{P} = [a_1, b_1] \times [a_2, b_2]$ d'une application continue sur P vérifie d'après le théorème de Fubini les relations :

$$\int_{\mathrm{P}} f = \int_{a_1}^{b_1} \left(\int_{a_2}^{b_2} f \right) = \int_{a_2}^{b_2} \left(\int_{a_1}^{b_1} f \right).$$

D'un point de vue pratique, afin de mieux préciser les variables successives par rapport auxquelles on intègre, on a coutume d'écrire ces relations de la manière suivante :

$$\begin{aligned}\int_{\mathrm{P}} f(x_1,x_2)\,\mathrm{d}x_1\mathrm{d}x_2 &= \int_{a_1}^{b_1}\left(\int_{a_2}^{b_2} f(x_1,x_2)\,\mathrm{d}x_2\right)\mathrm{d}x_1\\ &= \int_{a_2}^{b_2}\left(\int_{a_1}^{b_1} f(x_1,x_2)\,\mathrm{d}x_1\right)\mathrm{d}x_2.\end{aligned}$$

La condition de continuité de f sur le pavé $\mathrm{P} = [a_1,b_1]\times[a_2,b_2]$ est une condition suffisante pour que l'égalité

$$\int_{a_1}^{b_1}\left(\int_{a_2}^{b_2} f(x_1,x_2)\,\mathrm{d}x_2\right)\mathrm{d}x_1 = \int_{a_2}^{b_2}\left(\int_{a_1}^{b_1} f(x_1,x_2)\,\mathrm{d}x_1\right)\mathrm{d}x_2$$

ait lieu. Bien que cette condition ne soit pas nécessaire, il serait faux de penser que l'on a toujours égalité.

Par exemple, pour l'application f définie sur $[0,1]^2$ par $f(x_1,x_2) = \dfrac{x_1-x_2}{(x_1+x_2)^3}$ si $(x_1,x_2)\neq 0$ et $f(0,0)=0$, on a

$$\int_0^1\left(\int_0^1 f(x_1,x_2)\,\mathrm{d}x_2\right)\mathrm{d}x_1 = \int_0^1\left(\int_0^1 \frac{2x_1}{(x_1+x_2)^3}-\frac{1}{(x_1+x_2)^2}\,\mathrm{d}x_2\right)\mathrm{d}x_1$$

Or

$$\int_0^1 \frac{2x_1}{(x_1+x_2)^3}-\frac{1}{(x_1+x_2)^2}\,\mathrm{d}x_2 = \left[\frac{-x_1}{(x_1+x_2)^2}+\frac{1}{x_1+x_2}\right]_0^1 = \frac{1}{(x_1+1)^2}$$

d'où

$$\int_0^1\left(\int_0^1 f(x_1,x_2)\,\mathrm{d}x_2\right)\mathrm{d}x_1 = \int_0^1 \frac{1}{(x_1+1)^2}\,\mathrm{d}x_1 = \frac{1}{2}$$

mais[25]

$$\int_0^1\left(\int_0^1 f(x_1,x_2)\,\mathrm{d}x_1\right)\mathrm{d}x_2 = -\frac{1}{2}.$$

On vérifiera que l'application f n'est pas continue en 0.

Exemples

1. Considérons l'application $f : (x_1,x_2)\in[0,1]^2 \mapsto x_1x_2$. Cette application étant continue sur le pavé $[0,1]^2$, d'après le théorème de Fubini on a

$$\int_{[0,1]^2} f(x_1,x_2)\,\mathrm{d}x_1\mathrm{d}x_2 = \int_0^1\left(\int_0^1 f(x_1,x_2)\,\mathrm{d}x_2\right)\mathrm{d}x_1.$$

[25] Il suffit d'échanger les rôles de x_1 et x_2 dans le calcul précédent.

Pour tout $x_1 \in [0,1]$, on a

$$\int_0^1 f(x_1,x_2)\,\mathrm{d}x_2 = \int_0^1 x_1x_2\,\mathrm{d}x_2 = x_1\int_0^1 x_2\,\mathrm{d}x_2 = \frac{x_1}{2},$$

de sorte que

$$\int_{[0,1]^2} f(x_1,x_2)\,\mathrm{d}x_1\mathrm{d}x_2 = \int_0^1 \frac{1}{2}x_1\,\mathrm{d}x_1 = \frac{1}{4}.$$

On retrouve le résultat obtenu à la p. 863 où le calcul avait été effectué en ayant recours à la définition 12.6 ; on pourra comparer les deux méthodes de calcul.

2. Calculons l'intégrale sur le pavé $\mathrm{P} = [1,2]\times[0,1]$ de l'application $f : (x_1,x_2) \in \mathrm{P} \mapsto (x_1^2+x_2)^{-\frac{1}{2}}$. Comme pour tout $(x_1,x_2) \in \mathrm{P}$ on a $x_1^2+x_2 > 0$ et que la fonction puissance $-\frac{1}{2}$ est continue sur $\mathbb{R}_+^*$, l'application f est une application continue sur P. On en déduit, d'une part que f est bien Riemann-intégrable[26] sur P, et d'autre part que, d'après le théorème de Fubini,

$$\int_{\mathrm{P}} f(x_1,x_2)\,\mathrm{d}x_1\mathrm{d}x_2 = \int_1^2 \left(\int_0^1 \frac{1}{\sqrt{x_1^2+x_2}}\,\mathrm{d}x_2\right)\,\mathrm{d}x_1.$$

Pour tout $x_1 \in [1,2]$ fixé, on a

$$\int_0^1 \frac{1}{\sqrt{x_1^2+x_2}}\,\mathrm{d}x_2 = \left[2\sqrt{x_1^2+x_2}\right]_{x_2=0}^{x_2=1} = 2\left(\sqrt{x_1^2+1}-x_1\right).$$

On en déduit que

$$\begin{aligned}\int_{\mathrm{P}} f(x_1,x_2)\,\mathrm{d}x_1\mathrm{d}x_2 &= 2\int_1^2 (\sqrt{x_1^2+1}-x_1)\,\mathrm{d}x_1 = 2\int_1^2 \sqrt{x_1^2+1}\,\mathrm{d}x_1 - \left[x_1^2\right]_1^2\\ &= -3+2\int_1^2 \sqrt{x_1^2+1}\,\mathrm{d}x_1.\end{aligned}$$

Calculons $\int_1^2 \sqrt{x_1^2+1}\,\mathrm{d}x_1$ en effectuant le changement de variables[27] $x_1 = \mathrm{sh}(t)$. L'application $\phi : t \in [\mathrm{argsh}(1),\mathrm{argsh}(2)] \mapsto \mathrm{sh}(t) \in [1,2]$ étant une application de classe $\mathcal{C}^1$ sur $[\mathrm{argsh}(1),\mathrm{argsh}(2)]$, la formule de changement de variables[28] pour une intégrale simple permet d'écrire

$$\begin{aligned}\int_1^2 \sqrt{x_1^2+1}\,\mathrm{d}x_1 &= \int_{\phi(\mathrm{argsh}(1))}^{\phi(\mathrm{argsh}(2))} \sqrt{x_1^2+1}\,\mathrm{d}x_1 = \int_{\mathrm{argsh}(1)}^{\mathrm{argsh}(2)} \sqrt{\phi^2(t)+1}\,\phi'(t)\,\mathrm{d}t\\ &= \int_{\mathrm{argsh}(1)}^{\mathrm{argsh}(2)} \sqrt{\mathrm{sh}^2(t)+1}\,\mathrm{ch}(t)\,\mathrm{d}t.\end{aligned}$$

---

[26] Voir la proposition 12.1 p. 864.

[27] Signalons que l'on peut également calculer cette intégrale en utilisant une intégration par parties.

[28] Voir le théorème 18.2 p. 863 du *Cours de première année*.

Pour tout $t \in \mathbb{R}$, on a $\sqrt{\mathrm{sh}^2(t)+1} = \sqrt{\mathrm{ch}^2(t)} = |\mathrm{ch}(t)| = \mathrm{ch}(t)$, donc[29]

$$\begin{aligned}\int_1^2 \sqrt{x_1^2+1}\,\mathrm{d}x_1 &= \int_{\mathrm{argsh}(1)}^{\mathrm{argsh}(2)} \mathrm{ch}^2(t)\,\mathrm{d}t = \frac{1}{2}\int_{\mathrm{argsh}(1)}^{\mathrm{argsh}(2)} (\mathrm{ch}(2t)+1)\,\mathrm{d}t \\ &= \tfrac{1}{2}\Big[\tfrac{1}{2}\,\mathrm{sh}(2t)+t\Big]_{\mathrm{argsh}(1)}^{\mathrm{argsh}(2)} \\ &= \tfrac{1}{2}\,\mathrm{argsh}(2) - \tfrac{1}{2}\,\mathrm{argsh}(1) + \tfrac{1}{4}\,\mathrm{sh}(2\,\mathrm{argsh}(2)) - \tfrac{1}{2}\,\mathrm{sh}(2\,\mathrm{argsh}(1)).\end{aligned}$$

Puisque $\mathrm{sh}(2t) = 2\,\mathrm{sh}(t)\,\mathrm{ch}(t) = 2\,\mathrm{sh}(t)\sqrt{1+\mathrm{sh}^2(t)}$, on en déduit que

$$\int_1^2 \sqrt{x_1^2+1}\,\mathrm{d}x_1 = \tfrac{1}{2}\,\mathrm{argsh}(2) - \tfrac{1}{2}\,\mathrm{argsh}(1) + \sqrt{5} - \tfrac{1}{2}\sqrt{2}$$

puis que $\displaystyle\int_{\mathrm{P}} f(x_1,x_2)\,\mathrm{d}x_1\mathrm{d}x_2 = \mathrm{argsh}(2) - \mathrm{argsh}(1) + 2\sqrt{5} - \sqrt{2} - 3$.

Nous avions mentionné que le principe du calcul d'une intégrale double sur un pavé en ayant recours au théorème de Fubini consiste à écrire l'intégrale de telle sorte que l'on soit amené à calculer deux intégrales simples. La maîtrise des techniques de calcul de l'intégrale simple, qui font l'objet d'un chapitre du *Cours de première année*, est donc indispensable pour mener à bien le calcul d'intégrales multiples comme l'illustre cet exemple.

Pour une application f continue sur un pavé P de $\mathbb{R}^2$ et à valeurs réelles on a, d'après le théorème de Fubini,

$$\begin{aligned}\int_{\mathrm{P}} f(x_1,x_2)\,\mathrm{d}x_1\mathrm{d}x_2 &= \int_{a_1}^{b_1}\left(\int_{a_2}^{b_2} f(x_1,x_2)\,\mathrm{d}x_2\right)\mathrm{d}x_1 \\ &= \int_{a_2}^{b_2}\left(\int_{a_1}^{b_1} f(x_1,x_2)\,\mathrm{d}x_1\right)\mathrm{d}x_2\end{aligned}$$

et le calcul de l'intégrale de f sur P peut être mené :

- soit en intégrant d'abord f sur $[a_2,b_2]$ relativement à x_2 puis en intégrant l'expression obtenue (qui dépend de x_1) relativement à x_1 sur $[a_1,b_1]$;
- soit en intégrant d'abord f sur $[a_1,b_1]$ relativement à x_1 puis en intégrant l'expression obtenue (qui dépend de x_2) relativement à x_2 sur $[a_2,b_2]$.

[29] On utilise ici la relation $\mathrm{ch}(2t) = 2\,\mathrm{ch}^2(t) - 1$ valable pour tout réel t.

On sera attentif au fait que si les deux approches doivent fournir le même résultat, elles ne sont pas toujours comparables en terme de calculs : l'une peut nécessiter beaucoup plus de calculs que l'autre, voire peut ne pas aboutir du point de vue calculatoire.

Reprenons pour illustrer cette remarque l'exemple précédent et calculons l'intégrale sur le pavé $\mathrm{P} = [1,2]\times[0,1]$ de l'application $f : (x_1, x_2) \in \mathrm{P} \mapsto (x_1^2+x_2)^{-\frac{1}{2}}$ en procédant de la manière suivante :

$$\int_{\mathrm{P}} f(x_1,x_2)\,\mathrm{d}x_1\mathrm{d}x_2 = \int_0^1 \left(\int_1^2 \frac{1}{\sqrt{x_1^2+x_2}}\,\mathrm{d}x_1\right)\,\mathrm{d}x_2 = \int_0^1 I_{x_2}\,\mathrm{d}x_2$$

où pour $x_2 \in]0,1]$ fixé, on a

$$I_{x_2} = \int_1^2 \frac{1}{\sqrt{x_1^2+x_2}}\,\mathrm{d}x_1 = \int_1^2 \frac{1}{\sqrt{\left(x_1/\sqrt{x_2}\right)^2+1}}\,\frac{1}{\sqrt{x_2}}\,\mathrm{d}x_1.$$

En effectuant le changement de variables $t = x_1/\sqrt{x_2}$, on obtient

$$I_{x_2} = \int_{\frac{1}{\sqrt{x_2}}}^{\frac{1}{\sqrt{x_2}}} \frac{1}{\sqrt{t^2+1}}\,\mathrm{d}t = \operatorname{argsh}\left(\frac{2}{\sqrt{x_2}}\right) - \operatorname{argsh}\left(\frac{1}{\sqrt{x_2}}\right),$$

mais la poursuite du calcul de $\int_{\mathrm{P}} f(x_1,x_2)\,\mathrm{d}x_1\mathrm{d}x_2$ nécessite de déterminer une primitive de la fonction $x_2 \mapsto \operatorname{argsh}\left(1/\sqrt{x_2}\right)$ ce qui est loin d'être aisé[30].

Exercice 12.2 Calculer l'intégrale sur le pavé $\mathrm{P} = [0,1]\times[0,1]$ de l'application $f : (x_1,x_2) \in \mathrm{P} \mapsto \sqrt{x_1^2+x_2}$.

Remarque Lorsqu'une application continue f définie sur un pavé $\prod_{i=1}^n [a_i,b_i]$ de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}$ est « à variables séparées », c'est-à-dire lorsqu'elle peut s'exprimer sous la forme $f(x_1,\ldots,x_n) = g_1(x_1)\times\cdots\times g_n(x_n)$, le théorème de Fubini permet d'écrire :

$$\int_{\mathrm{P}} f(x_1,\ldots,x_n)\,\mathrm{d}x_1\cdots\mathrm{d}x_n = \left(\int_{a_1}^{b_1} g_1(x_1)\,\mathrm{d}x_1\right)\times\cdots\times\left(\int_{a_n}^{b_n} g_n(x_n)\,\mathrm{d}x_n\right).$$

On se ramène donc au calcul de n intégrales simples indépendantes les unes des autres. ◇

[30] En se montrant un petit peu astucieux, ou en ayant recours à MAPLE, on peut établir qu'une primitive est $x_2 \mapsto x_2 \operatorname{argsh}(1/\sqrt{x_2}) + \sqrt{1+x_2}$.

12.2.2 Intégration par pile et par tranche

Le théorème de Fubini admet une généralisation au cas où l'on intègre une application continue f sur un ensemble X qui, pour un entier k tel que $1 \leqslant k \leqslant n$, s'exprime sous la forme

$$X = \Big\{ x = (x_1, ..., x_n) \in \mathbb{R}^n \ ; \ \widetilde{x} = (x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n) \in \widetilde{X} \quad \text{et} \quad g(\widetilde{x}) \leqslant x_k \leqslant h(\widetilde{x}) \Big\}$$

où $\widetilde{X}$ désigne un sous-ensemble de $\mathbb{R}^{n-1}$ et où g et h sont deux applications continues de $\widetilde{X}$ dans $\mathbb{R}$. On a alors l'intégrale de f sur X qui s'exprime comme

$$\int_X f(x)\, \mathrm{d}x = \int_{\widetilde{X}} \left(\int_{g(\widetilde{x})}^{h(\widetilde{x})} f(x_1, ..., x_n)\, \mathrm{d}x_k \right) \mathrm{d}x_1 \cdots \mathrm{d}x_{k-1} \mathrm{d}x_{k+1} \cdots \mathrm{d}x_{n-1}.$$

Cas des intégrales doubles

Dans le cas $n = 2$, le résultat précédent est fréquemment utilisé pour calculer des intégrales doubles sur des sous-ensembles de $\mathbb{R}^2$ de la forme[31]

$$X = \Big\{ (x_1, x_2) \in \mathbb{R}^2 \ ; \ a_1 \leqslant x_1 \leqslant b_1 \quad \text{et} \quad g_2(x_1) \leqslant x_2 \leqslant h_2(x_1) \Big\},$$

ou de la forme[32]

$$X = \Big\{ (x_1, x_2) \in \mathbb{R}^2 \ ; \ a_2 \leqslant x_2 \leqslant b_2 \quad \text{et} \quad g_1(x_2) \leqslant x_1 \leqslant h_1(x_2) \Big\},$$

où a_1, a_2, b_1 et b_2 sont des réels et où g_1, g_2, h_1 et h_2 sont des fonctions continues de $\mathbb{R}$ dans $\mathbb{R}$.

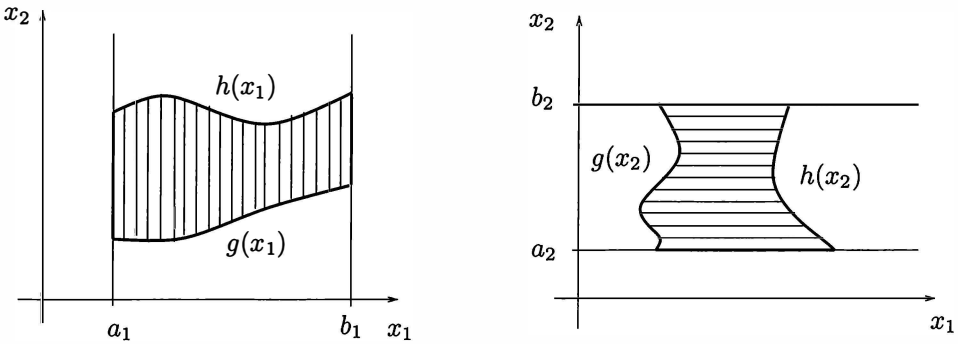


Fig. 11 Sous-ensembles de $\mathbb{R}^2$ décrits par pile (à gauche) et par tranche (à droite).

[31] Un tel ensemble est dit être décrit par pile.

[32] Un tel ensemble est dit être décrit par tranche.

Dans le premier cas (correspondant à la figure 11 à gauche), on dit que l'on intègre par pile et on écrit :

$$\int_X f = \int_{a_1}^{b_1} \left(\int_{g_2(x_1)}^{h_2(x_1)} f(x_1, x_2)\, \mathrm{d}x_2 \right) \mathrm{d}x_1,$$

et dans le second (correspondant à la figure 11 à droite) on dit que l'on intègre par tranche et on écrit :

$$\int_X f = \int_{a_2}^{b_2} \left(\int_{g_1(x_2)}^{h_1(x_2)} f(x_1, x_2)\, \mathrm{d}x_1 \right) \mathrm{d}x_2.$$

Exemples

1. Le triangle $T = \{(x_1, x_2) \in \mathbb{R}^2 \ ; \ 0 \leqslant x_1 \leqslant 1, \ 0 \leqslant x_2 \leqslant 1, \ x_1 + x_2 \leqslant 1\}$ est décrit par pile en prenant $a_1 = 0$, $b_1 = 1$, $g_2 : x_1 \mapsto 0$ et $h_2 : x_1 \mapsto 1 - x_1$.

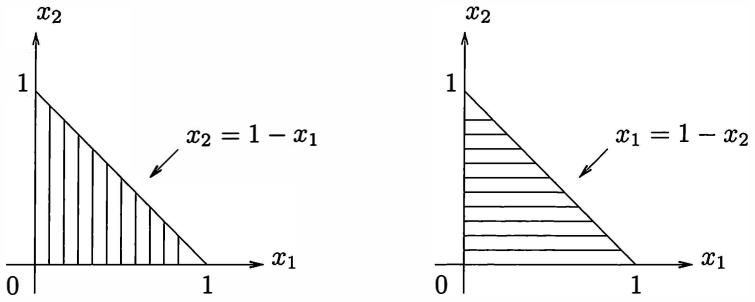


Fig. 12 Découpage du triangle T par pile (à gauche) et par tranche (à droite).

On intègre par pile l'application $f : (x_1, x_2) \in T \mapsto x_1 x_2 \in \mathbb{R}$ sur T de la façon suivante[33], voir la figure 12 :

$$\begin{aligned}
\int_T f(x_1, x_2)\, \mathrm{d}x_1 \mathrm{d}x_2 &= \int_0^1 \left(\int_0^{1-x_1} x_1 x_2\, \mathrm{d}x_2 \right) \mathrm{d}x_1 = \int_0^1 x_1 \left[\tfrac{1}{2} x_2^2 \right]_0^{1-x_1} \mathrm{d}x_1 \\
&= \frac{1}{2} \int_0^1 x_1 (1 - x_1)^2\, \mathrm{d}x_1 = \frac{1}{2} \int_0^1 \left(x_1^3 - 2x_1^2 + x_1 \right) \mathrm{d}x_1 \\
&= \frac{1}{2} \left[\tfrac{1}{4} x_1^4 - \tfrac{2}{3} x_1^3 + \tfrac{1}{2} x_1^2 \right]_0^1 = \frac{1}{24}.
\end{aligned}$$

L'ensemble T peut également être décrit par tranche en prenant $a_2 = 0$, $b_2 = 1$, $g_1 : x_2 \mapsto 0$ et $h_1 : x_2 \mapsto 1 - x_2$. On obtient dans ce cas,

$$\int_T f(x_1, x_2)\, \mathrm{d}x_1 \mathrm{d}x_2 = \int_0^1 \left(\int_0^{1-x_2} x_1 x_2\, \mathrm{d}x_1 \right) \mathrm{d}x_2 = \frac{1}{2} \int_0^1 x_2 (1 - x_2)^2\, \mathrm{d}x_2 = \frac{1}{24}.$$

[33] On comparera cette méthode à celle utilisée lors de l'exemple de la p. 866 pour calculer cette même intégrale double.

2. Calculons l'intégrale double $I = \int_D (x_1^2 + x_2^2)\, \mathrm{d}x_1 \mathrm{d}x_2$ où D est le disque de $\mathbb{R}^2$ de centre l'origine et de rayon 1 en utilisant une intégration par pile. Le disque D peut être décrit par pile de la façon suivante, voir la figure 13 :

$$D = \left\{ (x_1, x_2) \in \mathbb{R}^2 \ ; \ -1 \leqslant x_1 \leqslant 1 \quad \text{et} \quad -\sqrt{1 - x_1^2} \leqslant x_2 \leqslant \sqrt{1 - x_1^2} \right\}$$

de sorte que l'on a $I = \displaystyle\int_{-1}^{1} \left(\int_{-\sqrt{1-x_1^2}}^{\sqrt{1-x_1^2}} (x_1^2 + x_2^2)\, \mathrm{d}x_2 \right) \mathrm{d}x_1$.

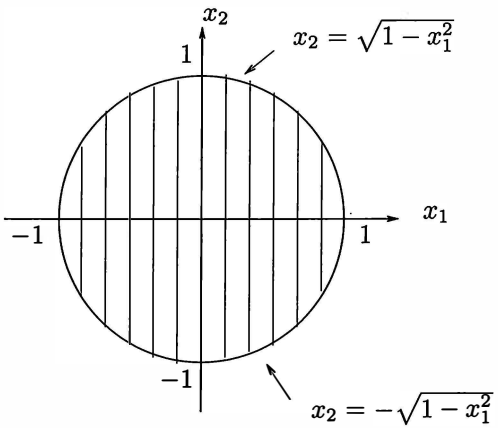


Fig. 13 Découpage du disque D par pile.

Pour tout $x_1 \in [-1, 1]$, on a

$$\int_0^{\sqrt{1-x_1^2}} (x_1^2 + x_2^2)\, \mathrm{d}x_2 = \left[x_1^2 x_2 + \tfrac{1}{3} x_2^3 \right]_{x_2=0}^{x_2=\sqrt{1-x_1^2}} = x_1^2 \sqrt{1 - x_1^2} + \tfrac{1}{3}(1 - x_1^2)^{3/2}.$$

Compte tenu du fait que la fonction $x_2 \mapsto x_1^2 + x_2^2$ est paire, on obtient

$$I = \int_{-1}^{1} \left(2x_1^2 \sqrt{1 - x_1^2} + \tfrac{2}{3}(1 - x_1^2)^{3/2} \right) \mathrm{d}x_1.$$

Pour calculer cette intégrale d'une variable réelle, on utilise le changement de variables $x_1 = \cos(\theta)$. L'application $\phi : \theta \in [0, \pi] \mapsto \cos(\theta) \in [-1, 1]$ est de classe C^1 sur $]0, \pi[$ et la formule de changement de variables[34] indique que

$$\begin{aligned} I &= \int_{\phi(\pi)}^{\phi(0)} \left(2x_1^2 \sqrt{1 - x_1^2} + \tfrac{2}{3}(1 - x_1^2)^{3/2} \right) \mathrm{d}x_1 \\ &= \int_{\pi}^{0} \left(2\,\phi(\theta)^2 \sqrt{1 - \phi(\theta)^2} + \tfrac{2}{3}(1 - \phi(\theta)^2)^{3/2} \right) \phi'(\theta)\, \mathrm{d}\theta. \end{aligned}$$

[34] Voir le théorème 18.2 p. 863 du *Cours de première année*.

On a $1-\phi(\theta)^2 = 1-\cos^2(\theta) = \sin^2(\theta)$ et pour $\theta \in [0,\pi]$, puisque $\sin(\theta) \geqslant 0$, on a $\sqrt{1-\phi(\theta)^2} = \sin(\theta)$. Ainsi,

$$\begin{aligned} I &= -\int_{\pi}^{0} \left(2\cos^2(\theta)\sin(\theta) + \tfrac{2}{3}\sin^3(\theta)\right)\sin(\theta)\ \mathrm{d}\theta \\ &= \int_{0}^{\pi} \left(2\cos^2(\theta)\sin^2(\theta) + \tfrac{2}{3}\sin^4(\theta)\right)\ \mathrm{d}\theta. \end{aligned}$$

En « linéarisant » les applications $\theta \mapsto \cos^2(\theta)\sin^2(\theta)$ et $\sin^4(\theta)$, on obtient[35] :

$$I = \int_0^{\pi} \left(\tfrac{1}{2} - \tfrac{1}{6}\cos(4\theta) - \tfrac{1}{3}\cos(2\theta)\right)\ \mathrm{d}\theta = \left[\tfrac{1}{2}\theta - \tfrac{1}{24}\sin(4\theta) - \tfrac{1}{6}\sin(2\theta)\right]_0^{\pi} = \frac{\pi}{2}.$$

Nous verrons ultérieurement[36] qu'un changement de variables effectué sur l'intégrale double $\int_D (x_1^2 + x_2^2)\ \mathrm{d}x_1\mathrm{d}x_2$ permet d'obtenir ce résultat beaucoup plus simplement.

3. Calculons l'intégrale double sur le rectangle D de $\mathbb{R}^2$ de sommets $(-1,0)$, $(0,1)$, $(1,-2)$ et $(2,-1)$ de l'application $f : (x_1,x_2) \in D \mapsto \mathrm{e}^{x_1+x_2}$. Pour cela « découpons » D en trois morceaux D_1, D_2 et D_3 où D_1 est le triangle de sommets $(-1,0)$, $(0,-1)$ et $(0,1)$, D_2 est le losange de sommets $(0,-1)$, $(1,-2)$ $(1,0)$ et $(0,1)$ et D_3 est le triangle de sommets $(1,0)$, $(2,-1)$ et $(1,-2)$, voir la figure 14.

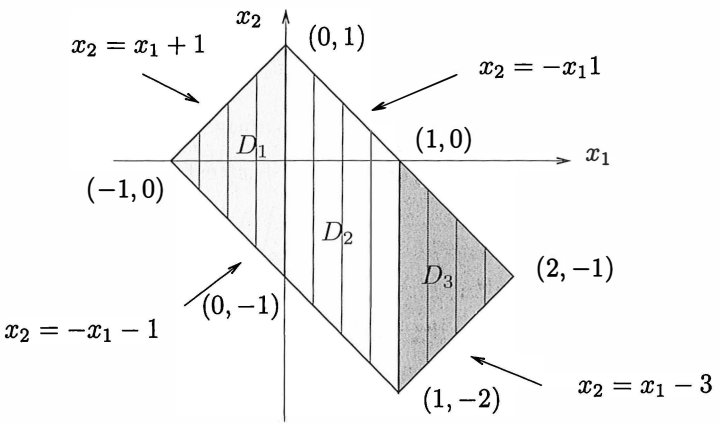


Fig. 14 Découpage du rectangle D pour une intégration par pile.

[35] On « linéarise » ces applications en ayant recours à la formule d'Euler et à la formule du binôme de Newton en suivant la méthode présentée dans le *Cours de première année* p. 158. On obtient $\cos^2(\theta)\sin^2(\theta) = \frac{1}{8} - \frac{1}{8}\cos(4\theta)$ et $\sin^4(\theta) = \frac{3}{8} + \frac{1}{8}\cos(4\theta) - \frac{1}{2}\cos(2\theta)$.

[36] Voir p. 902.

La relation de Chasles[37] permet d'écrire $\int_D f = \int_{D_1} f + \int_{D_2} f + \int_{D_3} f$.
On évalue chacune des trois intégrales séparément en intégrant par pile ; on obtient successivement

$$\begin{aligned}
\int_{D_1} f &= \int_{-1}^{0} \left(\int_{-1-x_1}^{1+x_1} \mathrm{e}^{x_1+x_2} \, \mathrm{d}x_2 \right) \mathrm{d}x_1 = \int_{-1}^{0} \Big[\mathrm{e}^{x_1+x_2} \Big]_{x_2=-1-x_1}^{x_2=1+x_1} \mathrm{d}x_1 \\
&= \int_{-1}^{0} (\mathrm{e}^{1+2x_1} - \mathrm{e}^{-1}) \, \mathrm{d}x_1 = \Big[\tfrac{1}{2} \mathrm{e}^{1+2x_1} - \mathrm{e}^{-1} x_1 \Big]_{-1}^{0} \\
&= \tfrac{1}{2}(\mathrm{e} + \mathrm{e}^{-1}) = \mathrm{ch}(1) \\
\int_{D_2} f &= \int_{0}^{1} \left(\int_{-1-x_1}^{1-x_1} \mathrm{e}^{x_1+x_2} \, \mathrm{d}x_2 \right) \mathrm{d}x_1 = \int_{0}^{1} \Big[\mathrm{e}^{x_1+x_2} \Big]_{x_2=-1-x_1}^{x_2=1-x_1} \mathrm{d}x_1 \\
&= \int_{0}^{1} (\mathrm{e} - \mathrm{e}^{-1}) \, \mathrm{d}x_1 = \mathrm{e} - \mathrm{e}^{-1} = 2\,\mathrm{sh}(1) \\
\int_{D_3} f &= \int_{1}^{2} \left(\int_{x_1-3}^{1-x_1} \mathrm{e}^{x_1+x_2} \, \mathrm{d}x_2 \right) \mathrm{d}x_1 = \int_{1}^{2} \Big[\mathrm{e}^{x_1+x_2} \Big]_{x_2=x_1-3}^{x_2=1-x_1} \mathrm{d}x_1 \\
&= \int_{1}^{2} (\mathrm{e} - \mathrm{e}^{-3+2x_1}) \, \mathrm{d}x_1 = \Big[\mathrm{e}\, x_1 - \tfrac{1}{2} \mathrm{e}^{-3+2x_1} \Big]_{1}^{2} \\
&= \tfrac{1}{2}(\mathrm{e} + \mathrm{e}^{-1}) = \mathrm{ch}(1).
\end{aligned}$$

On en conclut que $\int_D f = 2\,\mathrm{sh}(1) + 2\,\mathrm{ch}(1) = 2\mathrm{e}$.

Exercice 12.3 Calculer l'aire du disque de $\mathbb{R}^2$ de centre $(0,0)$ et de rayon $R \in \mathbb{R}_+^*$, autrement dit, calculer l'intégrale double $I = \int_X 1 \, \mathrm{d}x_1 \mathrm{d}x_2$ où $X = \{(x_1, x_2) \in \mathbb{R}^2 \ ; \ x_1^2 + x_2^2 \leqslant R^2\}$, en utilisant une intégration par pile.

Avant de se lancer dans un calcul d'intégrale double, il convient de prendre le temps de comparer les deux approches calculatoires que sont l'intégration par pile et l'intégration par tranche. Les calculs peuvent être plus ou moins rapides selon la méthode employée.

Par exemple, pour calculer le moment d'inertie d'un solide homogène ayant la forme du triangle T de sommets $(0,1)$, $(2,1)$ et $(1,0)$ par rapport à un axe perpendiculaire au triangle passant par l'origine, il est préférable d'effectuer une intégration par tranche qu'une intégration par pile. Calculer le moment d'inertie de T revient à calculer l'intégrale double $I = \int_T (x_1^2 + x_2^2) \, \mathrm{d}x_1 \mathrm{d}x_2$. En intégrant par tranche, on obtient

$$\begin{aligned}
I &= \int_0^1 \left(\int_{1-x_2}^{1+x_2} (x_1^2 + x_2^2) \, \mathrm{d}x_1 \right) \mathrm{d}x_2 = \int_0^1 \Big[\tfrac{1}{3} x_1^3 + x_2^2 x_1 \Big]_{1-x_2}^{1+x_2} \mathrm{d}x_2 \\
&= \int_0^1 2x_2 + \tfrac{8}{3} x_2^3 \, \mathrm{d}x_2 = \Big[x_2^2 + \tfrac{2}{3} x_2^4 \Big]_0^1 = \frac{5}{3}.
\end{aligned}$$

[37] Voir la proposition 12.7 p. 873.

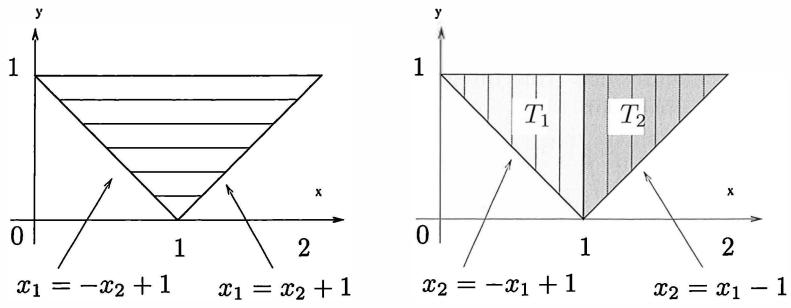


Fig. 15 Intégration par tranche (à gauche) et par pile (à droite).

Pour pouvoir intégrer par pile, il convient de découper le triangle T en deux triangles T_1 et T_2 et d'utiliser la relation de Chasles. On a $I = I_1 + I_2$ où

$$I_1 = \int_{T_1} (x_1^2 + x_2^2)\, \mathrm{d}x_1\, \mathrm{d}x_2 = \int_0^1 \left(\int_{1-x_1}^1 (x_1^2 + x_2^2)\, \mathrm{d}x_2 \right) \mathrm{d}x_1$$

$$\text{et } I_2 = \int_{T_2} (x_1^2 + x_2^2)\, \mathrm{d}x_1\, \mathrm{d}x_2 = \int_1^2 \left(\int_{x_1-1}^1 (x_1^2 + x_2^2)\, \mathrm{d}x_2 \right) \mathrm{d}x_1.$$

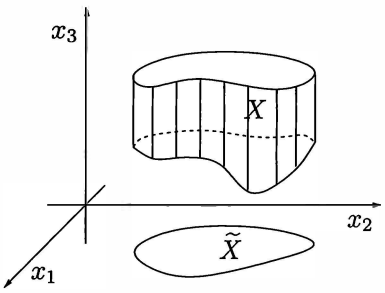
On a donc deux fois plus de calculs que dans le cas où l'intégration est effectuée par tranche.

Cas des intégrales triples

✗ Dans le cas où $n = 3$, si X est un sous-ensemble de $\mathbb{R}^3$ défini par

$$X = \Big\{ (x_1, x_2, x_3) \in \mathbb{R}^3, \quad (x_1, x_2) \in \widetilde{X} \\ \text{et} \quad g(x_1, x_2) \leqslant x_3 \leqslant h(x_1, x_2) \Big\}$$

où $\widetilde{X}$ est un sous-ensemble de $\mathbb{R}^2$, g et h sont deux fonctions de $\widetilde{X}$ dans $\mathbb{R}$, on a



$$\int_X f(x_1, x_2, x_3)\, \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}x_3 = \int_{\widetilde{X}} \left(\int_{g(x_1,x_2)}^{h(x_1,x_2)} f(x_1, x_2, x_3)\, \mathrm{d}x_3 \right) \mathrm{d}x_1 \mathrm{d}x_2$$

et on dit qu'on fait une *intégration par pile* dans la direction $\mathbf{e}_3$ (ou selon x_3).

Exemples

1. En utilisant une intégration par pile, déterminons le volume V d'un cône X de rayon R et de hauteur H. Considérons un repère orthonormé direct $(O; \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ où le vecteur $\mathbf{e}_3$ coïncide avec l'axe du cône et l'origine O est le centre de la base du cône. Le volume V du cône est obtenu en calculant l'intégrale triple $\int_X 1\, \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}x_3$.

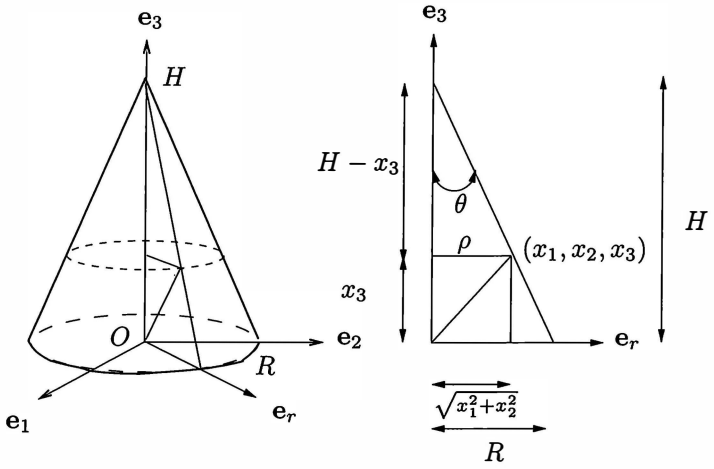


Fig. 16 Description par pile d'un cône.

Pour un point (x_1, x_2, x_3) appartenant à la surface latérale du cône, on a $(x_1, x_2) \in D(0, R)$ où $D(0, R) = \{(x_1, x_2) \in \mathbb{R}^2 \ ; \ x_1^2 + x_2^2 \leqslant R^2\}$ et, voir la figure 16,

$$\tan(\theta) = \frac{R}{H} = \frac{\rho}{H - x_3} \quad \text{où} \quad \rho = \sqrt{x_1^2 + x_2^2}.$$

On en déduit que x_3 est donné par $x_3 = H - \frac{H}{R^2}\sqrt{x_1^2 + x_2^2}$. Ainsi, le cône est défini par

$$X = \left\{(x_1, x_2, x_3) \in \mathbb{R}^3 \ ; \ (x_1, x_2) \in D(0, R) \quad \text{et} \quad 0 \leqslant x_3 \leqslant h(x_1, x_2)\right\}$$

avec $h : (x_1, x_2) \mapsto H - \frac{H}{R^2}\sqrt{x_1^2 + x_2^2}$. On obtient en intégrant par pile :

$$\begin{aligned} V &= \int_X 1 \, \mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3 = \int_{D(0,R)} \left(\int_0^{h(x_1,x_2)} 1 \, \mathrm{d}x_3 \right) \mathrm{d}x_1\mathrm{d}x_2 \\ &= \int_{D(0,R)} h(x_1, x_2) \, \mathrm{d}x_1\mathrm{d}x_2 \\ &= H \int_{D(0,R)} 1 \, \mathrm{d}x_1\mathrm{d}x_2 - \frac{H}{R} \int_{D(0,R)} \sqrt{x_1^2 + x_2^2} \, \mathrm{d}x_1\mathrm{d}x_2. \end{aligned}$$

La première intégrale s'interprète comme l'aire du disque centré à l'origine et de rayon R et vaut πR^2. Nous verrons ultérieurement que la seconde intégrale se calcule facilement en ayant recours à la formule de changement de variables et qu'elle vaut $\frac{2}{3}\pi R^3$ (voir l'exercice 12.6). Le volume du cône vaut donc

$$V = H\pi R^2 - \frac{H}{R}\,\frac{2}{3}\pi R^3 = \frac{1}{3}\pi R^2 H.$$

2. Calculons, en utilisant une intégration par pile, le volume[(38)] d'une « canette » cylindrique de hauteur H et de rayon R dont la base bombée correspond à une calotte sphérique de hauteur $h \in]0, R[$. Considérons un repère orthonormé direct $(O; \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ où le vecteur $\mathbf{e}_3$ coïncide avec l'axe du cylindre et l'origine O est le centre de la base du cylindre, voir la figure 17.

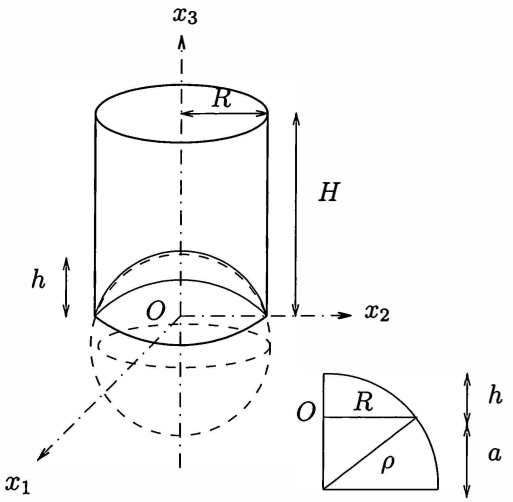


Fig. 17 Canette cylindrique de hauteur H et de rayon R dont la base bombée correspond à une calotte sphérique de hauteur h.

La calotte sphérique formant la base de la canette est issue de la sphère de centre $(0, 0, -a)$ et de rayon ρ où les réels positifs a et ρ vérifient :

$$a = \rho - h, \quad \rho^2 = a^2 + R^2.$$

En résolvant ces équations, on obtient $\rho = \frac{R^2+h^2}{2h}$ et $a = \frac{R^2-h^2}{2h}$. L'équation de cette sphère de centre $(0, 0, -a)$ et de rayon ρ est

$$x_1^2 + x_2^2 + (x_3 + a)^2 = \rho^2.$$

On a donc pour la demi-sphère supérieure $x_3 = -a + \sqrt{\rho^2 - (x_1^2 + x_2^2)}$.
Le volume V de la canette X est obtenu en calculant l'intégrale triple $V = \int_X 1 \, \mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3$. Utilisons une intégration par pile : on décrit l'ensemble X en faisant varier (x_1, x_2) dans le disque

$$D(0, R) = \left\{ (x_1, x_2) \in \mathbb{R}^2 \ ; \ x_1^2 + x_2^2 \leqslant R^2 \right\} = \widetilde{X}$$

et pour (x_1, x_2) fixé dans $\widetilde{X}$ en faisant varier x_3 entre

$$g(x_1, x_2) = -a + \sqrt{\rho^2 - (x_1^2 + x_2^2)} \quad \text{et} \quad h(x_1, x_2) = H.$$

(38) Le volume de cette canette peut également être obtenue en retranchant au volume du cylindre le volume de la calotte sphérique, voir l'exercice 12.4 p. 892.

On a donc

$$\begin{aligned} V &= \int_{\tilde{X}} \left(\int_{g(x_1,x_2)}^{H} 1 \, \mathrm{d}x_3 \right) \mathrm{d}x_1\mathrm{d}x_2 = \int_{D(0,R)} \big(H - g(x_1,x_2)\big) \, \mathrm{d}x_1\mathrm{d}x_2 \\ &= \int_{D(0,R)} (H+a) \, \mathrm{d}x_1\mathrm{d}x_2 - \int_{D(0,R)} \sqrt{\rho^2 - (x_1^2+x_2^2)} \, \mathrm{d}x_1\mathrm{d}x_2. \end{aligned}$$

Les deux intégrales se calculent de façon distinctes. On a d'une part

$$\int_{D(0,R)} (H+a) \, \mathrm{d}x_1\mathrm{d}x_2 = (H-a) \int_{D(0,R)} 1 \, \mathrm{d}x_1\mathrm{d}x_2 = \pi R^2 (H-a)$$

car $\int_{D(0,R)} 1 \, \mathrm{d}x_1\mathrm{d}x_2$ s'interprète comme l'aire du disque $D(0,R)$. Nous établirons en utilisant la formule de changement de variables (voir p. 902) que

$$\int_{D(0,R)} \sqrt{\rho^2 - (x_1^2+x_2^2)} \, \mathrm{d}x_1\mathrm{d}x_2 = \tfrac{2}{3}\pi \left(\rho^3 - (\rho^2 - R^2)^{\frac{3}{2}} \right).$$

On en conclut que la volume de la canette est

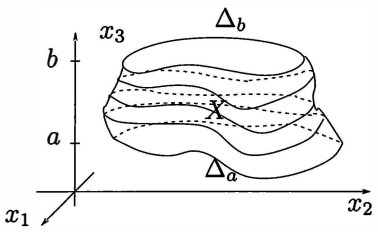
$$V = \pi R^2 (H+a) - \tfrac{2}{3}\pi \left(\rho^3 - (\rho^2 - R^2)^{\frac{3}{2}} \right) = \pi R^2 (H+a) - \tfrac{2}{3}\pi \left(\rho^3 - a^3 \right)$$

où $\rho = \frac{R^2+h^2}{2h}$ et $a = \frac{R^2-h^2}{2h}$. On peut déterminer l'expression du volume en fonction des seuls données H, h et R ; on obtient $V = \frac{\pi}{6}\left(3R^2(2H-h) - h^3\right)$.

Remarque Sur le même principe que celui qui vient d'être présenté, si la géométrie du domaine X s'y prête, on peut également avoir recours à une intégration par pile dans la direction $\mathbf{e}_2$ (ou selon x_2) ou encore dans la direction $\mathbf{e}_1$ (ou selon x_1). ◇

✗Dans le cas où $n = 3$, si X est un sous-ensemble de $\mathbb{R}^3$ défini par

$$X = \Big\{ (x_1,x_2,x_3) \in \mathbb{R}^3, \quad a \leqslant x_3 \leqslant b \quad \text{et} \quad (x_1,x_2) \in \Delta_{x_3} \Big\}$$



où Δ_{x_3} désigne un sous-ensemble de $\mathbb{R}^2$ paramétré par x_3, alors on a

$$\int_X f(x_1,x_2,x_3) \, \mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3 = \int_a^b \left(\int_{\Delta_{x_3}} f(x_1,x_2,x_3) \, \mathrm{d}x_1\mathrm{d}x_2 \right) \mathrm{d}x_3$$

et on dit qu'on fait une *intégration par tranche* dans la direction $\mathbf{e}_3$ (ou selon x_3).

Exemples

1. Calculons l'intégrale triple $\int_X x_3 \, dx_1 dx_2 dx_3$ lorsque X désigne la demi-boule de $\mathbb{R} \times \mathbb{R} \times \mathbb{R}^+$ de rayon $R \in \mathbb{R}_+^*$ et de centre $(0, 0, 0)$, voir la figure 18. On a

$$X = \left\{ (x_1, x_2, x_3) \in \mathbb{R}^3 \ ; \ x_3 \geqslant 0 \ \text{ et } \ x_1^2 + x_2^2 + x_3^2 = R^2 \right\}$$

et on peut décrire X par tranche de la manière suivante :

$$X = \left\{ (x_1, x_2, x_3) \in \mathbb{R}^3 \ ; \ 0 \leqslant x_3 \leqslant R \ \text{ et } \ (x_1, x_2) \in \Delta_{x_3} \right\},$$

où pour $x_3 \in [0, R]$ fixé l'ensemble Δ_{x_3} est défini par

$$\Delta_{x_3} = \left\{ (x_1, x_2) \in \mathbb{R}^2 \ ; \ x_1^2 + x_2^2 \leqslant R^2 - x_3^2 \right\}.$$

L'ensemble Δ_{x_3} est donc un disque de $\mathbb{R}^2$ de centre $(0, 0)$ et de rayon $\sqrt{R^2 - x_3^2}$.

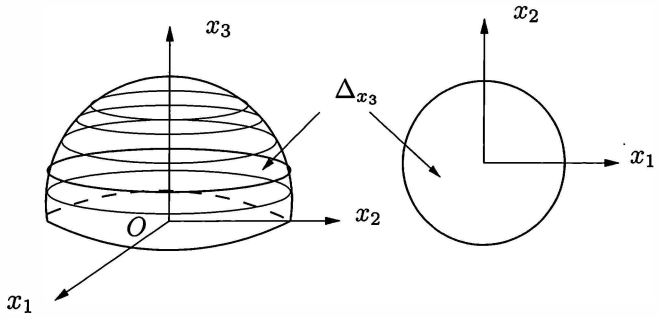


Fig. 18 Description d'une demi-boule par tranche.

En utilisant une intégration par tranche, on obtient

$$I = \int_X x_3 \, dx_1 dx_2 dx_3 = \int_0^R \left(\int_{\Delta_{x_3}} x_3 \, dx_1 dx_2 \right) dx_3$$
$$= \int_0^R x_3 \left(\int_{\Delta_{x_3}} 1 \, dx_1 dx_2 \right) dx_3.$$

L'intégrale $\int_{\Delta_{x_3}} 1 \, dx_1 dx_2$ s'interprète comme l'aire du disque Δ_{x_3} de centre $(0, 0)$ et de rayon $\sqrt{R^2 - x_3^2}$; elle vaut donc $\pi(R^2 - x_3^2)$. On a donc

$$I = \int_0^R \pi \, x_3 (R^2 - x_3^2) \, dx_3 = \pi \left[\tfrac{1}{2} R^2 x_3^2 - \tfrac{1}{4} x_3^4 \right]_0^R = \frac{1}{4} R^4.$$

Si la demi-boule représente un solide formé d'un matériau homogène, la valeur de cette intégrale divisée par le volume de la demi-boule ($\frac{2}{3}\pi R^3$) fournit la troisième coordonnée x_3^G de son centre d'inertie G, voir p. 875 ; on a donc $x_3^G = \frac{3}{8\pi} R$. Pour des raisons de symétrie, le centre d'inertie de la demi-boule est situé sur l'axe $(O; \mathbf{e}_3)$.

2. En utilisant une intégration par tranche, déterminons le volume V d'un cône X de rayon R et de hauteur H, voir la figure 19. Considérons un repère orthonormé direct $(O; \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ où le vecteur $\mathbf{e}_3$ coïncide avec l'axe du cône et l'origine 0 est le centre de la base du cône. La volume V du cône est obtenu en calculant l'intégrale triple $\int_X 1 \, \mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3$.

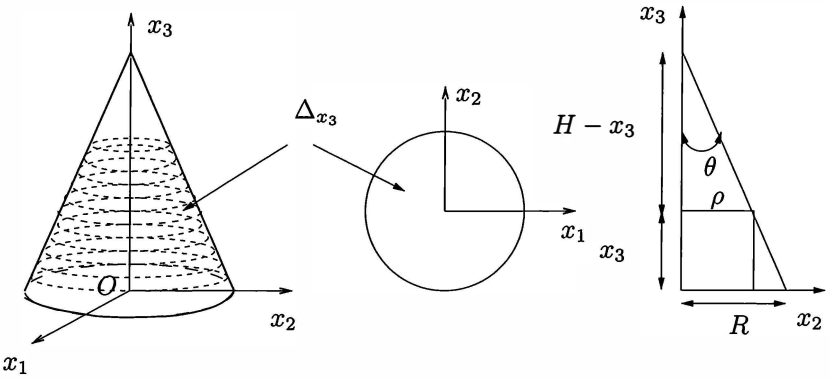


Fig. 19 Description d'un cône par tranche.

On peut décrire le cône X par tranche de la manière suivante :

$$X = \Big\{(x_1, x_2, x_3) \in \mathbb{R}^3 \ ; \ 0 \leqslant x_3 \leqslant H \ \text{ et } \ (x_1, x_2) \in \Delta_{x_3} \subset \mathbb{R}^2\Big\},$$

où pour $x_3 \in [0, R]$ fixé l'ensemble Δ_{x_3} un disque de $\mathbb{R}^2$ de centre $(0, 0)$ et de rayon ρ. La valeur de ρ qui dépend de x_3 peut être obtenue en remarquant que

$$\tan(\theta) = \frac{R}{H} = \frac{\rho}{H - x_3}.$$

On obtient $\rho = R(1 - x_3/H)$ et par conséquent

$$\Delta_{x_3} = \Big\{(x_1, x_2) \in \mathbb{R}^2 \ ; \ x_1^2 + x_2^2 \leqslant R^2(1 - x_3/H)^2\Big\}.$$

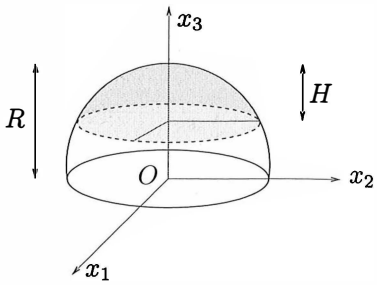
On en déduit que $V = \displaystyle\int_X 1 \, \mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3 = \int_0^H \left(\int_{\Delta_{x_3}} 1 \, \mathrm{d}x_1\mathrm{d}x_2\right) \mathrm{d}x_3.$

L'intégrale $\int_{\Delta_{x_3}} 1 \, \mathrm{d}x_1\mathrm{d}x_2$ s'interprète comme l'aire du disque Δ_{x_3} de centre $(0, 0)$ et de rayon $R(1 - x_3/H)$; elle vaut donc $\pi R^2(1 - x_3/H)^2$. On a donc

$$V = \int_0^H \pi R^2 \left(1 - \frac{x_3}{H}\right)^2 \mathrm{d}x_3 = \left[-\frac{\pi R^2 H}{3} \left(1 - \frac{x_3}{H}\right)^2\right]_0^H = \frac{1}{3}\pi R^2 H.$$

Exercice 12.4

1 - En utilisant une intégration par tranche, déterminer le volume d'une calotte sphérique obtenue dans un repère orthonormé direct $(O;\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3)$ par l'intersection de la boule de centre l'origine et de rayon $R \in \mathbb{R}_+^*$ avec le demi-espace $\{(x_1,x_2,x_3) \in \mathbb{R}^3 \ ; \ x_3 \geqslant H\}$ où H désigne un réel vérifiant $0 < H < R$, voir la figure ci-contre.



2 - En utilisant la relation de Chasles, retrouver en utilisant le résultat de la question précédente l'expression du volume de la canette obtenu p. 888.

Remarque Sur le même principe que celui qui vient d'être présenté, si la géométrie du domaine X s'y prête, on peut également avoir recours à une intégration par tranche dans la direction $\mathbf{e}_2$ (ou selon x_2) ou encore dans la direction $\mathbf{e}_1$ (ou selon x_1). ◇

Exercice 12.5

On considère une louche ayant la forme d'une demi-sphère de rayon R. À quelle hauteur faut-il remplir la louche pour que le volume de liquide soit égal à la moitié du volume total de la louche[39] (autrement dit, pour que la louche soit à moitié remplie) ?



Comme pour l'intégrale double, avant de se lancer dans un calcul d'intégrale triple, il convient de prendre le temps de comparer les deux approches calculatoires que sont l'intégration par pile et l'intégration par tranche. Les calculs peuvent être plus ou moins rapides selon la méthode employée. Ainsi, le volume d'un cône s'obtient relativement aisément en effectuant une intégration par tranche, mais plus difficilement en choisissant une intégration par tranche. Le choix est souvent dicté par la géométrie du domaine sur lequel on intègre.

[39] On pourra vérifier que l'équation polynomiale $t^3 - 3t - 1 = 0$ admet une seule solution supérieure à 1 qui est $2\cos(2\pi/9)$.

12.3 Changement de variables pour une intégrale multiple

12.3.1 Changement de variables dans $\mathbb{R}^n$

La formule du changement de variables pour l'intégrale d'une fonction de $\mathbb{R}$ dans $\mathbb{R}$ s'énonce de la façon suivante[40] :

Soient φ une application de classe $\mathcal{C}^1$ sur l'intervalle $[a,b]$ et f une application continue sur l'intervalle fermé borné $\varphi([a,b])$. On a

$$\int_{\varphi(a)}^{\varphi(b)} f(y)\,\mathrm{d}y = \int_a^b f(\varphi(x))\varphi'(x)\,\mathrm{d}x. \tag{1}$$

On dit que l'on effectue le changement de variables $y = \varphi(x)$ et « formellement » on remplace $\mathrm{d}y$ par $\varphi'(x)\,\mathrm{d}x$. Dans le cas des fonctions de plusieurs variables, un tel changement « intuitif » conduit à des résultats erronés. Considérons par exemple l'intégrale double

$$I = \int_X (x_1^2 + x_2^2)\,\mathrm{d}x_1\mathrm{d}x_2,$$

où X est le disque de $\mathbb{R}^2$ de centre $(0,0)$ et de rayon 1. Si on effectue le changement de variables $y_1 = -x_1$ et $y_2 = x_2$ en remplaçant « intuitivement » $\mathrm{d}x_1$ par $-\mathrm{d}y_1$ et $\mathrm{d}x_2$ par $\mathrm{d}y_2$ alors on obtient[41]

$$I = \int_X (x_1^2 + x_2^2)\,\mathrm{d}x_1\mathrm{d}x_2 = -\int_X (y_1^2 + y_2^2)\,\mathrm{d}y_1\mathrm{d}y_2 = -I$$

ce qui conduit à conclure que $I = 0$. Or, à la p. 883, on a établi en ayant recours à une intégration par pile que $I = \pi/2$.

La formulation correcte de la formule de changement de variables pour l'intégrale multiple est donnée par le théorème 12.2. Rappelons[42] qu'une application φ de U dans V où U et V désignent deux ouverts de $\mathbb{R}^n$ est qualifiée de $\mathcal{C}^1$-difféomorphisme si :

1. φ est de classe $\mathcal{C}^1$ sur U ;
2. φ est une bijection de U sur V ;
3. la bijection réciproque φ^{-1} de φ est de classe $\mathcal{C}^1$ sur V.

Rappelons également qu'en vertu du théorème d'inversion globale[43] une application φ de U dans V réalise un $\mathcal{C}^1$-difféomorphisme si et seulement elle satisfait aux conditions suivantes :

1. l'application φ est de classe $\mathcal{C}^1$ sur U,

---

[40] Voir le *Cours de première année*, p. 863.

[41] On notera que l'ensemble X est invariant par l'application $\varphi : (x_1, x_2) \mapsto (-x_1, x_2)$, i.e. $\varphi(X) = X$.

[42] Voir la définition 10.14 p. 663.

[43] Théorème 10.5 p. 665.

2. l'application φ est bijective de U sur V,

3. le jacobien de φ ne s'annule pas sur U.

L'intérêt du théorème d'inversion globale est qu'il dispense de déterminer la bijection réciproque φ^{-1} de φ et d'établir qu'elle est de classe $\mathcal{C}^1$ sur V. Rappelons enfin que l'application φ est surjective de U sur V si et seulement si $V = \varphi(U)$. Dans la pratique, on détermine donc l'image V de U par φ et on montre que φ est injective sur U pour établir que φ est bijective sur U sur V (condition 2 ci-dessus).

THÉORÈME 12.2 (Formule de changement de variables)
Soient U et V deux ouverts de $\mathbb{R}^n$ et φ un $\mathcal{C}^1$-difféomorphisme de U sur V de jacobien noté $\mathcal{J}_\varphi$. Si $f : x \in V \mapsto f(x) \in \mathbb{R}^n$ est une application continue sur V alors l'application $y \in U \mapsto f(\varphi(y))\,|\mathcal{J}_\varphi(y)|$ est continue sur U et on a

$$\int_V f(x)\ \mathrm{d}x = \int_U f(\varphi(y))\,|\mathcal{J}_\varphi(y)|\ \mathrm{d}y.$$

Remarques

1. La formule de changement de variables s'applique à des ensembles ouverts. Cette condition n'est pas contraignante d'un point de vue pratique car d'après la proposition 12.4 p. 872, la valeur de l'intégrale d'une fonction sur un ensemble X borné et de frontière négligeable est égale à la valeur de l'intégrale de cette fonction sur l'intérieur de X et l'intérieur d'un ensemble est un ouvert. En d'autres termes, si on souhaite appliquer la formule de changement de variables à l'intégrale d'une fonction f sur un ensemble (borné et de frontière négligeable) X qui n'est pas un ouvert, on commence par déterminer l'intérieur V de X et on applique la formule de changement de variables à $\int_V f(x)\ \mathrm{d}x$.

2. La formule de changement de variables est fréquemment utilisée dans l'un des deux cas suivants[44] :

– On cherche à calculer $\int_V f(x)\ \mathrm{d}x$ et l'ensemble V est « trop compliqué » pour que l'on puisse aisément effectuer une intégration par pile ou par tranche. En ayant recours à la formule de changement de variables, on « transforme » donc l'ensemble V sur lequel on doit intégrer en un ensemble U plus « simple » ; en général la fonction à intégrer prend elle une expression plus compliquée que celle que l'on avait initialement.

– On cherche à calculer $\int_V f(x)\ \mathrm{d}x$ mais la fonction f a une expression « trop compliquée ». En ayant recours à la formule de changement de variables, on « transforme » donc la fonction f à intégrer sur V en une nouvelle fonction $y \mapsto f(\varphi(y))\,|\mathcal{J}_\varphi(y)|$ d'expression plus simple que l'on doit intégrer sur l'ensemble $U = \varphi^{-1}(V)$.

[44] Il arrive que les deux cas se produisent simultanément.

Même si toute application φ réalisant un $\mathcal{C}^1$-difféomorphisme de U sur V peut être utilisée dans la formule de changement de variables pour calculer l'intégrale $\int_V f(x)\ \mathrm{d}x$, d'un point de vue pratique seules les applications φ pour lesquelles l'intégrale une fois transformée devient plus aisément calculable présentent évidement un intérêt. ◇

Interprétation de la formule de changement de variables

Considérons une application f définie sur l'ouvert $V =]0,\alpha[\times]0,\beta[$ de $\mathbb{R}^2$ (α et β désignant deux réels strictement positifs) et intéressons-nous à l'intégrale double $\int_V f(x_1,x_2)\ \mathrm{d}x_1\ \mathrm{d}x_2$. Considérons le changement de variables défini par l'application

$$\varphi : (y_1,y_2) \in]0,1[\times]0,1[\longmapsto (\alpha y_1, \beta y_2).$$

L'application φ est de classe $\mathcal{C}^1$ sur $U =]0,1[\times]0,1[$ et son jacobien vaut

$$\mathcal{J}_\varphi(y_1,y_2) = \begin{vmatrix} \dfrac{\partial\varphi_1}{\partial y_1}(y_1,y_2) & \dfrac{\partial\varphi_1}{\partial y_2}(y_1,y_2) \\ \dfrac{\partial\varphi_2}{\partial y_1}(y_1,y_2) & \dfrac{\partial\varphi_2}{\partial y_2}(y_1,y_2) \end{vmatrix} = \begin{vmatrix} \alpha & 0 \\ 0 & \beta \end{vmatrix} = \alpha\beta.$$

L'application φ réalise une bijection de U sur V dont la bijection réciproque $\varphi^{-1} : (x_1,x_2) \in]0,\alpha[\times]0,\beta[\mapsto (x_1/\alpha, x_2/\beta)$ est de classe $\mathcal{C}^1$ sur V. La formule de changement de variables indique que :

$$\int_V f(x_1,x_2)\ \mathrm{d}x_1\ \mathrm{d}x_2 = \int_U f(\varphi(y_1,y_2))\ |J_\varphi(y_1,y_2)|\ \ \mathrm{d}y_1\mathrm{d}y_2.$$

Interprétons cette relation. Le point (x_1,x_2) qui parcourt V pour définir l'intégrale $\int_V f(x_1,x_2)\ \mathrm{d}x_1\ \mathrm{d}x_2$ est transformé par la fonction φ de changement de variables en un point $(y_1,y_2) = \varphi^{-1}(x_1,x_2)$ qui parcourt U ; lorsque le point (x_1,x_2) de V prend la valeur $m = f(x_1,x_2)$, le point (y_1,y_2) de U qui lui correspond prend la valeur $f(\varphi(y_1,y_2))$ qui est égale à m puisque $(x_1,x_2) = \varphi(y_1,y_2)$. Ce point parcourt donc un autre domaine d'intégration, voir la figure 20, mais les valeurs prises par f et $f\circ\phi$ sont identiques.

Intéressons-nous maintenant à la présence dans la formule de changement de variables du jacobien de φ. Supposons que la fonction f soit constante égale à 1 ; l'intégrale double $\int_V f(x_1,x_2)\ \mathrm{d}x_1\ \mathrm{d}x_2$ qui s'interprète comme le volume du parallélépipède de base $V =]0,\alpha[\times]0,\beta[$ et de hauteur 1 vaut $\alpha\beta$ mais l'intégrale $\int_U f(\varphi(y_1,y_2))\ \mathrm{d}y_1\mathrm{d}y_2$ qui s'interprète comme le volume du parallélépipède de base $U =]0,1[\times]0,1[$ et de hauteur 1 est égale à 1. Par contre, comme $J_\varphi(y_1,y_2) = \alpha\beta$, l'intégrale $\int_U f(\varphi(y_1,y_2))\ |J_\varphi(y_1,y_2)|\ \mathrm{d}y_1\mathrm{d}y_2$ est elle égale à $\alpha\beta$ et on peut considérer que la présence du jacobien $J_\varphi(y_1,y_2)$ dans la formule de changement de variables vise à prendre en compte les variations d'aire du domaine d'intégration lorsqu'il est transformé géométriquement par la fonction de changement de variables.

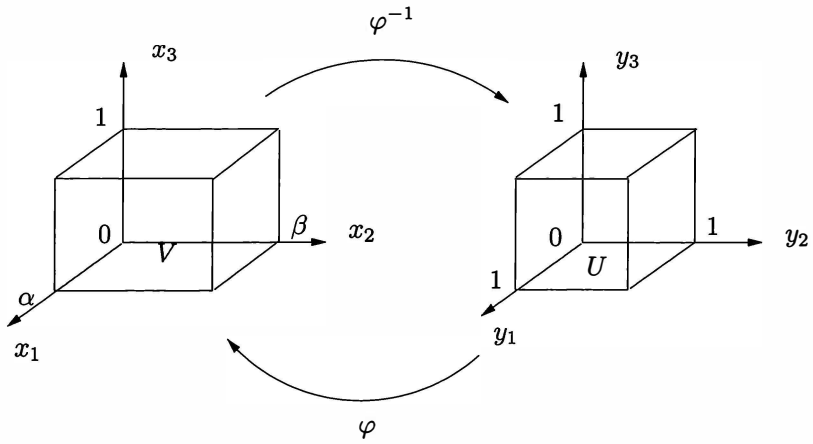


Fig. 20 Illustration de la formule de changement de variables.

Exemples

1. Calculons l'aire A du domaine délimité par l'ellipse d'équation

$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = 1 \qquad (a,b) \in \mathbb{R}_+^* \times \mathbb{R}_+^*$$

en considérant le changement de variables défini par l'application

$$\varphi : (x_1, x_2) \in \mathbb{R}^2 \longmapsto \left(\frac{x_1}{a}, \frac{x_2}{b}\right) \in \mathbb{R}^2.$$

Notons $V = \{(x_1, x_2) \in \mathbb{R}^2 \ ; \ x_1^2/a^2 + x_2^2/b^2 < 1\}$ l'ouvert correspondant à l'intérieur du domaine délimité par l'ellipse et désignons par U le disque ouvert de rayon 1 et de centre $(0,0)$. L'application φ réalise une bijection de l'ouvert U dans V. Par ailleurs, on a vu p. 895 que l'application φ est de classe $\mathcal{C}^1$, de même que sa bijection réciproque, et que son jacobien vaut ab. L'application φ est donc un $\mathcal{C}^1$-difféomorphisme de U dans V et la formule de changement de variables indique que l'aire de l'ellipse vaut :

$$A = \int_V 1 \, \mathrm{d}x_1 \mathrm{d}x_2 = \int_U 1 \ |ab| \ \mathrm{d}y_1 \mathrm{d}y_2 = ab \int_U 1 \, \mathrm{d}y_1 \mathrm{d}y_2.$$

L'intégrale double $\int_U 1 \, \mathrm{d}y_1 \mathrm{d}y_2$ s'interprète comme l'aire du disque de rayon 1 et de centre $(0,0)$ et vaut donc π. On en conclut que l'aire délimitée par l'ellipse vaut πab.

2. Calculons $\int_V f$ lorsque V est le rectangle de $\mathbb{R}^2$ de sommets $A = (-1,0)$, $B = (0,1)$, $C = (2,-1)$ et $D = (1,-2)$ et l'application f est définie par $f : (x_1, x_2) \in V \mapsto \mathrm{e}^{x_1 + x_2}$. L'ensemble V peut être caractérisé par les conditions $-1 < x_1 + x_2 < 1$ et $-1 < x_1 - x_2 < 3$ de sorte que si on décide de prendre comme nouvelles variables $y_1 = x_1 + x_2$ et $y_2 = x_1 - x_2$ alors pour $(x_1, x_2) \in V$ on a $(y_1, y_2) \in U$ où $U =]-1, 1[\times]-1, 3[$. On a l'expression suivante des anciennes variables x_1 et x_2 en fonction des nouvelles variables y_1 et y_2 :

$$x_1 = \tfrac{1}{2}(y_1 + y_2) \quad \text{et} \quad x_2 = \tfrac{1}{2}(y_1 - y_2).$$

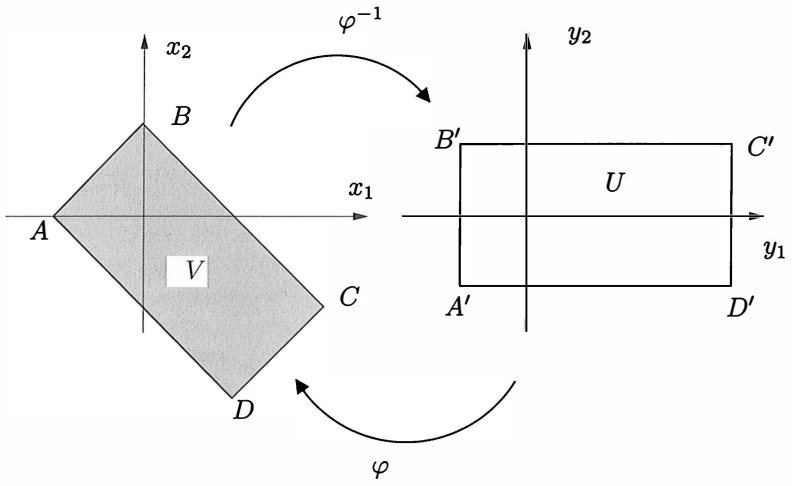


Fig. 21 Application de la fonction de changement de variables sur le rectangle de sommets $A = (-1, 0)$, $B = (0, 1)$, $C = (2, -1)$ et $D = (1, -2)$.

Cela nous amène à introduire l'application de changement de variables

$$\varphi : (y_1, y_2) \in U \longmapsto \left(\tfrac{1}{2}(y_1 + y_2), \tfrac{1}{2}(y_1 - y_2)\right) \in V.$$

L'application φ est de classe $\mathcal{C}^1$ sur U et réalise une bijection de U sur V. Son jacobien en $(y_1, y_2) \in U$ vaut

$$\mathcal{J}_\varphi(y_1, y_2) = \det \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix} = -\frac{1}{2}$$

et ne s'annule pas sur U. D'après le théorème d'inversion locale, on en déduit que φ est un $\mathcal{C}^1$-difféomorphisme de U sur V.

En utilisant la formule de changement de variables, on obtient

$$\begin{aligned} \int_V f(x_1, x_2)\, \mathrm{d}x_1 \mathrm{d}x_2 &= \int_U f(\varphi(y_1, y_2))\, |\mathcal{J}_\varphi(y_1, y_2)|\, \mathrm{d}y_1 \mathrm{d}y_2 = \int_U \tfrac{1}{2}\, \mathrm{e}^{y_1}\, \mathrm{d}y_1 \mathrm{d}y_2 \\ &= \frac{1}{2} \int_{-1}^{1} \left(\int_{-1}^{3} \mathrm{e}^{y_1} \mathrm{d}y_2 \right) \mathrm{d}y_1 = 2 \int_{-1}^{1} \mathrm{e}^{y_1} \mathrm{d}y_1 = 2\,(e - 1/e)\,. \end{aligned}$$

On retrouve le résultat obtenu par une autre méthode à la p. 884. Cet exemple illustre l'utilisation du théorème de changement de variables pour transformer l'ensemble V sur lequel on doit initialement intégrer en un nouvel ensemble sur lequel le calcul devient plus aisé.

3. Déterminons le moment d'inertie par rapport à son axe de symétrie d'un tore[45] T de rayon R et de distance à son centre a ($a > R$), supposé plein et formé d'un matériau homogène de masse M, voir la figure 22, en utilisant le changement de variables

$$\varphi : (r, \theta, t) \longmapsto ((a + r\cos(t))\cos(\theta), (a + r\cos(t))\sin(\theta), R\sin(t)).$$

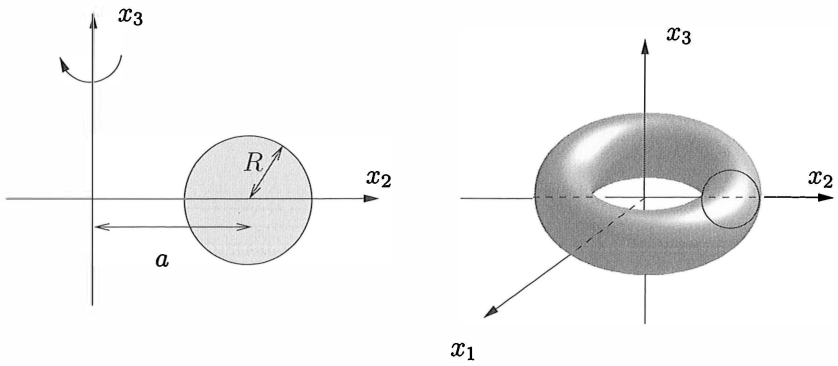


Fig. 22 Un tore engendré par la rotation d'un disque autour de l'axe $(O; \mathbf{e}_3)$.

Commençons par déterminer un ouvert U de $\mathbb{R}^3$ tel que φ réalise une bijection de U dans un ouvert V dont l'adhérence correspond au tore T. On peut découper par tranche le tore selon des plans d'équation $x_3 = c$ avec c variant entre $-R$ et R. Le réel c peut s'écrire sous la forme $x_3 = R\sin(t)$ où t varie entre $-\frac{\pi}{2}$ et $\frac{\pi}{2}$. L'intersection du tore avec un plan $x_3 = R\sin(t)$ fixé est un anneau, voir la figure 23. Notons ρ_1 le rayon intérieur de l'anneau et ρ_2 son rayon extérieur ; cet anneau admet la paramétrisation

$$(u, \theta) \in]\rho_1, \rho_2[\times]0, 2\pi[\longmapsto \big(u\cos(\theta), u\sin(\theta)\big).$$

On détermine les valeurs de ρ_1 et ρ_2 en fonction du plan $x_3 = R\sin(t)$ en ayant recours au théorème de Pythagore, voir la figure 23.

Compte tenu du fait que $t \in]-\frac{\pi}{2}, \frac{\pi}{2}[$, on a $\cos(t) > 0$ et

$$\rho_1 = a - \sqrt{R^2 - x_3^2} = a - \sqrt{R^2 - R^2\sin^2(t)} = a - R\,|\cos(t)| = a - R\cos(t)$$

et $\rho_2 = a + R\cos(t)$. La variable u qui varie entre ρ_1 et ρ_2 peut donc s'écrire sous la forme $u = a + r\cos(t)$ où la variable r varie entre $-R$ et R. On obtient donc pour paramétrage de l'anneau

$$(r, \theta) \in]-R, R[\times]0, 2\pi[\longmapsto \big((a + r\cos(t))\cos(\theta), (a + r\cos(t))\sin(\theta)\big).$$

[45] Un tore est la surface engendrée par la rotation d'un cercle autour d'une droite située dans son plan, mais ne passant pas par son centre.

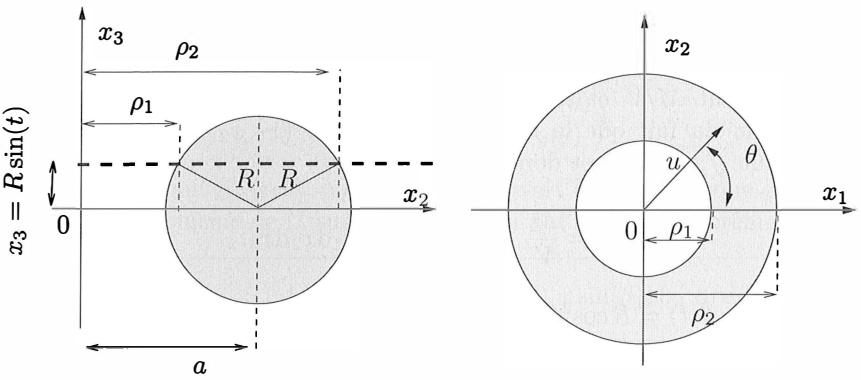


Fig. 23 Vue de coupe du tore dans le plan $x_1 = 0$ (à gauche) et $x_3 = R\sin(t)$ (à droite).

On en conclut que φ réalise une bijection de $U =]-R, R[\times]0, 2\pi[\times]\frac{\pi}{2}, \frac{\pi}{2}[$ dans un ouvert $V = \varphi(U)$ dont l'adhérence est le tore T. Comme les fonctions sinus et cosinus sont de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$, l'application φ est de classe $\mathcal{C}^1$ sur U. Son jacobien est donné par

$$\mathcal{J}_\varphi(r,\theta,t) = \begin{vmatrix} \cos(t)\cos(\theta) & -(a + r\cos(t))\sin(\theta) & -r\sin(t)\cos(\theta) \\ \cos(t)\sin(\theta) & (a + r\cos(t))\cos(\theta) & -r\sin(t)\sin(\theta) \\ 0 & 0 & R\cos(t) \end{vmatrix}.$$

On calcule facilement ce déterminant en le développant selon la dernière ligne ; on obtient $\mathcal{J}_\varphi(r,\theta,t) = R\cos^2(t)\,(a + r\cos(t))$. On notera que le jacobien est strictement positif car $a > R > 0$. Comme le jacobien $\mathcal{J}_\varphi(r,\theta,t)$ ne s'annule pas sur U, on peut conclure à l'aide du théorème d'inversion locale que φ est un $\mathcal{C}^1$-difféomorphisme de U sur V.

En utilisant la formule de changement de variables φ, le volume du tore T s'exprime sous la forme

$$\begin{aligned} V &= \int_T 1\,\mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3 = \int_U |\mathcal{J}_\varphi(r,\theta,t)|\,\mathrm{d}t\mathrm{d}\theta\mathrm{d}r \\ &= \int_{-R}^{R}\int_0^{2\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \left(Ra\cos^2(t) + Rr\cos^3(t)\right)\mathrm{d}t\mathrm{d}\theta\mathrm{d}r \\ &= 2\pi Ra\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\cos^2(t)\left(\int_{-R}^{R} 1\,\mathrm{d}r\right)\mathrm{d}t + 2\pi R\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\cos^2(t)\left(\int_{-R}^{R} r\,\mathrm{d}r\right)\mathrm{d}t. \end{aligned}$$

La seconde intégrale est nulle car $\int_{-R}^{R} r\,\mathrm{d}r = 0$ puisqu'on intègre une fonction impaire sur un intervalle de centre 0. En utilisant la relation $\cos(2t) = 2\cos^2(t) - 1$, on obtient

$$V = 2\pi R^2 a\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}(1 + \cos(2t))\,\mathrm{d}t = 2\pi R^2 a\Big[t + \tfrac{1}{2}\sin(2t)\Big]_{-\frac{\pi}{2}}^{\frac{\pi}{2}} = 2\pi^2 R^2 a.$$

Le volume du tore est donc $2\pi^2 R^2 a$.

Le tore étant supposé rempli d'un matériau homogène de masse M, sa masse volumique vaut M/V et son moment d'inertie par rapport à l'axe $(O;\mathbf{e}_3)$, compte tenu du fait que la distance d'un point (x_1, x_2, x_3) du tore à l'axe $(O;\mathbf{e}_3)$ vaut $\sqrt{x_1^2 + x_2^2}$, est donné par

$$J = \frac{M}{V} \int_T (x_1^2 + x_2^2)\, \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}x_3.$$

Puisque $\mathcal{J}_\varphi(r, \theta, t) = R\cos^2(t)\,(a + r\cos(t))$ et que $x_1^2 + x_2^2 = (a + r\cos(t))^2$, en utilisant le changement de variables φ, on obtient

$$\frac{V}{M}\, J = \int_{-R}^{R} \int_0^{2\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} R(a + r\cos(t))^3 \cos^2(t)\, \mathrm{d}t\mathrm{d}\theta\mathrm{d}r.$$

L'intégrant $g(r, \theta, t) = R(a + r\cos(t))^3 \cos^2(t)$ s'écrit encore, d'après la formule du binôme de Newton, sous la forme

$$g(r, \theta, t) = a^3 R\cos^2(t) + 3aRr^2\cos^4(t) + 3a^2 Rr\cos^3(t) + Rr^3\cos^5(t).$$

Comme $\int_{-R}^{R} r\ \mathrm{d}r = \int_{-R}^{R} r^3\ \mathrm{d}r = 0$, seuls les deux premiers termes vont fournir une contribution non nulle ; on a donc

$$\frac{V}{M}\, J = 4\pi a^3 R^2 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^2(t)\ \mathrm{d}t + 6\pi aR \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^4(t) \left(\int_{-R}^{R} r^2\ \mathrm{d}r \right)\ \mathrm{d}t.$$

On a $\displaystyle\int_{-R}^{R} r^2\ \mathrm{d}r = \frac{2}{3}R^3$, $\displaystyle\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^2(t)\ \mathrm{d}t = \frac{1}{2}\pi$ et $\displaystyle\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^4(t)\ \mathrm{d}t = \frac{3}{8}\pi$, ce qui permet de conclure que

$$J = \frac{M}{V}\left(2\pi^2 a^3 R^2 + \tfrac{3}{2}\pi^2 aR^4\right) = \frac{M}{V} 2\pi^2 aR^2 \left(a^2 + \tfrac{3}{4}R^2\right) = M\left(a^2 + \tfrac{3}{4}R^2\right),$$

compte tenu du fait que $V = 2\pi^2 aR^2$.

Les changements de variables les plus utilisés sont les changements de variables dits polaire, cylindrique et sphérique. Étant donné leur importance pratique, nous allons détailler leur formulation dans les paragraphes suivants.

12.3.2 Changement de variables polaire

Le changement de variables polaire est un changement de variables dans $\mathbb{R}^2$ défini par l'application

$$\begin{aligned} \varphi\ :\ \mathbb{R}^2 &\longrightarrow \quad \mathbb{R}^2 \\ (r, \theta) &\longmapsto (r\cos(\theta), r\sin(\theta)). \end{aligned}$$

L'application φ est de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$ car les fonctions sinus, cosinus et identité sont de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$. Sa matrice jacobienne est

$$\mathrm{J}_\varphi(r,\theta) = \begin{pmatrix} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta) \end{pmatrix}$$

et son jacobien vaut $\mathcal{J}_\varphi(r,\theta) = r$. Pour tout ouvert U de $\mathbb{R}^2$ tel que φ définisse un $\mathcal{C}^1$-difféomorphisme de U sur $V = \varphi(U)$, on a d'après le théorème 12.2 :

$$\boxed{\int_{\varphi(U)} f(x_1,x_2)\,\mathrm{d}x_1\mathrm{d}x_2 = \int_U f(r\cos(\theta), r\sin(\theta))\,|r|\,\mathrm{d}r\mathrm{d}\theta.}$$

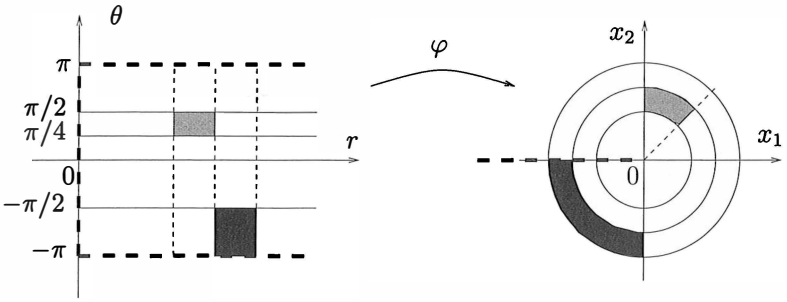


Fig. 24 Illustration du changement de variables polaire.

Intéressons-nous au cas où l'ouvert U est $]0,+\infty[\times]-\pi,\pi[$. L'image de U par φ est l'ouvert $V = \mathbb{R}^2 \setminus X$ où $X = \{(x_1,x_2) \in \mathbb{R}^2 \ ; \ x_1 \leqslant 0 \ \text{ et } \ x_2 = 0\}$, voir la figure 24. Comme le jacobien de φ ne s'annule pas sur U, d'après le théorème d'inversion locale, pour établir que φ est un $\mathcal{C}^1$-difféomorphisme de U sur V, il suffit de montrer que φ est une injection de U dans V. Pour (r_1,θ_1) et (r_2,θ_2) dans U on a

$$\varphi(r_1,\theta_1) = \varphi(r_2,\theta_2) \iff \begin{cases} r_1\cos(\theta_1) &= r_2\cos(\theta_2) \\ r_1\sin(\theta_1) &= r_2\sin(\theta_2) \end{cases}$$

En sommant ces deux équations après les avoir élevées au carré, on obtient $r_1^2 = r_2^2$ ce qui, compte tenu du fait que r_1 et r_2 sont positifs, permet d'établir que $r_1 = r_2$. On en déduit alors que les angles θ_1 et θ_2, qui appartiennent tous les deux à $]-\pi,\pi[$, ont même sinus et cosinus ; ils sont donc égaux. On en conclut que l'application φ est injective et par conséquent qu'il s'agit d'un $\mathcal{C}^1$-difféomorphisme de U sur V.

On pourrait montrer que la bijection réciproque de φ est l'application $\varphi^{-1} : (x_1,x_2) \in V \longmapsto (\varphi_1^{-1}(x_1,x_2), \varphi_2^{-1}(x_1,x_2))$ où $\varphi_1^{-1}(x_1,x_2) = \sqrt{x_1^2+x_2^2}$ et

$$\varphi_2^{-1}(x_1,x_2) = \begin{cases} \arccos(x_1/\sqrt{x_1^2+x_2^2}) & \text{si } x_2 > 0 \\ 0 & \text{si } x_2 = 0 \\ -\arccos(x_1/\sqrt{x_1^2+x_2^2}) & \text{si } x_2 < 0 \end{cases}.$$

Exemples

1. Nous avons calculé p. 883 l'intégrale double $I = \int_D (x_1^2 + x_2^2)\, \mathrm{d}x_1\mathrm{d}x_2$ où D désigne le disque de $\mathbb{R}^2$ centré à l'origine et de rayon 1 en effectuant une intégration par pile. Nous allons voir qu'un changement de variables polaire permet de calculer beaucoup plus rapidement la valeur de cette intégrale. Notons $f : (x_1, x_2) \in D \mapsto (x_1^2 + x_2^2)$. Commençons par remarquer que l'ensemble D est l'adhérence de l'ouvert

$$V = \Big\{(x_1, x_2) \in \mathbb{R}^2 \ ; \ 0 < x_1^2 + x_2^2 < 1\Big\} \setminus \Big\{(a, 0) \in \mathbb{R}^2 \ ; \ a < 0\Big\}$$

et que V est l'image par l'application φ de changement de variables polaire de l'ouvert $U =]0, 1[\times]-\pi, \pi[$. On a donc

$$\int_D f(x_1, x_2)\, \mathrm{d}x_1\mathrm{d}x_2 = \int_V f(x_1, x_2)\, \mathrm{d}x_1\mathrm{d}x_2 = \int_{\varphi(U)} f(x_1, x_2)\, \mathrm{d}x_1\mathrm{d}x_2.$$

Puisque φ est un $\mathcal{C}^1$-difféomorphisme de U sur V, on obtient en utilisant la formule de changement de variables :

$$\begin{aligned} I &= \int_{\varphi(U)} f(x_1, x_2)\, \mathrm{d}x_1\mathrm{d}x_2 = \int_U f(r\cos(\theta), r\sin(\theta))\ r\ \mathrm{d}r\mathrm{d}\theta \\ &= \int_0^{2\pi} \left(\int_0^1 r^3\, \mathrm{d}r \right) \mathrm{d}\theta = \int_0^{2\pi} \frac{1}{4}\, \mathrm{d}\theta = \frac{\pi}{2}. \end{aligned}$$

2. Calculons l'intégrale double $I = \int_D \sqrt{\rho^2 - (x_1^2 + x_2^2)}\, \mathrm{d}x_1\mathrm{d}x_2$ où D désigne le disque de $\mathbb{R}^2$ centré à l'origine et de rayon $R > 0$ et ρ un réel tel que $\rho > R$. Considérons la fonction φ de changement de variables polaire sur l'ensemble ouvert $U = \{(r, \theta) \in \mathbb{R}^2 \ ; \ r \in]0, R[\ \text{ et } \ \theta \in]-\pi, \pi[\}$. L'application φ réalise un $\mathcal{C}^1$-difféomorphisme de U sur $V = \varphi(U)$ où

$$V = \Big\{(x_1, x_2) \in \mathbb{R}^2 \ ; \ 0 < x_1^2 + x_2^2 < R^2\Big\} \setminus \Big\{(a, 0) \in \mathbb{R}^2 \ ; \ a < 0\Big\}$$

et l'adhérence de $\varphi(U)$ est le disque D de centre $(0, 0)$ et de rayon R. On a donc $I = \int_{\varphi(U)} \sqrt{\rho^2 - (x_1^2 + x_2^2)}\, \mathrm{d}x_1\mathrm{d}x_2$ et la formule de changement de variables indique que

$$\begin{aligned} I &= \int_U \sqrt{\rho^2 - \big((r\cos(\theta))^2 + (r\cos(\theta))^2\big)}\ r\ \mathrm{d}r\mathrm{d}\theta \\ &= \int_0^R \left(\int_0^{2\pi} r\sqrt{\rho^2 - r^2}\, \mathrm{d}\theta \right) \mathrm{d}r = 2\pi \int_0^R r\sqrt{\rho^2 - r^2}\, \mathrm{d}r \\ &= 2\pi \Big[-\tfrac{1}{3}\big(\rho^2 - r^2\big)^{\frac{3}{2}} \Big]_0^R = \frac{2}{3}\pi \ \Big(\rho^3 - (\rho^2 - R^2)^{\frac{3}{2}}\Big). \end{aligned}$$

3. Calculons l'aire de la région du plan définie par

$$D = \Big\{(x_1, x_2) \in \mathbb{R}^2 \ ; \ x_1^2 + x_2^2 \leqslant R^2 \ \text{ et } \ (x_1 - a)^2 + x_2^2 \leqslant a^2\Big\}$$

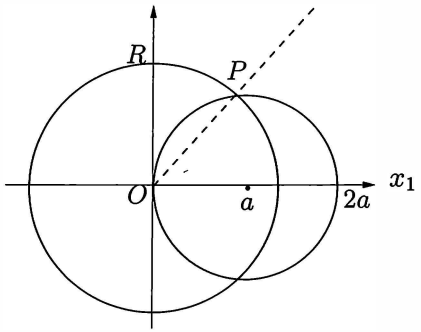


Fig. 25 Représentation de l'ensemble D intersection de deux disques.

où a et R désignent deux réels strictement positifs tels que $R \leqslant 2a$. Graphiquement l'ensemble D correspond à l'intersection des deux disques, D_1 centré à l'origine de rayon R et D_2 centré en $(a, 0)$ et de rayon a, voir la figure 25.

Les coordonnées du point P d'intersection des deux cercles situé dans le demi-plan d'équation $x_2 \geqslant 0$ s'obtiennent en résolvant le système

$$\begin{cases} x_1^2 + x_2^2 = R^2 \\ (x_1 - a)^2 + x_2^2 = a^2 \\ x_2 \geqslant 0 \end{cases}.$$

En développant le carré figurant dans la deuxième équation et en combinant cette expression avec la première équation, on établit que

$$a^2 - 2ax_1 + \underbrace{x_1^2 + x_2^2}_{=R^2} = a^2 \quad \text{d'où} \quad x_1 = \frac{R^2}{2a}.$$

Les relations $x_1^2 + x_2^2 = R^2$ et $x_2 \geqslant 0$ permettent alors d'en déduire que $x_2 = \frac{R}{2a}\sqrt{4a^2 - R^2}$. Les coordonnées du point P sont donc $x_{1,P} = \frac{R^2}{2a}$ et $x_{2,P} = \frac{R}{2a}\sqrt{4a^2 - R^2}$.

Pour calculer l'aire A de l'intersection D des deux disques D_1 et D_2, on peut commencer par remarquer que pour des raisons de symétrie, cette aire est égale à deux fois l'aire A' de la portion du domaine D située dans le demi-plan d'équation $x_2 \geqslant 0$. L'aire A' se décompose en la somme de l'aire des deux secteurs angulaires limités par la droite (OP), voir la figure 25. Pour calculer l'aire de ces deux secteurs angulaires, on peut effectuer un changement de variables polaire. Les coordonnées polaires du point P sont (r_P, θ_P) avec $r_P = R$ car P appartient au cercle centré à l'origine de rayon R et

$$\theta_P = \arctan(x_{2,P}/x_{1,P}) = \arctan\left(\sqrt{4(a/R)^2 - 1}\right).$$

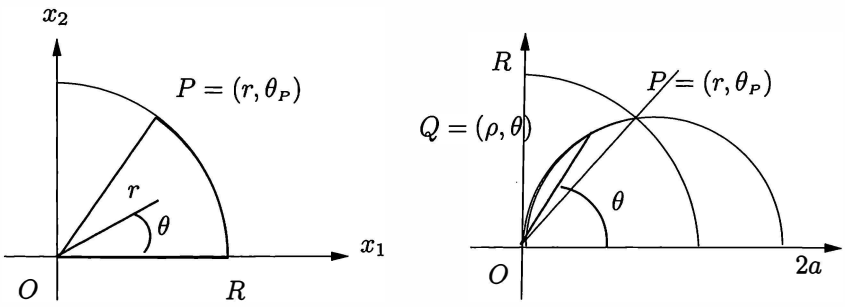


Fig. 26 Paramétrisation en coordonnées polaires des deux secteurs angulaires.

L'aire du premier secteur angulaire S_1 représenté à gauche sur la figure 26 se calcule sans difficulté en utilisant un changement de variables polaire ; on a

$$A_1' = \int_{S_1} 1 \, \mathrm{d}x_1 \, \mathrm{d}x_2 = \int_0^{\theta_P} \left(\int_0^R r \, \mathrm{d}r \right) \mathrm{d}\theta = \int_0^{\theta_P} \frac{1}{2} R^2 \, \mathrm{d}\theta = \frac{1}{2} R^2 \theta_P .$$

Pour pouvoir calculer l'aire du second secteur angulaire S_2 (représenté à droite sur la figure 26) en utilisant un changement de variables polaire, il convient de décrire S_2 à l'aide des coordonnées polaires. On peut le faire de la manière suivante : l'angle polaire θ varie entre θ_P et $\pi/2$; pour une valeur fixée $\theta \in [\theta_P, \pi/2]$ la variable radiale r varie entre 0 et une valeur maximale ρ dépendant de θ que l'on peut expliciter en exploitant le fait que le point Q de coordonnées polaires (ρ, θ) appartient au cercle de centre $(a, 0)$ et de rayon a. On a donc

$$(\rho \cos(\theta) - a)^2 + \rho^2 \sin^2(\theta) = a^2 \quad \text{ce qui donne } \rho = 2a \cos(\theta).$$

L'aire de S_2 s'exprime donc sous la forme

$$\begin{aligned} A_2' &= \int_{S_2} 1 \, \mathrm{d}x_1 \, \mathrm{d}x_2 = \int_{\theta_P}^{\frac{\pi}{2}} \left(\int_0^{\rho(\theta)} r \, \mathrm{d}r \right) \mathrm{d}\theta = \int_{\theta_P}^{\frac{\pi}{2}} \frac{1}{2} \rho^2(\theta) \, \mathrm{d}\theta \\ &= 2a^2 \int_{\theta_P}^{\frac{\pi}{2}} \cos^2(\theta) \, \mathrm{d}\theta . \end{aligned}$$

Pour calculer cette intégrale, on utilise la formule : $2\cos^2(\theta) = 1 + \cos(2\theta)$; on obtient

$$A_2' = a^2 \int_{\theta_P}^{\frac{\pi}{2}} (1 + \cos(2\theta)) \, \mathrm{d}\theta = a^2 \left[\theta + \tfrac{1}{2} \sin(2\theta) \right]_{\theta_P}^{\frac{\pi}{2}} = a^2 \left(\frac{\pi}{2} - \theta_P \right) - \frac{a^2}{2} \sin(2\theta_P).$$

Comme $\theta_P \in]0, \frac{\pi}{2}[$, on a

$$\sin(2\theta_P) = 2\cos(\theta)\sin(\theta) = 2\cos^2(\theta)\tan(\theta_P) = \frac{2\tan(\theta_P)}{1 + \tan^2(\theta_P)}.$$

Par ailleurs, on a

$$\tan(\theta_P) = \frac{x_{2,P}}{x_{1,P}} = \sqrt{4(a/R)^2 - 1} = \frac{1}{R}\sqrt{(2a)^2 - R^2}$$

ce qui permet d'établir que $\sin(2\theta_P) = \frac{1}{2a^2} R\sqrt{(2a)^2 - R^2}$. On en conclut que

$$A_2' = a^2\left(\tfrac{1}{2}\pi - \theta_P\right) - \tfrac{1}{4}R\sqrt{(2a)^2 - R^2}$$

puis que l'aire de l'intersection des deux disques $A = 2A_1' + 2A_2'$ vaut

$$A = \pi a^2 + (R^2 - 2a^2)\arctan\left(\sqrt{4\left(\frac{a}{R}\right)^2 - 1}\right) - \frac{1}{2}R\sqrt{(2a)^2 - R^2}.$$

Exercice 12.6
Calculer l'intégrale double $\int_D \sqrt{x_1^2 + x_2^2}\, \mathrm{d}x_1\mathrm{d}x_2$ lorsque
1 - $D = \{(x_1, x_2) \in \mathbb{R}^2 \ ; \ x_1^2 + x_2^2 \leqslant R^2\}$;
2 - $D = \{(x_1, x_2) \in (\mathbb{R}^+)^2 \ ; \ x_1^2 + x_2^2 \leqslant 1 \text{ et } x_1^2 + x_2^2 \geqslant 2x_2\}$.

12.3.3 Changement de variables cylindrique

Le changement de variables cylindrique est un changement de variables dans $\mathbb{R}^3$ défini par l'application

$$\begin{aligned} \varphi \ : \quad \mathbb{R}^3 \quad &\longrightarrow \quad \mathbb{R}^3 \\ (r, \theta, z) &\longmapsto (r\cos(\theta), r\sin(\theta), z). \end{aligned}$$

Cette application est de classe $\mathcal{C}^1$ sur $\mathbb{R}^3$. Sa matrice jacobienne est

$$\mathrm{J}_\varphi(r, \theta, z) = \begin{pmatrix} \cos(\theta) & -r\sin(\theta) & 0 \\ \sin(\theta) & r\cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

et son jacobien vaut $\mathcal{J}_\varphi(r, \theta, z) = r$. Pour tout ouvert $U \subset \mathbb{R}^* \times \mathbb{R} \times \mathbb{R}$ tel que φ soit un $\mathcal{C}^1$-difféomorphisme de U sur $V = \varphi(U)$, la formule de changement de variables s'écrit :

$$\boxed{\int_{\varphi(U)} f(x_1, x_2, x_3)\ \mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3 = \int_U f(r\cos(\theta), r\sin(\theta), z)\,|r|\ \mathrm{d}r\mathrm{d}\theta\mathrm{d}z.}$$

L'application φ est un $\mathcal{C}^1$-difféomorphisme de l'ouvert $U =]0, +\infty[\times]-\pi, \pi[\times \mathbb{R}$ dans l'ouvert $V = \mathbb{R}^3 \setminus X$ où $X = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \ ; \ x_1 \leqslant 0 \ \text{ et } \ x_2 = 0\}$ et

$$\varphi^{-1} : (x_1, x_2, x_3) \in V \longmapsto (\varphi_1^{-1}(x_1, x_2, x_3), \varphi_2^{-1}(x_1, x_2, x_3), x_3)$$

où $\varphi_1^{-1}(x_1, x_2, x_3) = \sqrt{x_1^2 + x_2^2}$ et

$$\varphi_2^{-1}(x_1, x_2, x_3) = \begin{cases} \arccos(x_1/\sqrt{x_1^2 + x_2^2}) & \text{si } x_2 > 0 \\ 0 & \text{si } x_2 = 0 \\ -\arccos(x_1/\sqrt{x_1^2 + x_2^2}) & \text{si } x_2 < 0 \end{cases}.$$

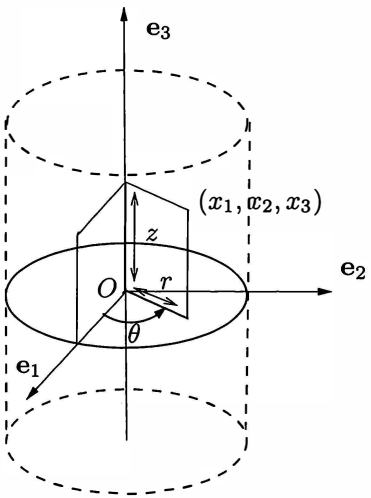


Fig. 27 Changement de variables cylindrique.

Exemples

1. Calculons le moment d'inertie d'un solide homogène de masse M ayant la forme d'un cylindre de hauteur H et de rayon R par rapport à son axe. Le solide étant homogène de volume $V = \pi R^2 H$, sa densité volumique de masse est $\rho = M/V = M/(\pi R^2 H)$. Considérons un repère orthonormé direct $(O; \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ d'origine le centre du disque formant la base du cylindre tel que le vecteur $\mathbf{e}_3$ soit orienté selon l'axe du cylindre. La distance d'un point du cylindre à son axe $(O; \mathbf{e}_3)$ vaut $d = \sqrt{x_1^2 + x_2^2}$. Pour déterminer le moment d'inertie du cylindre, il s'agit de calculer l'intégrale triple $I = \rho \int_D (x_1^2 + x_2^2)\ \mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3$ où

$$D = \left\{(x_1, x_2, x_3) \in \mathbb{R}^3\ ;\ 0 \leqslant x_3 \leqslant H \ \text{ et } \ x_1^2 + x_2^2 \leqslant R^2\right\}.$$

Notons $f : (x_1, x_2, x_3) \in \mathbb{R}^3 \mapsto x_1^2 + x_2^2$. Compte tenu de la géométrie particulière du domaine d'intégration, utilisons la formule de changement de variables cylindrique pour calculer I. L'application φ réalise un $\mathcal{C}^1$-difféomorphisme de l'ouvert $U =]0, R[\times]-\pi, \pi[\times]0, H[$ sur l'ouvert $V = \varphi(U)$ dont l'adhérence est D. La formule de changement de variables indique que

$$\begin{aligned} I &= \rho \int_D f(x_1, x_2, x_3)\ \mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3 \\ &= \rho \int_0^R \int_{-\pi}^{\pi} \int_0^H f(r\cos(\theta), r\sin(\theta), z)\ r\ \mathrm{d}r\mathrm{d}\theta\mathrm{d}z = \rho \int_0^R \int_{-\pi}^{\pi} \int_0^H r^3\ \mathrm{d}r\mathrm{d}\theta\mathrm{d}z. \end{aligned}$$

En utilisant le théorème de Fubini[(46)], on obtient

$$I = \rho \int_0^H \left(\int_{-\pi}^{\pi} \left(\int_0^R r^3\ \mathrm{d}r \right)\ \mathrm{d}\theta \right)\ \mathrm{d}z = \rho \int_0^H \left(\int_{-\pi}^{\pi} \tfrac{1}{4} R^4\ \mathrm{d}\theta \right)\ \mathrm{d}z = \tfrac{1}{2}\pi H R^4 \rho.$$

Le moment d'inertie d'un solide homogène de masse M ayant la forme d'un cylindre de hauteur H et de rayon R par rapport à son axe vaut donc $MR^2/2$.

2. Calculons $I = \displaystyle\int_D x_3 \cos(x_1^2 + x_2^2)\, \mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3$ où

$$D = \left\{(x_1, x_2, x_3) \in \mathbb{R}^3 \ ; \ x_3 \geqslant 0 \ \text{ et } \ x_1^2 + x_2^2 + x_3^2 \leqslant R^2\right\}$$

est la demi-boule de $\mathbb{R}^3$ de centre $(0,0,0)$ et de rayon R située dans le demi-espace $x_3 \geqslant 0$. Pour appliquer la formule de changement de variables cylindrique, il faut déterminer l'ouvert U tel que φ réalise un $\mathcal{C}^1$-difféomorphisme de U sur V avec $\overline{V} = D$. En « découpant » la demi-boule D en tranches dans la direction $\mathbf{e}_3$, on obtient pour intersection de la demi-boule avec les plans d'équation $x_3 = z$ où $z \in [0, R]$, des disques de centre $(0,0,z)$ et dont le rayon, obtenu par le théorème de Pythagore, vaut $r = \sqrt{R^2 - z^2}$, voir la figure 28.

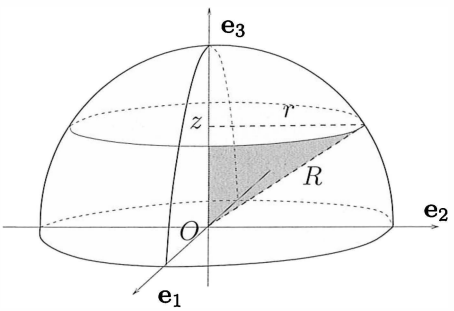


Fig. 28 Demi-boule paramétrée en coordonnées cylindriques.

On en déduit que $D = \overline{\varphi(U)}$ où

$$U = \left\{(r, \theta, z) \in \mathbb{R}^3 \ ; \ -\pi < \theta < \pi \ , \ 0 < z < R \ \text{ et } \ 0 < r < \sqrt{R^2 - z^2}\right\}.$$

La formule du changement de variables cylindrique indique que

$$I = \int_U z\, \cos(r^2)\, r\ \mathrm{d}r\mathrm{d}\theta\mathrm{d}z = \int_0^R \int_0^{\sqrt{R^2-z^2}} zr\cos(r^2) \left(\int_0^{2\pi} \mathrm{d}\theta\right) \mathrm{d}r\ \mathrm{d}z$$
$$= 2\pi \int_0^R z \left(\int_0^{\sqrt{R^2-z^2}} r\cos(r^2)\ \mathrm{d}r\right)\ \mathrm{d}z.$$

Comme pour tout réel β on a $\int_0^\beta r\, \cos(r^2)\, \mathrm{d}r = \left[\frac{1}{2}\sin(r^2)\right]_0^\beta = \frac{1}{2}\sin(\beta^2)$, on en déduit que

$$I = 2\pi \int_0^R \tfrac{1}{2}\, z\, \sin(R^2 - z^2)\ \mathrm{d}z = 2\pi\left[\tfrac{1}{4}\, \cos(R^2 - z^2)\right]_0^R = \frac{\pi}{2}\left(1 - \cos(R^2)\right).$$

(46) On notera que la fonction $(r, \theta, z) \in \mathbb{R}^3 \mapsto r^3$ est continue sur U ce qui permet l'emploi du théorème de Fubini.

12.3.4 Changement de variables sphérique

Le changement de variables sphérique est un changement de variables dans $\mathbb{R}^3$ défini par l'application

$$\begin{aligned} \varphi : \quad \mathbb{R}^3 &\longrightarrow \quad \mathbb{R}^3 \\ (r,\theta,\phi) &\longmapsto (r\cos(\theta)\sin(\phi), r\sin(\theta)\sin(\phi), r\cos(\phi)). \end{aligned}$$

Cette application est de classe $\mathcal{C}^1$ sur $\mathbb{R}^3$. Sa matrice jacobienne est

$$\mathrm{J}_\varphi(r,\theta,\phi) = \begin{pmatrix} \cos(\theta)\sin(\phi) & -r\sin(\theta)\sin(\phi) & r\cos(\theta)\cos(\phi) \\ \sin(\theta)\cos(\phi) & r\cos(\theta)\sin(\phi) & r\sin(\theta)\cos(\phi) \\ \cos(\phi) & 0 & -r\sin(\phi) \end{pmatrix}$$

et son jacobien vaut $\mathcal{J}_\varphi(r,\theta,\phi) = r^2\sin(\phi)$. Pour tout ouvert $U \subset \mathbb{R}^3$ tel que φ soit un $\mathcal{C}^1$-difféomorphisme, la formule de changement de variables s'écrit :

$$\boxed{\int_{\varphi(U)} f(x_1,x_2,x_3)\,\mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3 = \int_U f(\varphi(r,\theta,\phi))\; r^2|\sin(\phi)|\;\mathrm{d}r\mathrm{d}\theta\mathrm{d}\phi.}$$

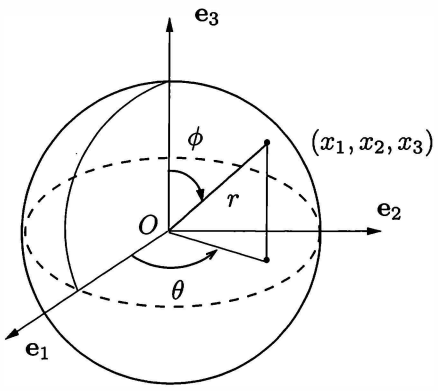


Fig. 29 Changement de variables sphérique.

Par exemple, φ est un $\mathcal{C}^1$-difféomorphisme de l'ouvert $U =]0,+\infty[\times]0,2\pi[\times]0,\pi[$ dans l'ouvert $V = \mathbb{R}^3 \setminus X$ où $X = \{(x_1,x_2,x_3) \in \mathbb{R}^3 \ ; \ x_1 \geqslant 0 \ \text{ et } \ x_2 = 0\}$ et

$$\varphi^{-1} : (x_1,x_2,x_3) \in V \longmapsto (\varphi_1^{-1}(x_1,x_2,x_3), \varphi_2^{-1}(x_1,x_2,x_3), \varphi_3^{-1}(x_1,x_2,x_3))$$

où

$$\varphi_1^{-1}(x_1,x_2,x_3) = \sqrt{x_1^2+x_2^2+x_3^2},$$

$$\varphi_3^{-1}(x_1,x_2,x_3) = \arccos\left(x_3/\sqrt{x_1^2+x_2^2+x_3^2}\right),$$

$$\varphi_2^{-1}(x_1,x_2,x_3) = \begin{cases} \arcsin\left(x_2/\sqrt{x_1^2+x_2^2}\right) & \text{si } x_1 \geqslant 0 \\ \pi - \arcsin\left(x_2/\sqrt{x_1^2+x_2^2}\right) & \text{si } x_1 < 0 \end{cases}.$$

Exemples

1. Dans l'espace, considérons deux points A et B situés sur un même méridien[47] de la sphère de rayon R dont les coordonnées sphériques sont respectivement $(R, 0, \phi_A)$ et $(R, 0, \phi_B)$. Déterminons le volume du secteur sphérique engendré par la rotation du secteur circulaire OAB autour de l'axe des pôles $(O; \mathbf{e}_3)$.

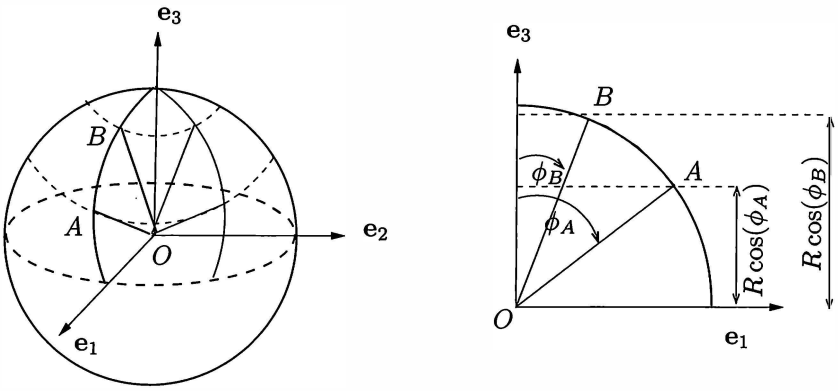


Fig. 30 Secteur sphérique engendré par la rotation du secteur circulaire OAB.

Le secteur sphérique D est difficile à décrire en coordonnées cartésiennes ; c'est l'adhérence de l'image par la fonction φ de changement de variables en coordonnées sphériques de l'ouvert $U =]0, R[\times]0, 2\pi[\times]\phi_B, \phi_A[$. Comme le jacobien de φ ne s'annule pas sur U, le théorème d'inversion locale indique que φ réalise un $\mathcal{C}^1$-difféomorphisme de U sur $\varphi(U)$. En appliquant le théorème de changement de variables, le volume de D s'exprime sous la forme

$$\begin{aligned} V &= \int_D 1 \, \mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3 = \int_U r^2 \sin(\phi) \, \mathrm{d}r\mathrm{d}\theta\mathrm{d}\phi \\ &= 2\pi \int_{\phi_B}^{\phi_A} \sin(\phi) \left(\int_0^R r^2 \, \mathrm{d}r \right) \mathrm{d}\phi = \frac{2}{3}\pi R^3(\cos(\phi_B) - \cos(\phi_A)). \end{aligned}$$

En notant $(x_{1,A}, x_{2,A}, x_{3,A})$ et $(x_{1,B}, x_{2,B}, x_{3,B})$ les coordonnées cartésiennes des points A et B, le volume du secteur sphérique vaut $V = \frac{2}{3}\pi R^2(x_{3,B} - x_{3,A})$.

2. Considérons dans l'espace une boule B_R de centre $(0,0,0)$ et de rayon $R \in \mathbb{R}_+^*$ supposée formée d'un matériau homogène de densité volumique de masse égale à 1 (sa masse M vaut donc $M = \frac{4}{3}\pi R^3$). Cette boule est responsable d'une attraction gravitationnelle en tout point de l'espace dont le potentiel newtonien créé au point $(0, 0, \gamma) \in \mathbb{R}^3$ est donné, lorsque le point $(0, 0, \gamma)$ est en dehors de la boule, par :

[47] Un méridien est un cercle tracé sur la sphère dont le centre coïncide avec le centre de la sphère et qui passe par les deux pôles.

$$\Sigma(\gamma) = \int_{B_R} \frac{1}{\sqrt{x_1^2 + x_2^2 + (x_3 - \gamma)^2}} \, \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}x_3.$$

Utilisons le changement de variables sphérique pour calculer $\Sigma(\gamma)$. La boule B_R est l'adhérence de l'image par la fonction φ de changement de variables en coordonnées sphériques de l'ouvert $U =]0, R[\times]0, 2\pi[\times]0, \pi[$. Comme le jacobien de φ ne s'annule pas sur U, le théorème d'inversion locale indique que φ réalise un $\mathcal{C}^1$-difféomorphisme de U sur $\varphi(U)$ et on obtient en appliquant le théorème de changement de variables :

$$\Sigma(\gamma) = \int_0^R \int_0^{2\pi} \int_0^{\pi} \frac{r^2 \sin(\phi)}{\sqrt{r^2 + \gamma^2 - 2r\gamma\cos(\phi)}} \, \mathrm{d}\phi \mathrm{d}\theta \mathrm{d}r.$$

Dans cette dernière intégrale, la quantité à intégrer ne dépend pas de θ ; on a donc

$$\Sigma(\gamma) = 2\pi \int_0^R \int_0^{\pi} \frac{r^2 \sin(\phi)}{\sqrt{r^2 + \gamma^2 - 2r\gamma\cos(\phi)}} \, \mathrm{d}\phi \mathrm{d}r = 2\pi \int_0^R I(r) \, \mathrm{d}r$$

où pour $r \in [0, R]$ on a posé $I(r) = \displaystyle\int_0^{\pi} \frac{r^2 \sin(\phi)}{\sqrt{r^2 + \gamma^2 - 2r\gamma\cos(\phi)}} \, \mathrm{d}\phi$.

Pour calculer l'intégrale $I(r)$, on remarque que l'intégrant est de la forme

$$C\frac{u'}{2\sqrt{u}} \quad \text{où } u(\phi) = r^2 + \gamma^2 - 2r\gamma\cos(\phi) \quad \text{et } C = r/\gamma.$$

Par conséquent, $I(r) = C[\sqrt{u(\phi)}]_0^{\pi} = \frac{r}{\gamma}\left(|r + \gamma| - |r - \gamma|\right)$. Puisque le point $(0, 0, \gamma)$ est en dehors de la boule B_R, on a $|\gamma| > R$ et on obtient

$$\int_0^{\pi} \frac{r^2 \sin(\phi)}{\sqrt{r^2 + \gamma^2 - 2r\gamma\cos(\phi)}} \, \mathrm{d}\phi = \frac{2r^2}{|\gamma|}.$$

On en conclut que

$$\Sigma(\gamma) = 2\pi \int_0^R \frac{2r^2}{|\gamma|} \, \mathrm{d}r = \frac{4\pi R^3}{3|\gamma|} = \frac{M}{|\gamma|}.$$

Cette expression indique que l'attraction créée par une boule en un point situé à l'extérieur de celle-ci ne dépend pas du rayon de la boule mais uniquement de sa masse et de la distance entre le point et le centre de la boule.

Exercice 12.7

1 - Soient a, b et c trois réels strictement positifs ; en considérant le changement de variables défini par $\varphi : (y_1, y_2, y_3) \mapsto (ay_1, by_2, cy_3)$, calculer le volume du domaine délimité par l'ellipsoïde d'équation $x_1^2/a^2 + x_2^2/b^2 + x_3^2/c^2 = 1$.

2 - On suppose que l'ellipsoïde est rempli d'un matériau dont la masse volumique au point (x_1, x_2, x_3) vaut $\rho(x_1, x_2, x_3) = x_1^2/a^2 + x_2^2/b^2 + x_3^2/c^2$. Déterminer la masse de l'ellipsoïde.

12.3.5 Exploitation de symétries

Il arrive assez fréquemment que les intégrales multiples issues de la modélisation d'une situation physique donnée présentent des « symétries », soit au niveau de la fonction intégrée, soit au niveau de l'ensemble sur lequel on intègre. Nous allons voir qu'il est possible le cas échéant d'exploiter ces symétries pour simplifier le calcul de l'intégrale multiple. Rappelons que dans le cas des fonctions de $\mathbb{R}$ dans $\mathbb{R}$, les symétries les plus fréquentes correspondent au cas où l'on intègre une fonction paire ou impaire sur un intervalle centré en 0 et que l'on a le résultat suivant[48] :

- si f est une application continue sur $[-a, a]$ paire alors $\int_{-a}^{a} f = 2\int_{0}^{a} f$;
- si f est une application continue sur $[-a, a]$ impaire alors $\int_{-a}^{a} f = 0$.

L'exploitation des symétries dans le cas de fonctions de $\mathbb{R}^n$ dans $\mathbb{R}$ nécessite l'introduction des définitions suivantes :

- on appelle *symétrie vectorielle* de $\mathbb{R}^n$ une application s_v de $\mathbb{R}^n$ dans $\mathbb{R}^n$ qui est linéaire et qui vérifie[49] $s_v \circ s_v = \mathrm{id}$;
- on appelle *symétrie affine* de $\mathbb{R}^n$ une application s de $\mathbb{R}^n$ dans $\mathbb{R}^n$ de la forme $s = s_v + c$ où s_v est une symétrie vectorielle de $\mathbb{R}^n$ et c une application constante de $\mathbb{R}^n$ dans $\mathbb{R}^n$.

Exemples

1. Dans $\mathbb{R}$, les seules symétries vectorielles sont l'application identité id et son opposée $-\mathrm{id}$.

2. Dans $\mathbb{R}^2$ la symétrie s_1 par rapport à la droite D d'équation $x_1 = 1$ (voir la figure 31 à gauche) est une symétrie affine ; elle est définie pour tout $(x_1, x_2) \in \mathbb{R}^2$ par $s_1(x_1, x_2) = (2 - x_1, x_2)$.

3. De même, la symétrie s_2 par rapport au point $(1, 1)$ (voir la figure 31 à droite) est une symétrie affine ; elle est définie pour tout $(x_1, x_2) \in \mathbb{R}^2$ par $s_2(x_1, x_2) = (2 - x_1, 2 - x_2)$.

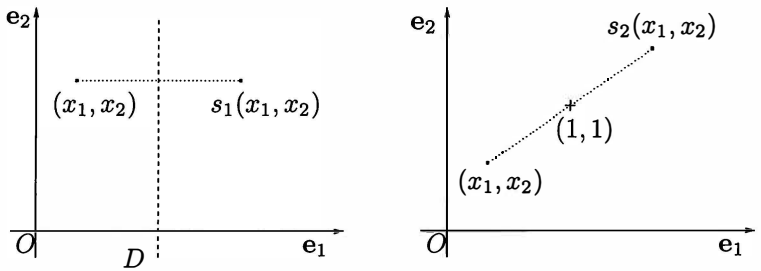


Fig. 31 Symétrie par rapport à une droite et symétrie par rapport à un point.

[48] Voir le *Cours de première année* p. 867.

[49] Rappelons que id désigne l'application identité ; ici id : $x \in \mathbb{R}^n \mapsto x \in \mathbb{R}^n$.

Une symétrie affine s se décomposant sous la forme $s = s_v + c$ où s_v est une symétrie vectorielle de $\mathbb{R}^n$ et c une application constante de $\mathbb{R}^n$ dans $\mathbb{R}^n$, les applications s et s_v ont mêmes dérivées partielles premières. Comme s_v est une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^n$, la matrice jacobienne de l'application s_v correspond à la matrice M représentant s_v dans la base canonique de $\mathbb{R}^n$. Par définition d'une symétrie vectorielle, on a $s_v \circ s_v = \mathrm{id}$, ce qui matriciellement se traduit par $\mathrm{M}^2 = \mathrm{I}_n$. On déduit des propriétés du déterminant[50] que

$$(\det(\mathrm{M}))^2 = \det(\mathrm{M}^2) = \det(\mathrm{I}_n) = 1.$$

Par conséquent, le jacobien de s_v est égal à 1 où -1 et comme s et s_v ont mêmes dérivées partielles premières, le jacobien de s qui est égal à celui de s_v est lui aussi égal à 1 où -1. On en conclut que le jacobien d'une symétrie affine ou d'une symétrie vectorielle de $\mathbb{R}^n$ vaut 1 ou -1.

Considérons un sous-ensemble X de $\mathbb{R}^n$ et une application f continue de X dans $\mathbb{R}$ et supposons qu'il existe une symétrie affine s de $\mathbb{R}^n$ telle que les conditions suivantes soient satisfaites :

- $s(X) = X$ et $f \circ s = f$;
- il existe un sous-ensemble X_1 de X tel que la famille $\{X_1, s(X_1)\}$ forme une partition[51] de X.

D'après la relation de Chasles[52], on a

$$\int_X f(x)\,\mathrm{d}x = \int_{X_1} f(x)\,\mathrm{d}x + \int_{s(X_1)} f(x)\,\mathrm{d}x. \tag{2}$$

Considérons le changement de variables $y = s(x)$; comme par hypothèse on a $f \circ s = f$ et que l'on a vu que le jacobien de s vaut 1 ou -1, on obtient[53]

$$\int_{s(X_1)} f(x)\,\mathrm{d}x = \int_{X_1} f(s(y))\,|\mathcal{J}_s(y)|\,\mathrm{d}y = \int_{X_1} f(y)\,\mathrm{d}y.$$

On déduit alors de la relation (2) que

$$\int_X f(x)\,\mathrm{d}x = 2\int_{X_1} f(x)\,\mathrm{d}x.$$

Sur le même principe, on établit que si f est une application de X dans $\mathbb{R}$ continue et qu'il existe une symétrie affine s de $\mathbb{R}^n$ telle que

- $s(X) = X$ et $f \circ s = -f$;
- il existe un sous-ensemble X_1 de X tel que la famille $\{X_1, s(X_1)\}$ forme une partition de X,

---

[50] Voir la proposition 11.3 p. 478 du *Cours de première année*.

[51] La définition d'une partition est donnée p. 873.

[52] Voir la proposition 12.7 p. 873.

[53] Voir le théorème 12.2 p. 894.

alors on a $\int_X f(x)\ \mathrm{d}x = 0$.

On a en définitive démontré la proposition suivante.

PROPOSITION 12.10 *Soient X un sous-ensemble de $\mathbb{R}^n$ et f une application continue de X dans $\mathbb{R}$. On suppose qu'il existe une symétrie affine s de $\mathbb{R}^n$ telle que $s(X) = X$ et qu'il existe un sous-ensemble X_1 de X tel que la famille $\{X_1, s(X_1)\}$ forme une partition de X.*

- *Si $f \circ s = f$ alors $\int_X f = 2\int_{X_1} f$.*
- *Si $f \circ s = -f$ alors $\int_X f = 0$.*

Remarque Pour $n = 1$, on retrouve les résultats rappelés en préambule concernant l'intégration des fonctions paires ou impaires sur un intervalle centré en 0 en considérant $X = [-a, a]$, $s : x \in \mathbb{R} \mapsto -x$ et $X_1 = [0, a]$ dans la proposition 12.10. ◇

Exemples

1. Calculons l'intégrale $I = \int_X f(x_1, x_2)\ \mathrm{d}x_1\mathrm{d}x_2$ où $f : (x_1, x_2) \in X \mapsto |x_1| + |x_2|$ et $X = [-1, 1]^2$. On peut dans un premier temps considérer la symétrie vectorielle d'axe $(O; \mathbf{e}_2)$ définie par $s_1 : (x_1, x_2) \in \mathbb{R}^2 \mapsto (-x_1, x_2)$. On vérifie que $s_1(X) = X$, que $f \circ s_1 = f$ et que si $X_1 = [0, 1] \times [-1, 1]$ alors $\{X_1, s_1(X_1)\}$ est une partition de X. On déduit de la proposition 12.10 que

$$I = \int_X (|x_1| + |x_2|)\ \mathrm{d}x_1\mathrm{d}x_2 = 2\int_{X_1} (|x_1| + |x_2|)\ \mathrm{d}x_1\mathrm{d}x_2.$$

Considérons à présent la symétrie vectorielle d'axe $(O; \mathbf{e}_1)$ définie par $s_2 : (x_1, x_2) \in \mathbb{R}^2 \mapsto (x_1, -x_2)$. On a $s_2(X_1) = X_1$, $f \circ s_2 = f$ et si $X_2 = [0, 1] \times [0, 1]$ alors $\{X_2, s_2(X_2)\}$ est une partition de X_1. On déduit de la proposition 12.10 que

$$\int_{X_1} (|x_1| + |x_2|)\ \mathrm{d}x_1\mathrm{d}x_2 = 2\int_{X_2} (|x_1| + |x_2|)\ \mathrm{d}x_1\mathrm{d}x_2$$

et donc que $I = 4\int_{X_2} (|x_1| + |x_2|)\ \mathrm{d}x_1\mathrm{d}x_2$.
Pour $(x_1, x_2) \in X_2$ on a $|x_1| + |x_2| = x_1 + x_2$ de sorte que

$$I = 4\int_0^1 \left(\int_0^1 (x_1 + x_2)\ \mathrm{d}x_1\right)\ \mathrm{d}x_2 = 4\int_0^1 \left(\tfrac{1}{2} + x_2\right)\ \mathrm{d}x_2 = 4.$$

2. Calculons le volume V du domaine D formé par l'intersection entre deux cylindres de même rayon R et de même hauteur H (où $H > 2R$) se coupant à mi-hauteur perpendiculairement à leur axe, voir la figure 32. Supposons que dans un repère orthonormé direct $(O; \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ les deux cylindres soient définis par

$$C_1 = \Big\{(x_1, x_2, x_3) \in \mathbb{R}^3 \ ; \ x_1^2 + x_2^2 \leqslant R^2 \ \text{ et } \ x_3 \in [-H, H]\Big\},$$
$$C_2 = \Big\{(x_1, x_2, x_3) \in \mathbb{R}^3 \ ; \ x_1^2 + x_3^2 \leqslant R^2 \ \text{ et } \ x_2 \in [-H, H]\Big\}.$$

Le volume de D est obtenu en calculant la valeur de l'intégrale triple :

$$\int_D 1 \, \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}x_3.$$

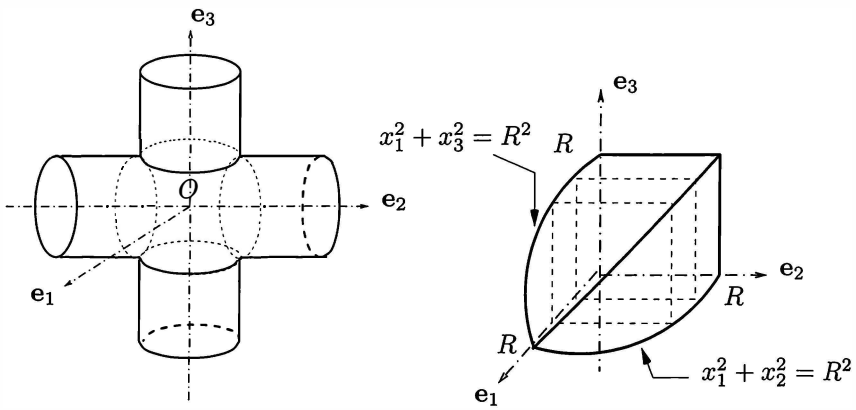


Fig. 32 Situation considérée : deux cylindres de même rayon et même hauteur s'intersectant perpendiculairement à mi-hauteur (à gauche) et représentation du domaine correspondant à l'intersection (à droite).

Compte tenu des symétries que possèdent les deux cylindres C_1 et C_2 par rapport aux plans d'équation $x_1 = 0$, $x_2 = 0$ et $x_3 = 0$, le volume de D est égal à 8 fois le volume de la partie D^+ de D situé dans la partie de l'espace $\mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R}^+$, i.e.

$$\begin{aligned} D^+ &= \Big\{(x_1, x_2, x_3) \in (\mathbb{R}^+)^3 \ ; \ (x_1, x_2, x_3) \in C_1 \cap C_2\Big\} \\ &= \Big\{(x_1, x_2, x_3) \in (\mathbb{R}^+)^3 \ ; \ x_1^2 + x_2^2 \leqslant R^2 \ \text{ et } \ x_1^2 + x_3^2 \leqslant R^2\Big\}. \end{aligned}$$

Afin de visualiser l'ensemble D^+, montrons que son intersection avec un plan d'équation $x_1 = c$ où $c \in]0, R[$ est un carré. Les points à l'intersection de D^+ et du plan $x_1 = c$ sont les points dont les coordonnées (x_1, x_2, x_3) vérifient

$$x_2 \geqslant 0, \quad x_3 \geqslant 0, \quad x_1 = c, \quad x_1^2 + x_2^2 \leqslant R^2 \ \text{ et } \ x_1^2 + x_3^2 \leqslant R^2.$$

On a donc $x_2^2 \leqslant R^2 - c^2$ et $x_3^2 \leqslant R^2 - c^2$. Ces points appartiennent au carré $A_c = [0, \sqrt{R^2 - c^2}] \times [0, \sqrt{R^2 - c^2}]$ dans le plan d'équation $x_1 = c$. L'ensemble D^+ est représenté figure 32.

Pour intégrer sur l'ensemble D^+, on peut intégrer par pile dans la direction $\mathbf{e}_1$; on a :

$$V = 8 \int_{D^+} 1 \, \mathrm{d}x_1 \, \mathrm{d}x_2 \, \mathrm{d}x_3 = 8 \int_0^R \left(\int_{A_c} 1 \, \mathrm{d}x_2 \, \mathrm{d}x_3 \right) \mathrm{d}x_1.$$

L'intégrale $\int_{A_c} 1 \, \mathrm{d}x_2 \, \mathrm{d}x_3$ s'interprète comme l'aire du pavé $A_c = [0, \sqrt{R^2 - c^2}]^2$ et vaut donc $R^2 - c^2$. On a par conséquent :

$$V = 8 \int_0^R (R^2 - x_1^2) \, \mathrm{d}x_1 = 8 \Big[R^2 x_1 - \tfrac{1}{3} x_1^3 \Big]_0^R = \frac{16}{3} R^3.$$

Il est indispensable qu'à la fois la fonction à intégrer et l'ensemble sur lequel on intègre possèdent des propriétés de symétrie « compatibles » pour pouvoir les exploiter au niveau du calcul. Par exemple, la fonction $(x_1, x_2) \mapsto \mathrm{sh}(x_1)\,\mathrm{ch}(x_2)$ et l'ensemble $X = [1, 2]^2$ possèdent des propriétés de symétrie, mais celles-ci ne sont pas compatibles entre elles et ne peuvent être exploitées lors du calcul de $\int_X \mathrm{sh}(x_1)\,\mathrm{ch}(x_2) \, \mathrm{d}x_1 \mathrm{d}x_2$.

12.4 Intégrale multiple généralisée

Jusqu'à présent nous n'avons défini l'intégrale multiple que de fonctions f bornées sur un ensemble borné X de $\mathbb{R}^n$. Comme cela a été fait pour l'intégrale simple[(54)], on peut chercher à étendre la notion d'intégrale multiple :

- à des fonctions bornées sur un ensemble X non borné (par exemple lorsque $X = \mathbb{R}^n$) ;
- à des fonctions f non bornées sur un ensemble borné (par exemple lorsque $f : (x_1, x_2, x_3) \in \mathbb{R}^3_* \mapsto (x_1^2 + x_2^2 + x_3^2)^{-1/2}$ sur la boule $B(0, 1)$).

Dans le cas de l'intégrale simple, cette extension de la notion d'intégrale avait été réalisée par « passage à la limite » de l'intégrale considérée sur intervalle $[a, x]$ fermé et borné à un intervalle non fermé $[a, b[$ (cas de l'intégration d'une fonction non bornée en b) ou à un intervalle non borné $[a, +\infty[$ (cas de l'intégration sur un intervalle non borné). Une telle approche pour l'intégrale multiple est envisageable mais se révèle beaucoup plus compliquée ; en particulier la notion de limite d'une suite d'ensembles dans $\mathbb{R}^n$ a besoin d'être précisée.

(54) Voir le chap. 19 du *Cours de première année*.

Nous nous contenterons dans cette partie de donner, sans justification, quelques résultats permettant d'appréhender simplement le concept d'intégrale multiple généralisée[55] pour les besoins pratiques pouvant apparaître dans les applications. Signalons qu'un cadre théorique beaucoup mieux adapté pour traiter cette extension de la notion d'intégrale multiple est fourni par la théorie de l'intégrale de Lebesgue[56].

Afin de pouvoir considérer la notion de limite d'une suite d'ensembles de $\mathbb{R}^n$, nous avons besoin d'introduire la définition suivante.

DÉFINITION 12.9 *Étant donné un ensemble X de $\mathbb{R}^n$, on appelle* suite exhaustive de compacts de X *toute suite d'ensembles $(K_\ell)_{\ell\in\mathbb{N}}$ vérifiant les conditions suivantes :*

- *pour tout $\ell \in \mathbb{N}$, l'ensemble K_ℓ est compact*[57] *;*
- *pour tout $\ell \in \mathbb{N}$, l'ensemble K_ℓ en inclus dans X ;*
- *pour tout $\ell \in \mathbb{N}$, $K_\ell \subset K_{\ell+1}$ (autrement dit la suite est croissante) ;*
- *pour tout sous-ensemble compact K de X, il existe $\ell \in \mathbb{N}$ tel que $K \subset K_\ell$.*

Exemples

1. La suite $(K_\ell)_\ell$, où $K_\ell = \{(x_1, x_2) \in \mathbb{R}^2 \ ; \ x_1^2 + x_2^2 \leqslant \ell^2\}$ correspond au disque fermé de centre l'origine et de rayon ℓ dans le plan, constitue une suite exhaustive de compacts de $\mathbb{R}^2$, voir la figure 33 (à gauche). En effet, pour tout $\ell \in \mathbb{N}$, l'ensemble K_ℓ qui correspond à la boule fermée de centre 0 et de rayon ℓ pour la norme euclidienne, est un ensemble fermé et borné donc compact. Il est clair que pour tout $\ell \in \mathbb{N}$ on a $K_\ell \subset K_{\ell+1}$. Enfin, si K est un compact de $\mathbb{R}^2$, K est borné et d'après la définition 7.11 p. 382 il existe un réel strictement positif r tel que $K \subset \overline{B}(0,r)$; on en déduit que $K \subset \overline{B}(0,\ell) = K_\ell$ pour $\ell = \mathrm{E}(r) + 1$. Plus généralement, la suite $(K_\ell)_\ell$ où $K_\ell = \{\mathbf{x} \in \mathbb{R}^n \ ; \ \|\mathbf{x}\|_2 \leqslant \ell\}$ constitue une suite exhaustive de compacts de $\mathbb{R}^n$.

2. La suite $(K_\ell)_\ell$ où $K_\ell = \{(x_1, x_2) \in \mathbb{R}^2 \ ; \ 1/\ell^2 \leqslant x_1^2 + x_2^2 \leqslant R^2\}$ est la couronne centrée à l'origine de rayons $1/\ell$ et R (avec $R > 1$) est une suite exhaustive de compacts de $X = \{(x_1, x_2) \in \mathbb{R}^2 \ ; \ x_1^2 + x_2^2 \leqslant R^2\} = \overline{B}(0,R)$, voir la figure 33 (à droite).

---

[55] On parle aussi d'intégrale multiple impropre.

[56] Henri LEBESGUE est un mathématicien français né à Beauvais en 1875 et mort à Paris en 1941. Après avoir été reçu à l'agrégation de mathématiques en 1897, il enseigna quelques années en classes préparatoires à Nancy, en préparant simultanément une thèse qu'il soutint en 1902. Dans cette thèse intitulée *Intégrale, longueur, aire*, H. Lebesgue présente la théorie d'une nouvelle intégrale qui va considérablement simplifier l'utilisation pratique des outils de calcul intégral ; elle requiert toutefois des préliminaires théoriques importants.

[57] Autrement dit, d'après le théorème 7.7 p. 415 il s'agit d'un ensemble fermé et borné.

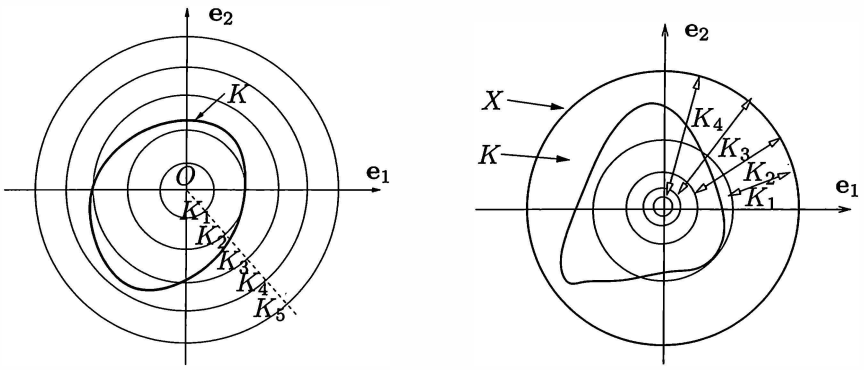


Fig. 33 Suite exhaustive de compacts de $X = \mathbb{R}^2$ (à gauche) et suite exhaustive de compacts de $X = \overline{B}(0, R)$ (à droite).

3. Soit U un ouvert borné de $\mathbb{R}^2$; on désigne par $d(\mathbf{x}, \complement U)$ la distance euclidienne d'un élément $\mathbf{x}$ de U au complémentaire[58] de U dans $\mathbb{R}^2$. La suite $(K_\ell)_\ell$ où

$$K_\ell = \left\{ \mathbf{x} = (x_1, x_2) \in \mathbb{R}^2 \ ; \ x_1^2 + x_2^2 \leqslant \ell^2 \ \text{ et } \ d(\mathbf{x}, \mathbb{R}^n \setminus U) \geqslant 1/\ell \right\}$$

est une suite exhaustive de compacts de U.

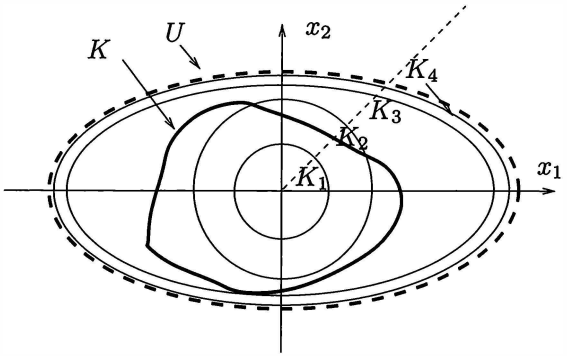


Fig. 34 Suite exhaustive de compacts de l'ouvert borné U représenté en pointillés.

Plus généralement, si U désigne un ouvert borné de $\mathbb{R}^n$, la suite $(K_\ell)_\ell$ où

$$K_\ell = \left\{ \mathbf{x} \in \mathbb{R}^n \ ; \ \|\mathbf{x}\|_2 \leqslant \ell \ \text{ et } \ d(\mathbf{x}, \complement U) \geqslant 1/\ell \right\}$$

est une suite exhaustive de compacts de U.

[58] La définition de la distance d'un point à un ensemble est donnée p. 386.

Remarque Si $(K_\ell)_{\ell\in\mathbb{N}}$ est une suite exhaustive de compacts d'un ensemble X de $\mathbb{R}^n$ alors $X = \bigcup_{\ell\in\mathbb{N}} K_\ell$. ◇

Exercice 12.8 Montrer que la suite $(K_\ell)_\ell$, dont le terme général est $K_\ell = [-\ell, \ell] \times [-\ell, \ell]$, constitue une suite exhaustive de compacts de $\mathbb{R}^2$.

Disposant de la notion de suite exhaustive de compacts, on peut définir la notion d'*intégrale multiple généralisée* de la manière suivante[59].

DÉFINITION 12.10 (Intégrale multiple généralisée convergente)
Soient X un ensemble de $\mathbb{R}^n$ et f une application de X dans $\mathbb{R}$. On dit que f admet une intégrale généralisée convergente sur X s'il existe une suite exhaustive $(K_\ell)_\ell$ de compacts de X telle que :

- *pour tout $\ell \in \mathbb{N}$, f est Riemann-intégrable sur le compact K_ℓ ;*
- *la suite numérique $(J_\ell)_\ell$ de terme général $J_\ell = \int_{K_\ell} |f(x)|\ \mathrm{d}x$ est une suite convergente.*

On appelle alors intégrale de f sur X le réel

$$\int_X f(x)\ \mathrm{d}x = \lim_{\ell\to+\infty} \int_{K_\ell} f(x)\ \mathrm{d}x.$$

Remarques

1. Si l'intégrale généralisée $\int_X f(x)\ \mathrm{d}x$ converge pour une suite exhaustive de compacts de X alors on montre qu'elle converge pour toute suite exhaustive de compacts de X. De plus, la valeur obtenue pour l'intégrale généralisée ne dépend pas de la suite exhaustive de compacts de X utilisée.

2. Reprenant la terminologie utilisée pour l'intégrale simple généralisée, si l'intégrale généralisée $\int_X f(x)\ \mathrm{d}x$ ne converge pas alors on dit qu'elle diverge.

3. On notera que d'après la définition 12.10 la convergence de l'intégrale généralisée $\int_X f(x)\ \mathrm{d}x$ équivaut à la convergence de l'intégrale généralisée $\int_X |f(x)|\ \mathrm{d}x$. Il n'existe donc pas de notion d'intégrale multiple semi-convergente contrairement au cas de l'intégrale simple[60]. ◇

[59] Nous admettrons le fait que cette définition est bien cohérente avec le but recherché, en particulier le fait que la valeur d'une intégrale généralisée convergente ne dépend pas de la suite exhaustive de compacts choisie.

[60] Voir le *Cours de première année* p. 940.

Exemples

1. Étudions l'intégrabilité sur $\mathbb{R}^2$ de $f : (x_1, x_2) \in \mathbb{R}^2 \mapsto \mathrm{e}^{-x_1^2-x_2^2}$. La suite $(K_\ell)_\ell$, où $K_\ell = \{(x_1, x_2) \in \mathbb{R}^2 \ ; \ x_1^2 + x_2^2 \leqslant \ell^2\}$ est le disque centré à l'origine de rayon ℓ, constitue une suite exhaustive de compacts de $\mathbb{R}^2$. Pour tout $\ell \in \mathbb{N}$, le compact K_ℓ est de frontière négligeable[61] et l'application f est continue et bornée sur K_ℓ ; l'application f est donc Riemann-intégrable[62] sur K_ℓ. De plus, comme f est positive on a

$$J_\ell = \int_{K_\ell} |f(x)| \, \mathrm{d}x = \int_{K_\ell} f(x) \, \mathrm{d}x = \int_{K_\ell} \mathrm{e}^{-x_1^2-x_2^2} \, \mathrm{d}x_1 \mathrm{d}x_2.$$

En effectuant un changement de variables polaire, on obtient

$$J_\ell = \int_0^{2\pi} \int_0^\ell \mathrm{e}^{-r^2} \, r \, \mathrm{d}r \, \mathrm{d}\theta = \int_0^{2\pi} \left[-\tfrac{1}{2}\mathrm{e}^{-r^2}\right]_0^\ell \mathrm{d}\theta = \pi(1 - \mathrm{e}^{-\ell^2}).$$

Comme $\lim_{\ell \to +\infty} J_\ell = \pi$, on en déduit que l'intégrale généralisée $\int_{\mathbb{R}^2} f(x) \, \mathrm{d}x$ converge. De plus, comme $\int_{K_\ell} f(x) \, \mathrm{d}x = J_\ell$ on a $\int_{\mathbb{R}^2} \mathrm{e}^{-x_1^2-x_2^2} \, \mathrm{d}x_1 \mathrm{d}x_2 = \pi$.

2. Considérons à nouveau l'intégrale généralisée $\int_{\mathbb{R}^2} \mathrm{e}^{-x_1^2-x_2^2} \, \mathrm{d}x_1 \mathrm{d}x_2$ et la suite exhaustive $(K_\ell)_\ell$ de compacts de $\mathbb{R}^2$ de terme général $K_\ell = [-\ell, \ell]^2$. Comme l'intégrale généralisée converge, on a d'après le calcul effectué à l'exercice précédent

$$\pi = \int_{\mathbb{R}^2} \mathrm{e}^{-x_1^2-x_2^2} \, \mathrm{d}x_1 \mathrm{d}x_2 = \lim_{\ell \to +\infty} \int_{K_\ell} \mathrm{e}^{-x_1^2-x_2^2} \, \mathrm{d}x_1 \mathrm{d}x_2.$$

Or

$$\begin{aligned} \int_{K_\ell} \mathrm{e}^{-x_1^2-x_2^2} \, \mathrm{d}x_1 \mathrm{d}x_2 &= \int_{-\ell}^{\ell} \int_{-\ell}^{\ell} \mathrm{e}^{-x_1^2-x_2^2} \, \mathrm{d}x_1 \mathrm{d}x_2 = \int_{-\ell}^{\ell} \mathrm{e}^{-x_1^2} \left(\int_{-\ell}^{\ell} \mathrm{e}^{-x_2^2} \, \mathrm{d}x_2 \right) \mathrm{d}x_1 \\ &= \left(\int_{-\ell}^{\ell} \mathrm{e}^{-x_2^2} \, \mathrm{d}x_2 \right) \left(\int_{-\ell}^{\ell} \mathrm{e}^{-x_1^2} \, \mathrm{d}x_1 \right) = \left(\int_{-\ell}^{\ell} \mathrm{e}^{-t^2} \, \mathrm{d}t \right)^2 \end{aligned}$$

et par conséquent

$$\lim_{\ell \to +\infty} \int_{K_\ell} \mathrm{e}^{-x_1^2-x_2^2} \, \mathrm{d}x_1 \mathrm{d}x_2 = \left(\int_{-\infty}^{+\infty} \mathrm{e}^{-t^2} \, \mathrm{d}t \right)^2.$$

On en déduit que

$$\boxed{\int_{-\infty}^{+\infty} \mathrm{e}^{-t^2} \, \mathrm{d}t = \sqrt{\pi}.}$$

En effectuant le changement de variables $s = t\sqrt{2}$, on obtient que

$$\boxed{\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \mathrm{e}^{-s^2/2} \, \mathrm{d}s = 1.}$$

[61] Cette affirmation a été établie p. 870.

[62] Voir la proposition 12.2 p. 871.

Cette relation joue un rôle important en théorie des probabilités où l'application $x \in \mathbb{R} \mapsto \mathrm{e}^{-x^2/2}$ est la densité de probabilité de la loi normale centrée réduite (encore appelée loi gaussienne).

3. Étudions l'intégrabilité sur $X = \{(x_1, x_2) \in \mathbb{R}^2 \ ; \ 0 < x_2^2 + x_2^2 \leqslant 1\}$ de l'application $f : (x_1, x_2) \in \mathbb{R}^2 \mapsto (x_1^2 + x_2^2)^{-\frac{1}{2}}$. La suite $(K_\ell)_\ell$ de terme général $K_\ell = \{(x_1, x_2) \in \mathbb{R}^2 \ ; \ 1/\ell^2 \leqslant x_1^2 + x_2^2 \leqslant 1\}$ constitue une suite exhaustive de compacts de X. Pour tout $\ell \in \mathbb{N}$, le compact K_ℓ est de frontière négligeable[63] et l'application f est continue et bornée sur K_ℓ ; l'application f est donc Riemann-intégrable[64] sur K_ℓ. De plus, comme f est positive on a

$$J_\ell = \int_{K_\ell} |f(x)| \, \mathrm{d}x = \int_{K_\ell} f(x) \, \mathrm{d}x = \int_{K_\ell} (x_1^2 + x_2^2)^{-\frac{1}{2}} \, \mathrm{d}x_1 \mathrm{d}x_2.$$

En effectuant un changement de variables polaire, on obtient

$$J_\ell = \int_0^{2\pi} \int_{1/\ell} \frac{1}{r} r \, \mathrm{d}r \, \mathrm{d}\theta = 2\pi \left(1 - \frac{1}{\ell}\right).$$

Comme $\lim_{\ell \to +\infty} J_\ell = 2\pi$, on en déduit que l'intégrale généralisée $\int_X f(x) \, \mathrm{d}x$ converge. De plus, comme $\int_{K_\ell} f(x) \, \mathrm{d}x = J_\ell$ on a $\int_X (x_1^2 + x_2^2)^{-\frac{1}{2}} \, \mathrm{d}x_1 \mathrm{d}x_2 = 2\pi$.

4. Étudions l'intégrabilité sur $X = \{(x_1, x_2) \in \mathbb{R}^2 \ ; \ 0 < x_2^2 + x_2^2 \leqslant 1\}$ de l'application $f : (x_1, x_2) \in \mathbb{R}^2 \mapsto 1/(x_1^2 + x_2^2)$. La suite $(K_\ell)_\ell$ de terme général $K_\ell = \{(x_1, x_2) \in \mathbb{R}^2 \ ; \ 1/\ell^2 \leqslant x_1^2 + x_2^2 \leqslant 1\}$ constitue une suite exhaustive de compacts de X. Pour tout $\ell \in \mathbb{N}$, le compact K_ℓ est de frontière négligeable et l'application f est continue et bornée sur K_ℓ ; l'application f est donc Riemann-intégrable sur K_ℓ. De plus, comme f est positive on a

$$J_\ell = \int_{K_\ell} |f(x)| \, \mathrm{d}x = \int_{K_\ell} f(x) \, \mathrm{d}x = \int_{K_\ell} (x_1^2 + x_2^2)^{-1} \, \mathrm{d}x_1 \mathrm{d}x_2.$$

En effectuant un changement de variables polaire, on obtient

$$J_\ell = \int_0^{2\pi} \int_{1/\ell}^1 \frac{1}{r^2} r \, \mathrm{d}r \, \mathrm{d}\theta = 2\pi \int_{1/\ell}^1 \frac{1}{r} \, \mathrm{d}r = 2\pi \Big[\ln(r)\Big]_{1/\ell}^1 = 2\pi \ln(\ell).$$

Comme $\lim_{\ell \to +\infty} J_\ell = +\infty$, on en déduit que l'intégrale généralisée $\int_X 1/(x_1^2 + x_2^2) \, \mathrm{d}x_1 \mathrm{d}x_2$ diverge.

---

[63] Cette affirmation peut être établie à partir des exemples de la p. 870.

[64] Voir la proposition 12.2 p. 871.

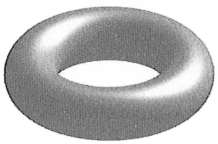
12.5 Exercices de synthèse

Exercice 12.9 ***Second théorème de Guldin***[65]

1 - Dans l'espace muni d'un repère orthonormé direct $(O; \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$, on considère un tore T engendré par la rotation autour de l'axe $(O; \mathbf{e}_3)$ d'un disque $D = \{(0, x_2, x_3) \in \mathbb{R}^3 \ ; \ (x_2 - a)^2 + x_3^2 \leqslant R^2\}$ de centre $(0, a, 0)$ et de rayon R avec $0 < R < a$.

a) En considérant un changement de variables en coordonnées cylindriques, montrer que le volume V du tore T est égal à $2\pi \int_\Delta r \, \mathrm{d}r \, \mathrm{d}z$ où $\Delta = \{(r, z) \in \mathbb{R}^2 \ ; \ (r - a)^2 + z^2 \leqslant R^2\}$.

b) Calculer le volume du tore T.



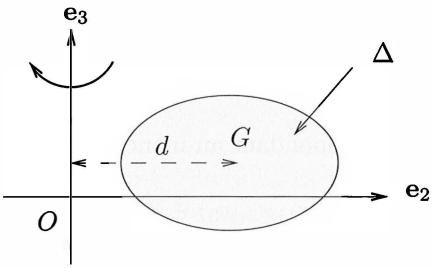


Fig. 35 Un tore (à gauche) ; solide engendré par la rotation du domaine Δ autour de l'axe $(O; \mathbf{e}_3)$ (à droite).

2 - Dans l'espace muni d'un repère orthonormé direct $(O; \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$, on considère un domaine Δ situé dans le demi-plan $\{(x_1, x_2, x_3) \in \mathbb{R}^3 \ ; \ x_1 = 0, \ x_2 > 0\}$. On désigne par A l'aire du domaine Δ et par G le centre d'inertie[66] de Δ. Montrer que le volume du solide de révolution Ω engendré par la rotation de Δ autour de l'axe $(O; \mathbf{e}_3)$ est égal à[67] $2\pi Ad$ où d désigne la distance du point G à la droite $(O; \mathbf{e}_3)$.

Exercice 12.10 ***Moments d'inertie d'un balai***

1 - Déterminer[68] les moments d'inertie d'une boule de $\mathbb{R}^3$ de rayon $\widetilde{R}$ et de densité volumique de masse constante ρ par rapport aux trois axes du repère canonique de $\mathbb{R}^3$.

---

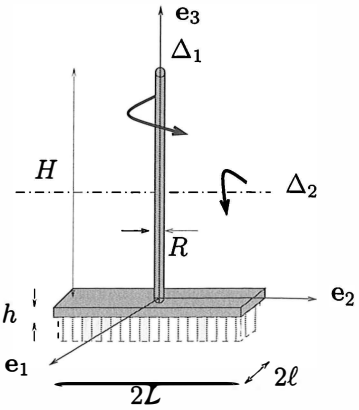
[65] Paul Guldin (1577-1643) Astronome et mathématicien suisse. Le premier théorème de Guldin exprime l'aire de la surface engendrée par la rotation d'un arc de courbe. Le second théorème de Guldin fait l'objet du présent exercice.

[66] Voir p. 875.

[67] En d'autre termes, le volume du solide engendré par la révolution d'un domaine plan autour d'un axe situé dans son plan et ne le coupant pas est égal au produit de l'aire A du domaine plan par la longueur de la circonférence décrite par son centre d'inertie ($2\pi d$). Ce résultat est connu sous le nom de second théorème de Guldin.

[68] La notion de moment d'inertie est présentée p. 875.

Considérons un balai tel celui représenté ci-contre formé d'un manche en bois de hauteur H, de rayon R et de densité volumique de masse constante $\rho = 380\,\mathrm{kg}\cdot\mathrm{m}^{-3}$ (cette densité correspond à du bois de peuplier) et d'une brosse du même bois ayant la forme d'un parallélépipède rectangle de longueur L, de largeur ℓ et de hauteur h. On introduit un repère orthonormé direct $(O;\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3)$ dont l'origine O est le point de contact entre l'axe de symétrie du manche avec la brosse, le vecteur $\mathbf{e}_3$ étant orienté vers le haut selon l'axe du balai.



Désignons par Ω le domaine de l'espace correspondant à ce balai, par

$$\Omega_1 = \left\{(x_1,x_2,x_3)\in\mathbb{R}^3\ ;\ 0\leqslant x_3\leqslant H\ \text{ et }\ 0\leqslant x_1^2+x_2^2\leqslant R^2\right\}$$

celui correspondant au manche du balai et par

$$\Omega_2 = \left\{(x_1,x_2,x_3)\in\mathbb{R}^3\ ;\ -L\leqslant x_1\leqslant L,\ -\ell\leqslant x_2\leqslant \ell,\ -h\leqslant x_3\leqslant 0\right\}$$

celui correspondant à la brosse (on omet de prendre en considération les poils du balai).

2 - Calculer le moment d'inertie J_{Δ_1} du balai par rapport à l'axe Δ_1 coïncidant avec l'axe de symétrie du manche.

3 - Calculer le moment d'inertie J_{Δ_2} du balai par rapport à un axe Δ_2 horizontal, coupant le manche du balai perpendiculairement à celui-ci à mi-hauteur.

4 - On suppose que le balai à les dimensions suivantes :

$$H = 1.5\ \mathrm{m},\quad R = 2\ \mathrm{cm},\quad L = 20\ \mathrm{cm},\quad \ell = 6\ \mathrm{cm}\quad\text{et}\quad h = 3\ \mathrm{cm}.$$

Déterminer les valeurs approchées de J_{Δ_1} et J_{Δ_2}. Pour quel rayon $\widetilde{R}$ le moment d'inertie de la boule de $\mathbb{R}^3$ de rayon $\widetilde{R}$ et de densité $380\,\mathrm{kg}\cdot\mathrm{m}^{-3}$ par rapport à un axe du repère canonique de $\mathbb{R}^3$ est-il égal à J_{Δ_1} ? à J_{Δ_2} ?

12.6 Solutions des exercices

Solution de l'exercice 12.1

D'après la définition 12.4, l'intégrale de ψ_n sur le pavé $[0,1] \times [0,1]$ vaut

$$\int_{[0,1]^2} \psi_n = \sum_{i=1}^{n} \sum_{j=1}^{n} m(C_{ij}) \; \frac{i-1}{n} \times \frac{j-1}{n}.$$

Étant donné que la mesure de toutes les cellules C_{ij} est identique et vaut $1/n^2$, on obtient :

$$\begin{aligned} \int_{[0,1]^2} \psi_n &= \frac{1}{n^4} \sum_{i=1}^{n} \sum_{j=1}^{n} (i-1)(j-1) = \frac{1}{n^4} \left(\sum_{k=0}^{n-1} k \right) \left(\sum_{\ell=0}^{n-1} \ell \right) \\ &= \frac{1}{n^4} \frac{(n-1)n}{2} \frac{(n-1)n}{2} = \frac{(n-1)^2}{4n^2} \end{aligned}$$

car $\displaystyle\sum_{i=1}^{n-1} i = \frac{1}{2} n(n-1)$.

Solution de l'exercice 12.2

L'application f est continue sur P et d'après le théorème de Fubini, on a

$$\begin{aligned} I &= \int_{\mathrm{P}} f(x_1, x_2) \, \mathrm{d}x_1 \, \mathrm{d}x_2 = \int_0^1 \left(\int_0^1 \sqrt{x_1^2 + x_2} \, \mathrm{d}x_2 \right) \mathrm{d}x_1 \\ &= \int_0^1 \left[\tfrac{2}{3} \left(x_1^2 + x_2 \right)^{\frac{3}{2}} \right]_0^1 \mathrm{d}x_1 = \frac{2}{3} \int_0^1 \left((x_1^2 + 1)^{\frac{3}{2}} - x_1^3 \right) \mathrm{d}x_1 \\ &= -\frac{1}{6} + \int_0^1 (x_1^2 + 1)^{\frac{3}{2}} \, \mathrm{d}x_1. \end{aligned}$$

On a choisi d'intégrer d'abord relativement à la variable x_2 car la primitive de la fonction $x_2 \mapsto \sqrt{x_1^2 + x_2}$ est facilement accessible alors que celle de la fonction $x_1 \mapsto \sqrt{x_1^2 + x_2}$ ne l'est pas.

Pour calculer l'intégrale simple $\int_0^1 (x_1^2 + 1)^{\frac{3}{2}} \, \mathrm{d}x_1$ plusieurs méthodes sont envisageables. En intégrant une première fois par partie, on obtient

$$J = \int_0^1 (x_1^2 + 1)^{\frac{3}{2}} \, \mathrm{d}x_1 = \left[x_1 (x_1^2 + 1)^{\frac{3}{2}} \right]_0^1 - 3 \int_0^1 x_1^2 \, (x_1^2 + 1)^{\frac{1}{2}} \, \mathrm{d}x_1 = 2^{\frac{3}{2}} - 3K$$

où, en écrivant $x_1^2 = (x_1^2 + 1) - 1$,

$$K = \int_0^1 x_1^2 \, (x_1^2 + 1)^{\frac{1}{2}} \, \mathrm{d}x_1 = \int_0^1 (x_1^2 + 1)^{\frac{3}{2}} \, \mathrm{d}x_1 - \int_0^1 (x_1^2 + 1)^{\frac{1}{2}} \, \mathrm{d}x_1.$$

On a $K = J - L$ où $L = \int_0^1 (x_1^2 + 1)^{\frac{1}{2}} \, \mathrm{d}x_1$. En intégrant une seconde fois par partie, on obtient

$$L = \left[x_1 (x_1^2 + 1)^{\frac{1}{2}} \right]_0^1 - \int_0^1 x_1^2 \, (x_1^2 + 1)^{-\frac{1}{2}} \, \mathrm{d}x_1 = \sqrt{2} - M,$$

où, en écrivant $x_1^2 = (x_1^2+1) - 1$,

$$\begin{aligned} M &= \int_0^1 x_1^2\,(x_1^2+1)^{-\frac{1}{2}}\,\mathrm{d}x_1 = \int_0^1 (x_1^2+1)^{\frac{1}{2}}\,\mathrm{d}x_1 - \int_0^1 (x_1^2+1)^{-\frac{1}{2}}\,\mathrm{d}x_1 \\ &= L - \Big[\operatorname{argsh}(x_1)\Big]_0^1 = L - \operatorname{argsh}(1). \end{aligned}$$

On a donc $M = L - \operatorname{argsh}(1)$ et $L = \sqrt{2} - M$; en résolvant ce système linéaire de deux équations à deux inconnues, on établit que $L = \frac{1}{\sqrt{2}} + \frac{1}{2}\operatorname{argsh}(1)$. Par ailleurs, on a $K = J - L$ et $J = 2^{\frac{3}{2}} - 3K$ ce qui nous donne $4J = 2^{\frac{3}{2}} + 3L$ donc $J = \frac{7}{4\sqrt{2}} + \frac{3}{8}\operatorname{argsh}(1)$. On en conclut que

$$I = \int_{\mathrm{P}} f(x_1,x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2 = -\frac{1}{6} + \frac{7}{4\sqrt{2}} + \frac{3}{8}\operatorname{argsh}(1).$$

Signalons qu'une autre méthode de calcul de l'intégrale $J = \int_0^1 (x_1^2+1)^{\frac{3}{2}}\,\mathrm{d}x_1$ consiste à effectuer le changement de variables $x_1 = \operatorname{sh}(t)$.

Solution de l'exercice 12.3

On peut écrire l'ensemble X sous la forme

$$X = \Big\{(x_1,x_2) \in \mathbb{R}^2 \ ;\ a_1 \leqslant x_1 \leqslant b_1 \quad \text{et} \quad g_2(x_1) \leqslant x_2 \leqslant h_2(x_1)\Big\}$$

où $a_1 = -R$, $a_2 = R$, $g_2 : x_1 \mapsto -\sqrt{R^2 - x_1^2}$ et $h_2 : x_1 \mapsto \sqrt{R^2 - x_1^2}$. En intégrant par pile, l'intégrale double $I = \int_X 1\,\mathrm{d}x_1\mathrm{d}x_2$ s'écrit

$$I = \int_{-R}^{R} \left(\int_{-\sqrt{R^2-x_1^2}}^{\sqrt{R^2-x_1^2}} 1\,\mathrm{d}x_2 \right) \mathrm{d}x_1 = \int_{-R}^{R} 2\sqrt{R^2 - x_1^2}\,\mathrm{d}x_1.$$

Considérons le changement de variables défini par l'application de classe $\mathcal{C}^1$ $\varphi : \theta \in]-\pi,\pi[\mapsto R\sin(\theta)$. On a $\varphi'(\theta) = R\cos(\theta)$ et

$$\sqrt{R^2 - \varphi(\theta)^2} = R\sqrt{1-\sin(\theta)} = R\sqrt{\cos^2(\theta)} = R|\cos(\theta)|.$$

La formule de changement de variables pour l'intégrale simple[(69)] permet d'écrire

$$\begin{aligned} I &= 2\int_{-R}^{R} \sqrt{R^2 - x_1^2}\,\mathrm{d}x_1 = 2\int_{\varphi(-\frac{\pi}{2})}^{\varphi(-\frac{\pi}{2})} \sqrt{R^2 - x_1^2}\,\mathrm{d}x_1 \\ &= 2\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sqrt{R^2 - \varphi(\theta)^2}\ \varphi'(\theta)\ \mathrm{d}\theta = 2\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} R^2|\cos(\theta)|\cos(\theta)\ \mathrm{d}\theta. \end{aligned}$$

Comme la fonction cosinus est positive sur $[-\frac{\pi}{2},\frac{\pi}{2}]$, on a

$$|\cos(\theta)|\cos(\theta) = \cos(\theta)^2 = \tfrac{1}{2}(\cos(2\theta)+1)$$

---

(69) Voir le théorème 18.2 p. 863 du *Cours de première année*.

et comme la fonction cosinus est 2π-périodique, d'intégrale nulle sur une période, on obtient

$$I = R^2 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} (\cos(2\theta) + 1) \ \mathrm{d}\theta = R^2 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \mathrm{d}\theta = \pi R^2.$$

SOLUTION DE L'EXERCICE 12.4

1 - Le volume V de la calotte sphérique X est obtenu en calculant l'intégrale triple $V = \int_X 1 \ \mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3$. Utilisons une intégration par tranche pour calculer V.

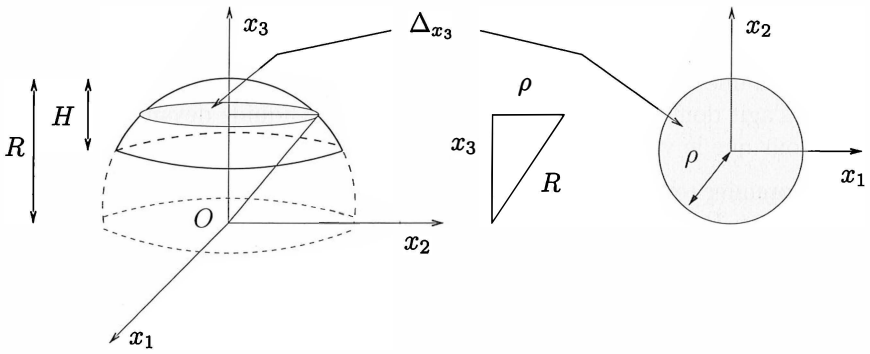


Fig. 36 Illustration de l'intégration par tranche.

L'ensemble X est décrit par tranche de la manière suivante :

$$X = \left\{(x_1, x_2, x_3) \in \mathbb{R}^3 \ ; \ R - H \leqslant x_3 \leqslant R \ \text{ et } \ (x_1, x_2) \in \Delta_{x_3}\right\}$$

où Δ_{x_3} est le sous-ensemble de $\mathbb{R}^2$ correspondant au disque de centre $(0,0)$ et dont le rayon ρ, déterminé en utilisant le théorème de Pythagore, vaut $\rho = \sqrt{R^2 - x_3^2}$.

En utilisant une intégration par tranche, on obtient

$$V = \int_X 1 \ \mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3 = \int_{R-H}^{R} \left(\int_{\Delta_{x_3}} 1 \ \mathrm{d}x_1\mathrm{d}x_2 \right) \mathrm{d}x_3.$$

L'intégrale double $\int_{\Delta_{x_3}} 1 \ \mathrm{d}x_1\mathrm{d}x_2$ s'interprète comme l'aire du disque Δ_{x_3} ; elle est donc égale à $\pi\rho^2$. On en déduit que

$$\begin{aligned} V &= \pi \int_{R-H}^{R} (R^2 - x_3^2) \ \mathrm{d}x_3 = \pi \int_{R-H}^{R} R^2 \ \mathrm{d}x_3 - \pi \left[\tfrac{1}{3}x_3^3\right]_{R-H}^{R} \\ &= \pi R^2 H - \frac{\pi}{3}\left(R^3 - (R-H)^3\right) = \frac{\pi}{3} H^2 (3R - H). \end{aligned}$$

2 - La canette X est composée d'un cylindre X_1 auquel on retranche la calotte X_2. La relation de Chasles permet d'écrire

$$V = \int_X 1 \ \mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3 = \int_{X_1} 1 \ \mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3 - \int_{X_2} 1 \ \mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3 = V_1 - V_2$$

où V_1 désigne le volume du cylindre et V_2 celui de la calotte. En reprenant les notations de l'exemple traité p. 888, on a $V_1 = \pi R^2 H$ et $V_2 = \frac{1}{3}\pi h^2(3\rho - h)$ où $\rho = \frac{R^2+h^2}{2h}$. On obtient

$$V = \pi R^2 H - \frac{1}{3}\pi h^2 \left(3\,\frac{R^2+h^2}{2h} - h\right) = \frac{\pi}{6}\left(3R^2(2H-h) - h^3\right)$$

ce qui correspond au résultat obtenu p. 888 en effectuant une intégration par pile.

Solution de l'exercice 12.5

Le volume d'une sphère de rayon R étant égal à $\frac{4}{3}\pi R^3$, celui de la louche vaut $\frac{2}{3}\pi R^3$. Il s'agit donc de déterminer la hauteur H de liquide devant être dans la louche pour que le volume du liquide soit égal à $\frac{1}{3}\pi R^3$.

Dans un premier temps, nous avons à déterminer le volume d'une calotte sphérique de rayon R et de hauteur H avec $H \in]0, R[$ puis nous aurons à déterminer pour quelle valeur de H on a $V = \frac{1}{3}\pi R^3$.

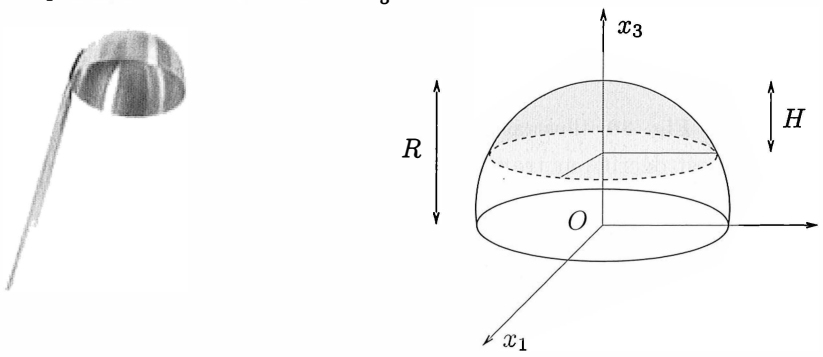


Fig. 37 Illustration de la situation étudiée.

Le volume V d'une calotte sphérique de rayon R et de hauteur H avec $H \in]0, R[$ a été calculé à l'exercice 12.4. On a $V = \frac{\pi}{3}H^2(3R - H)$. Il reste à déterminer la valeur de H pour que $V = \frac{1}{3}\pi R^3$, autrement dit à résoudre l'équation suivante d'inconnue H :

$$\frac{\pi}{3}H^2(3R - H) = \frac{\pi}{3}R^3$$

En posant $t = R/H$, résoudre l'équation précédente revient à trouver les solutions $t > 1$ de l'équation $t^3 - 3t + 1 = 0$.

L'étude de la fonction $t \mapsto t^3 - 3t + 1$ permet d'établir qu'il existe une unique solution à cette équation supérieure à 1. Vérifions que cette solution est bien $t_0 = 2\cos(2\pi/9)$. On peut remarquer que $0 < 2\pi/9 < 3\pi/9 = \pi/3$; comme la fonction cosinus est décroissante sur $[0, \pi]$, on a $\cos(2\pi/9) > \cos(\pi/3) = 1/2$ donc $t_0 > 1$.

D'après les formules d'Euler, on a $t_0 = e^{i2\pi/9} + e^{-i2\pi/9}$ donc en utilisant la formule du binôme de Newton, on obtient

$$\begin{aligned} t_0^3 &= \left(e^{i2\pi/9} + e^{-i2\pi/9}\right)^3 = e^{i2\pi/3} + 3e^{i4\pi/9}e^{-i2\pi/9} + 3e^{i2\pi/9}e^{-i4\pi/9}e^{-i2\pi/3} \\ &= e^{i2\pi/3} + 3e^{i2\pi/9} + 3e^{-i2\pi/9} + e^{-i2\pi/3}. \end{aligned}$$

On en déduit que

$$t_0^3 - 3t_0 = e^{i2\pi/3} + e^{-i2\pi/3} = 2\cos(2\pi/3) = 1.$$

Le réel $t_0 = 2\cos(2\pi/9)$ est donc bien l'unique solution supérieure à 1 de l'équation $t^3 - 3t + 1 = 0$. On en conclut que pour que la louche soit à moitié pleine, elle doit être remplie à une hauteur

$$H = \frac{R}{t_0} \approx 0,653R.$$

Il faut donc que la louche soit approximativement remplie aux 2/3 de sa hauteur pour qu'elle soit à moitié pleine de liquide.

Solution de l'exercice 12.6

1 - Considérons la fonction $\varphi : (r,\theta) \mapsto (r\cos(\theta), r\sin(\theta))$ de changement de variables polaire sur l'ouvert $U = \{(r,\theta) \in \mathbb{R}^2 \ ; \ r \in]0,R[\text{ et } \theta \in]0,2\pi[\}$. L'application φ réalise un $\mathcal{C}^1$-difféomorphisme de U sur $\varphi(U)$ et l'adhérence de $\varphi(U)$ est D. On a donc

$$\begin{aligned} \int_D \sqrt{x_1^2 + x_2^2}\, dx_1 dx_2 &= \int_0^R \int_0^{2\pi} \sqrt{(r\cos(\theta))^2 + (r\cos(\theta))^2}\ r\ d\theta\ dr \\ &= \int_0^R \int_0^{2\pi} r^2\ d\theta\ dr = \int_0^R r^2 \left(\int_0^{2\pi} 1\ d\theta\right) dr \\ &= \int_0^R 2\pi r^2\ dr = \left[\tfrac{2}{3}\pi r^3\right]_0^R = \tfrac{2}{3}\pi R^3. \end{aligned}$$

2 - Commençons par visualiser l'ensemble D. Pour $(x_1,x_2) \in \mathbb{R}^2$, l'inéquation $x_1^2 + x_2^2 \geqslant 2x_2$ se réécrit $x_1^2 + (x_2 - 1)^2 \geqslant 1$; géométriquement il s'agit de l'extérieur du disque fermé de centre $(0,1)$ et de rayon 1. Par ailleurs, l'inéquation $x_1^2 + x_2^2 \leqslant 1$ décrit le disque fermé de centre $(0,0)$ et de rayon 1. On en déduit que l'ensemble D correspond géométriquement à l'intersection des demi-plans $x_1 \geqslant 0$ et $x_2 \geqslant 0$, du disque de centre $(0,0)$ et de rayon 1 et de l'extérieur du disque de centre $(0,1)$ et de rayon 1. Cet ensemble est représenté en hachuré sur la figure 38. Les coordonnées (x_1,x_2) du point Q intersection des deux cercles dans $\mathbb{R}^+ \times \mathbb{R}^+$ s'obtiennent en considérant simultanément les équations des deux cercles et en résolvant le système

$$\begin{cases} x_1^2 + x_2^2 &= 2x_2 \\ x_1^2 + x_2^2 &= 1 \end{cases}.$$

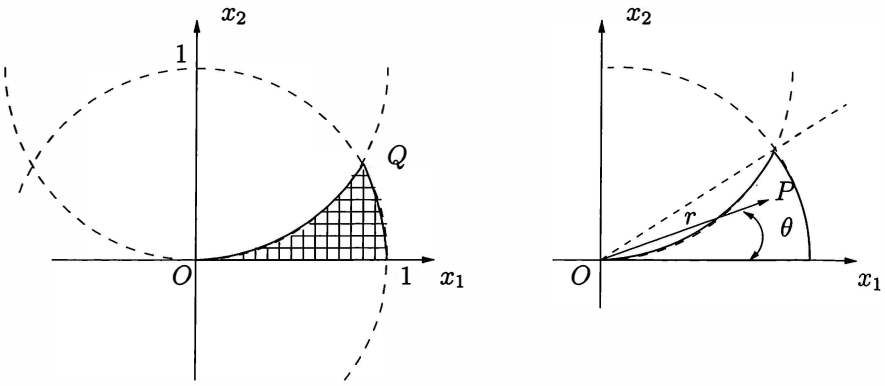


Fig. 38 Représentation graphique de l'ensemble D.

On obtient facilement que $x_2 = 1/2$, ce qui implique ensuite que $x_1^2 = 3/4$ et par conséquent que $x_1 = \sqrt{3}/2$.

Pour appliquer la formule de changement de variables polaire, il faut déterminer un ouvert U de $\mathbb{R}^+ \times \mathbb{R}$ tel que la fonction

$$\varphi : (r, \theta) \in \mathbb{R}^+ \times \mathbb{R} \longmapsto (r\cos(\theta), r\sin(\theta))$$

soit une bijection de U dans $\varphi(U)$ et que l'adhérence de $\varphi(U)$ soit D.

Remarquons que si $(x_1, x_2) = \varphi(r, \theta)$ alors $x_1^2 + x_2^2 \leqslant 1$ équivaut à $r \leqslant 1$ et $x_1^2 + x_2^2 \geqslant 2x_2$ équivaut à $r \geqslant 2\sin(\theta)$. Par ailleurs, le point Q a pour coordonnées cartésiennes $(\sqrt{3}/2, 1/2)$ ce qui implique qu'en polaire ses coordonnées sont $r = 1$ et $\theta = \frac{\pi}{6}$. On en déduit que $D = \varphi(\Delta)$ où

$$\Delta = \left\{(r, \theta) \in [0, 1] \times [0, \pi/6] \ ; \ r \geqslant 2\sin(\theta)\right\}$$

ce qui nous amène à prendre ouvert U l'ensemble suivant :

$$U = \left\{(r, \theta) \in]0, 1[\times]0, \pi/6[\ ; \ r > 2\sin(\theta)\right\}.$$

D'après la formule de changement de variables, on obtient :

$$\begin{aligned} I &= \int_D \sqrt{x_1^2 + x_2^2}\, \mathrm{d}x_1\mathrm{d}x_2 = \int_{\varphi(U)} \sqrt{x_1^2 + x_2^2}\, \mathrm{d}x_1\mathrm{d}x_2 = \int_U r^2\, \mathrm{d}\theta\, \mathrm{d}r \\ &= \int_0^{\frac{\pi}{6}} \left(\int_{2\sin(\theta)}^1 r^2\, \mathrm{d}r \right) \mathrm{d}\theta = \int_0^{\pi/6} \left[\tfrac{1}{3}r^3\right]_{2\sin(\theta)}^1 \mathrm{d}\theta \\ &= \int_0^{\pi/6} \tfrac{1}{3}(1 - 8\sin^3(\theta))\, \mathrm{d}\theta = \frac{1}{18}\pi - \frac{8}{3}\int_0^{\pi/6} \sin^3(\theta)\, \mathrm{d}\theta. \end{aligned}$$

Or

$$\begin{aligned}\int_0^{\pi/6} \sin^3(\theta)\ \mathrm{d}\theta &= \int_0^{\pi/6} \sin(\theta)(1-\cos^2(\theta))\ \mathrm{d}\theta \\ &= \int_0^{\pi/6} \sin(\theta)\ \mathrm{d}\theta - \int_0^{\pi/6} \cos^2(\theta)\sin(\theta)\ \mathrm{d}\theta \\ &= \big[-\cos(\theta)\big]_0^{\pi/6} + \big[\tfrac{1}{3}\cos^3(\theta)\big]_0^{\pi/6} = -\frac{3\sqrt{3}}{8} + \frac{2}{3}.\end{aligned}$$

Finalement, on obtient

$$\int_D \sqrt{x_1^2+x_2^2}\ \mathrm{d}x_1\mathrm{d}x_2 = \sqrt{3} + \frac{\pi}{18} - \frac{16}{9}.$$

Solution de l'exercice 12.7

1 - L'application $\varphi : (y_1,y_2,y_3) \mapsto (ay_1, by_2, cy_3) \in \mathbb{R}^3$ est une application de classe $\mathcal{C}^1$ dont le jacobien vaut

$$\begin{vmatrix} \frac{\partial\varphi_1}{\partial y_1}(y_1,y_2,y_3) & \frac{\partial\varphi_1}{\partial y_2}(y_1,y_2,y_3) & \frac{\partial\varphi_1}{\partial y_3}(y_1,y_2,y_3) \\ \frac{\partial\varphi_2}{\partial y_1}(y_1,y_2,y_3) & \frac{\partial\varphi_2}{\partial y_2}(y_1,y_2,y_3) & \frac{\partial\varphi_2}{\partial y_3}(y_1,y_2,y_3) \\ \frac{\partial\varphi_3}{\partial y_1}(y_1,y_2,y_3) & \frac{\partial\varphi_3}{\partial y_2}(y_1,y_2,y_3) & \frac{\partial\varphi_3}{\partial y_3}(y_1,y_2,y_3) \end{vmatrix} = \begin{vmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{vmatrix} = abc.$$

On notera que ce jacobien est constant et non nul. Par ailleurs, l'ellipsoïde

$$E = \left\{(x_1,x_2,x_3) \in \mathbb{R}^3\ ;\ \frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} + \frac{x_3^2}{c^2} < 1\right\}$$

est l'image par φ de la boule $B = \left\{(y_1,y_2,y_3) \in \mathbb{R}^3\ ;\ y_1^2+y_2^2+y_3^2 < 1\right\}$. On déduit du théorème d'inversion locale que φ est un $\mathcal{C}^1$-difféomorphisme de B dans E et en utilisant la formule de changement de variables, on obtient

$$\mathrm{Volume}(E) = \int_E 1\ \mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3 = \int_B |abc|\ \mathrm{d}y_1\mathrm{d}y_2\mathrm{d}y_3 = abc\,\mathrm{Volume}(B).$$

Comme on sait que le volume de la boule B de rayon 1 vaut $\frac{4}{3}\pi$, on en déduit que le volume délimité par l'ellipsoïde vaut $\frac{4}{3}\pi abc$.

2 - La masse M de l'ellipsoïde est donnée par $M = \int_E \rho(x_1,x_2,x_3)\ \mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3$. En utilisant le changement de variables défini par φ, on obtient

$$M = \int_B \rho(\varphi(y_1,y_2,y_3))\ abc\ \mathrm{d}y_1\mathrm{d}y_2\mathrm{d}y_3 = abc\int_B (y_1^2+y_2^2+y_3^2)\ \ \mathrm{d}y_1\mathrm{d}y_2\mathrm{d}y_3.$$

Pour calculer cette intégrale, utilisons un changement de variables en coordonnées sphériques. On obtient

$$\begin{aligned} M &= abc\int_0^1\int_0^{2\pi}\int_0^{\pi} r^2\ r\sin(\phi)\ \mathrm{d}r\ \mathrm{d}\theta\ \mathrm{d}\phi = 2\pi abc\int_0^1\int_0^{\pi} r^3\sin(\phi)\ \mathrm{d}r\ \mathrm{d}\phi \\ &= 2\pi abc\int_0^1 r^3\Big[-\cos(\phi)\Big]_0^{\pi}\ \mathrm{d}r = 4\pi abc\int_0^1 r^3\ \mathrm{d}r = \pi abc.\end{aligned}$$

La masse de l'ellipsoïde vaut donc πabc unités de masse.

Solution de l'exercice 12.8

Pour tout $\ell \in \mathbb{N}$, l'ensemble K_ℓ est la boule fermée de centre 0 et de rayon ℓ dans $\mathbb{R}^2$ pour la norme de la convergence uniforme $\|\cdot\|_\infty$; c'est donc un ensemble fermé et borné, autrement dit un ensemble compact d'après le théorème 7.7 p. 415. Il est clair que pour tout $\ell \in \mathbb{N}$ on a $K_\ell \subset K_{\ell+1}$. Si K désigne un compact de $\mathbb{R}^2$, K est borné et d'après la définition 7.11 p. 382, il existe un réel strictement positif r tel que $K \subset \overline{B}(0,r)$; on en déduit que $K \subset \overline{B}(0,\ell) = K_\ell$ pour $\ell = \mathrm{E}(r) + 1$.

Notons que le fait que sur $\mathbb{R}^2$ toutes les normes sont équivalentes (voir le théorème 7.2 p. 367) laisse la possibilité de faire le choix de la norme pour laquelle le raisonnement est le plus simple. En l'occurrence, ici la norme de la convergence uniforme est mieux adaptée que la norme euclidienne pour le raisonnement.

Plus généralement, on montre sur le même principe que la suite $(K_\ell)_\ell$ où $K_\ell = [-\ell, \ell]^n$ constitue une suite exhaustive de compacts de $\mathbb{R}^n$.

Solution de l'exercice 12.9

1 - a) Le volume V du tore T est donné par $V = \int_T 1 \, \mathrm{d}x$. Puisque le tore T est engendré par la rotation autour de l'axe $(O; \mathbf{e}_3)$ du disque D de centre $(0, a, 0)$ et de rayon R avec $0 < R < a$, dans la base des coordonnées cylindriques (r, θ, z) le tore T est décrit par l'ensemble $\{(r,\theta,z) \in \mathbb{R}^3 \ ; \ \theta \in [0, 2\pi] \ \text{ et } \ (r,z) \in \Delta\}$ où $\Delta = \{(r,z) \in \mathbb{R}^2 \ ; \ (r-a)^2 + z^2 \leqslant R^2\}$.

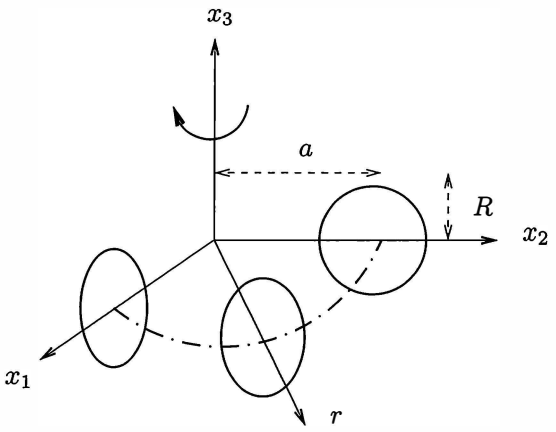


Fig. 39 Obtention d'un tore par rotation d'un disque autour d'un axe.

Considérons le changement de variables cylindrique défini par l'application $\varphi : (r,\theta,z) \in U \mapsto (r\cos(\theta), r\sin(\theta), z)$ où

$$U = \Big\{(r,\theta,z) \in \mathbb{R}^3 \ ; \ \theta \in]0, 2\pi[\ \text{ et } \ (r-a)^2 + z^2 < R^2\Big\}.$$

L'application φ est un $\mathcal{C}^1$-difféomorphisme de U sur $\varphi(U)$ et on a $\overline{\varphi(U)} = T$.

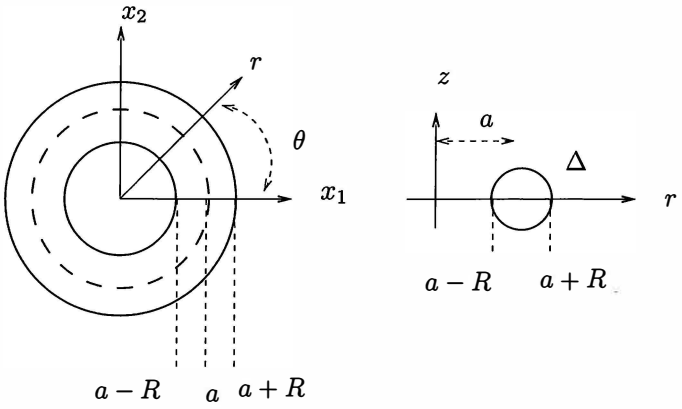


Fig. 40 Représentation en coupe du tore T dans les plans d'équation $x_3 = 0$ (à gauche) et $\theta = 0$ (à droite).

La formule de changement de variables indique que

$$\begin{aligned} V &= \int_T 1 \, \mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3 = \int_{\varphi(U)} 1 \, \mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3 = \int_U r \, \mathrm{d}r\mathrm{d}\theta\mathrm{d}z \\ &= \int_\Delta \left(\int_0^{2\pi} r \, \mathrm{d}\theta \right) \mathrm{d}r\mathrm{d}z = 2\pi \int_\Delta r \, \mathrm{d}r\mathrm{d}z. \end{aligned}$$

b) Pour obtenir la valeur du volume V, il convient de calculer l'intégrale double $I = \int_\Delta r \, \mathrm{d}r\mathrm{d}z$ où $\Delta = \{(r,z) \in \mathbb{R}^2 \, ; \, (r-a)^2 + z^2 \leqslant R^2\}$ est le disque de centre $(a,0)$ et de rayon R du plan, voir la figure 40. Considérons le changement de variables $\varphi : (u,v) \in \Delta_0 \mapsto (u+a, v) \in \Delta$ où Δ_0 désigne le disque ouvert de rayon R centré à l'origine. L'application φ est un $\mathcal{C}^1$-difféomorphisme de Δ_0 sur l'intérieur de Δ et la formule de changement de variables indique que

$$I = \int_\Delta r \, \mathrm{d}r\mathrm{d}z = \int_{\varphi(\Delta_0)} r \, \mathrm{d}r\mathrm{d}z = \int_{\Delta_0} (u+a) \, \mathrm{d}u\mathrm{d}v = \int_{\Delta_0} u \, \mathrm{d}u\mathrm{d}v + a \int_{\Delta_0} \mathrm{d}u\mathrm{d}v.$$

On obtient rapidement les valeurs des deux intégrales du membre de droite en remarquant que l'intégrale $\int_{\Delta_0} \mathrm{d}u\mathrm{d}v$ s'interprète comme l'aire du disque Δ_0 (qui vaut πR^2) et que l'intégrale $\frac{1}{2\pi R^2} \int_{\Delta_0} u \, \mathrm{d}u\mathrm{d}v$ s'interprète comme[(70)] l'abscisse du centre d'inertie du disque Δ_0 supposé homogène (l'intégrale est donc nulle car le centre d'inertie de Δ_0 est l'origine pour des raisons de symétrie). On en déduit que $I = \pi R^2 a$ et par conséquent que $V = 2\pi^2 R^2 a$.

[(70)] Voir p. 875. On retrouve ce résultat en remarquant que

$$\frac{1}{2\pi R^2} \int_{\Delta_0} u \, \mathrm{d}u\mathrm{d}v = \int_{-R}^{R} \left(\int_{-\sqrt{R^2-v^2}}^{\sqrt{R^2-v^2}} u \, \mathrm{d}u \right) \mathrm{d}v$$

et que $\int_{-\sqrt{R^2-v^2}}^{\sqrt{R^2-v^2}} u \, \mathrm{d}u = 0$ car on intègre une fonction impaire sur un intervalle de centre 0, voir la proposition 18.15 p.867 du *Cours de première année*.

On peut également calculer ces intégrales en effectuant un changement de variables polaire. On a[71]

$$\int_{\Delta_0} u \, \mathrm{d}u\mathrm{d}v = \int_0^R \left(\int_0^{2\pi} \rho^2 \cos(\alpha) \, \mathrm{d}\alpha \right) \mathrm{d}\rho = \int_0^R \rho^2 \left(\int_0^{2\pi} \cos(\alpha) \, \mathrm{d}\alpha \right) \mathrm{d}\rho = 0$$

car $\int_0^{2\pi} \cos(\alpha) \, \mathrm{d}\alpha = 0$. De même,

$$\int_{\Delta_0} \mathrm{d}u\mathrm{d}v = \int_0^R \left(\int_0^{2\pi} \rho \, \mathrm{d}\alpha \right) \mathrm{d}\rho = 2\pi \int_0^R \rho \, \mathrm{d}\rho = \pi R^2.$$

2 - Pour établir le second théorème de Guldin, reprenons le raisonnement effectué pour calculer le volume du tore. Le volume du solide de révolution Ω engendré par la rotation de l'ensemble Δ autour de l'axe $(O; \mathbf{e}_3)$ est égal à $V = \int_\Omega 1 \, \mathrm{d}x$.

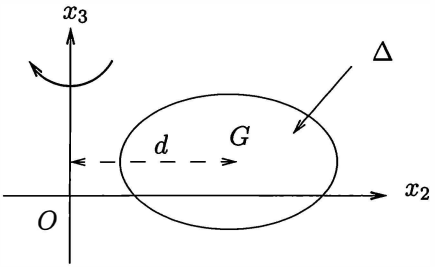


Fig. 41 Représentation du domaine Δ.

Considérons le changement de variables cylindrique défini par l'application $\varphi : (r, \theta, z) \in U \mapsto (r\cos(\theta), r\sin(\theta), z)$ où

$$U = \Big\{ (r, \theta, z) \in \mathbb{R}^3 \ ; \ \theta \in]0, 2\pi[\ \text{ et } \ (r, z) \in \Delta \Big\}.$$

L'application φ est un $\mathcal{C}^1$-difféomorphisme de U sur $\varphi(U)$ avec $\overline{\varphi(U)} = \Omega$ et la formule de changement de variables indique que

$$\begin{aligned} V &= \int_\Omega 1 \, \mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3 = \int_{\varphi(U)} 1 \, \mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3 = \int_U r \, \mathrm{d}r\mathrm{d}\theta\mathrm{d}z \\ &= \int_\Delta \left(\int_0^{2\pi} r \, \mathrm{d}\theta \right) \mathrm{d}r\mathrm{d}z = 2\pi \int_\Delta r \, \mathrm{d}r\mathrm{d}z. \end{aligned}$$

En désignant par A l'aire du domaine Δ, l'intégrale $\frac{1}{A} \int_\Delta r \, \mathrm{d}r\mathrm{d}z = \frac{1}{A} \int_\Delta x_2 \, \mathrm{d}x_2\mathrm{d}x_3$ s'interprète comme[72] l'ordonnée x_2^G du centre d'inertie G du domaine Δ supposé homogène. On a donc $V = 2\pi A \, x_2^G$ et dans le plan $x_1 = 0$, le réel x_2^G correspond bien à la distance d du point G à la droite $(O; \mathbf{e}_3)$, voir la figure 42.

---

[71] On désigne ici les coordonnées polaires par (ρ, α) afin d'éviter toute confusion avec les notations (r, θ) utilisées auparavant pour les coordonnées cylindriques.

[72] Voir p. 875.

Solution de l'exercice 12.10

1 - Puisque la densité volumique de masse de la boule est supposée constante, le moment d'inertie de la boule par rapport à l'axe $(O;\mathbf{e}_1)$ est donné par

$$J_1 = \rho \int_{B(0,\widetilde{R})} d(\mathbf{x})^2 \, \mathrm{d}x.$$

où $d(\mathbf{x})$ désigne la distance d'un point $\mathbf{x}$ de la boule de coordonnées (x_1, x_2, x_3) à l'axe $(O;\mathbf{e}_1)$ et vaut $\sqrt{x_2^2 + x_3^2}$. On a donc

$$J_1 = \rho \int_{B(0,\widetilde{R})} (x_2^2 + x_3^2) \, \mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3.$$

De même, les moments d'inertie de la boule par rapport aux axes $(O;\mathbf{e}_2)$ et $(O;\mathbf{e}_3)$ sont donnés par

$$J_2 = \rho \int_{B(0,\widetilde{R})} (x_1^2 + x_3^2) \, \mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3, \quad J_3 = \rho \int_{B(0,\widetilde{R})} (x_1^2 + x_2^2) \, \mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3.$$

Pour des raisons de symétrie, les moments d'inertie d'une boule de $\mathbb{R}^3$ de rayon $\widetilde{R}$ et de densité constante ρ par rapport aux trois axes sont égaux : $J_1 = J_2 = J_3$. Le calcul des moments d'inertie est facilité si on remarque que cette égalité entraîne que

$$J_1 = \tfrac{1}{3}(J_1 + J_2 + J_3) = \frac{2\rho}{3} \int_{B(0,\widetilde{R})} (x_1^2 + x_2^2 + x_3^2) \, \mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3.$$

En effectuant un changement de variables sphérique, on obtient

$$\begin{aligned} J_1 &= \frac{2\rho}{3} \int_0^{\widetilde{R}} \int_0^{\pi} \int_0^{2\pi} r^4 \sin(\phi) \, \mathrm{d}\theta\mathrm{d}\phi\mathrm{d}r = \frac{4\pi\rho}{3} \int_0^{\widetilde{R}} \int_0^{2\pi} r^4 \sin(\phi) \, \mathrm{d}\phi\mathrm{d}r \\ &= \frac{8\pi\rho}{3} \int_0^{\widetilde{R}} r^4 \, \mathrm{d}r = \frac{8\pi\rho\widetilde{R}^5}{15}. \end{aligned}$$

2 - Compte tenu de la relation de Chasles, le moment d'inertie du balai par rapport à l'axe Δ_1 coïncidant avec l'axe de symétrie du manche est :

$$J_{\Delta_1} = \int_\Omega \rho \, d(\mathbf{x}, \Delta_1)^2 \, \mathrm{d}x = \rho \int_{\Omega_1} d(\mathbf{x}, \Delta_1)^2 \, \mathrm{d}x + \rho \int_{\Omega_2} d(\mathbf{x}, \Delta_1)^2 \, \mathrm{d}x.$$

La distance d'un point $\mathbf{x}$ du balai de coordonnées (x_1, x_2, x_3) à l'axe Δ_1 est $d(\mathbf{x}, \Delta_1) = \sqrt{x_1^2 + x_2^2}$. On a donc

$$\begin{aligned} J_{11} &= \int_{\Omega_1} d(\mathbf{x}, \Delta_1)^2 \, \mathrm{d}x = \int_{\Omega_1} (x_1^2 + x_2^2) \, \mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3 \\ &= \int_0^H \int_0^{2\pi} \int_0^R r^3 \, \mathrm{d}r\mathrm{d}\theta\mathrm{d}x_3 = 2\pi H \int_0^R r^3 \, \mathrm{d}r = \frac{\pi}{2} R^4 H \end{aligned}$$

et

$$\begin{aligned} J_{12} &= \int_{\Omega_2} d(\mathbf{x}, \Delta_1)^2 \, \mathrm{d}x = \int_{\Omega_2} (x_1^2 + x_2^2) \, \mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3 \\ &= \int_{-h}^0 \int_0^L \int_0^\ell (x_1^2 + x_2^2) \, \mathrm{d}x_1 \, \mathrm{d}x_2 \, \mathrm{d}x_3 = \frac{1}{3} L\ell h(\ell^2 + L^2). \end{aligned}$$

On en déduit que le moment d'inertie du balai par rapport à l'axe Δ_1 vaut

$$J_{\Delta_1} = \frac{\pi}{2} R^4 H \rho + \frac{1}{3} L \ell h (\ell^2 + L^2) \rho.$$

3 - Le moment d'inertie du balai par rapport à un axe Δ_2 horizontal, coupant le manche du balai perpendiculairement à celui-ci à mi hauteur, est :

$$J_{\Delta_2} = \int_\Omega \rho\, d(\mathbf{x}, \Delta_2)^2 \, \mathrm{d}x = \rho \int_{\Omega_1} d(\mathbf{x}, \Delta_2)^2 \, \mathrm{d}x + \rho \int_{\Omega_2} d(\mathbf{x}, \Delta_2)^2 \, \mathrm{d}x.$$

La distance d'un point $\mathbf{x}$ du manche du balai de coordonnées (x_1, x_2, x_3) à l'axe Δ_2 est $d(\mathbf{x}, \Delta_1) = \sqrt{x_1^2 + (x_3 - H/2)^2}$. On a donc

$$J_{21} = \int_{\Omega_1} d(\mathbf{x}, \Delta_2)^2 \, \mathrm{d}x = \int_{\Omega_1} x_1^2 + (x_3 - \tfrac{H}{2})^2 \, \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}x_3$$

En utilisant un changement de variables cylindrique, on obtient

$$\begin{aligned}
J_{21} &= \int_0^H \int_0^{2\pi} \int_0^R (r^2(\cos(\theta))^2 + (x_3 - \tfrac{H}{2})^2) r \;\mathrm{d}r\mathrm{d}\theta\mathrm{d}x_3 \\
&= H \int_0^{2\pi} \int_0^R r^3 (\cos(\theta))^2 \;\mathrm{d}r\mathrm{d}\theta + 2\pi \int_0^H \int_0^R r (x_3 - \tfrac{H}{2})^2 \;\mathrm{d}r\mathrm{d}x_3 \\
&= H \left(\int_0^{2\pi} (\cos(\theta))^2 \;\mathrm{d}\theta \right) \left(\int_0^R r^3 \;\mathrm{d}r \right) + 2\pi \left(\int_0^H (x_3 - \tfrac{H}{2})^2 \;\mathrm{d}x_3 \right) \left(\int_0^R r \;\mathrm{d}r \right) \\
&= \frac{\pi H R^4}{4} + \frac{\pi H^3 R^2}{12}.
\end{aligned}$$

On calcule la contribution de la brosse du balai :

$$\begin{aligned}
J_{22} &= \int_{\Omega_2} d(\mathbf{x}, \Delta_2)^2 \, \mathrm{d}x = \int_{\Omega_2} x_1^2 + (\tfrac{1}{2}H - x_3)^2 \, \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}x_3 \\
&= \int_{-h}^0 \int_0^L \int_0^\ell x_1^2 \;\mathrm{d}x_1 \,\mathrm{d}x_2 \;\mathrm{d}x_3 + \int_{-h}^0 \int_0^L \int_0^\ell (\tfrac{1}{2}H - x_3)^2 \;\mathrm{d}x_1 \;\mathrm{d}x_2 \;\mathrm{d}x_3 \\
&= \frac{1}{3}\ell^3 h L + L\ell \Big[-\tfrac{1}{3}(\tfrac{1}{2}H - x_3)^3 \Big]_{-h}^0 = \frac{1}{3}\ell L \left(h\ell^2 + (h + \tfrac{1}{2}H)^3 - \frac{h^3}{8} \right)
\end{aligned}$$

On en déduit que le moment d'inertie du balai par rapport à l'axe Δ_2 vaut

$$J_{\Delta_2} = \frac{\pi H R^4 \rho}{4} + \frac{\pi H^3 R^2 \rho}{12} + \frac{1}{3}\ell L \left(h\ell^2 + (h + \tfrac{1}{2}H)^3 - \frac{h^3}{8} \right) \rho.$$

4 - En utilisant les valeurs numériques données dans l'énoncé, on obtient

$$J_{\Delta_1} \approx 0.0021314 \quad \text{et} \quad J_{\Delta_2} \approx 0.8558528.$$

Le moment d'inertie de la boule de rayon $\widetilde{R}$ et de densité 380 kg. m^{-3} par rapport à un axe du repère $(O; \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ est égal à J_{Δ_1} si

$$\frac{8\pi\rho\widetilde{R}^5}{15} = \frac{\pi}{2} R^4 H \rho + \frac{1}{3} L \ell h (\ell^2 + L^2) \rho,$$

autrement dit, si

$$\widetilde{R}^5 = \frac{15}{8\pi}\left(\frac{\pi}{2}R^4H + \frac{1}{3}L\ell h(\ell^2 + L^2)\right)$$

Numériquement, on obtient $\widetilde{R} \approx 8.03$ cm.

De même, le moment d'inertie de la boule de $\mathbb{R}^3$ de rayon $\widetilde{R}$ et de densité $380\,\mathrm{kg}\cdot\mathrm{m}^{-3}$ par rapport à un axe du repère $(O; \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ est égal à J_{Δ_2} si

$$\frac{8\pi\rho\widetilde{R}^5}{15} = \frac{\pi HR^4\rho}{4} + \frac{\pi H^3R^2\rho}{12} + \frac{1}{3}\ell L\left(h\ell^2 + (h + \tfrac{1}{2}H)^3 - \frac{h^3}{8}\right)\rho,$$

autrement dit, si

$$\widetilde{R}^5 = \frac{15}{8\pi}\left(\frac{\pi HR^4}{4} + \frac{\pi H^3R^2}{12} + \frac{1}{3}\ell L\left(h\ell^2 + (h + \tfrac{1}{2}H)^3 - \frac{h^3}{8}\right)\right).$$

Numériquement, on obtient $\widetilde{R} \approx 26.65$ cm.

CHAPITRE 13

Intégration sur des courbes et des surfaces

Savoir calculer la longueur d'une courbe, l'aire d'une surface ou le volume d'un objet a toujours été au cœur des préoccupations des mathématiciens : depuis les travaux d'Archimède sur le calcul des aires et des volumes cherchant une approximation du nombre π, jusqu'aux travaux actuels comme la résolution de la conjecture de Poincaré[1] qui valut la médaille Fields à Grigori Perelman[2] en 2003.

Nous avons depuis quelques chapitres mis en place tous les outils mathématiques permettant de définir correctement les notions d'intégrale le long d'une courbe et d'intégrale sur une surface qui sont le cœur de ce chapitre et qui, dans le cas où la fonction intégrée est la fonction constante égale à 1, fournissent le moyen de déterminer la longueur d'une courbe ou l'aire d'une surface. Dans la dernière partie de ce chapitre, nous montrerons brièvement que l'idée de calculer la longueur d'une courbe (ensemble à un paramètre) ou l'aire d'une surface (ensemble à deux paramètres) peut être définie de manière générale sur des ensembles à n paramètres. Nous évoquerons alors un résultat essentiel du calcul intégral : le *théorème de Stokes*, dont les corollaires les plus simples portent les noms de *formule de Green* et de *formule d'Ostrogradski*.

13.1 Intégrale curviligne

Nous considérons dans cette partie l'espace euclidien $\mathbb{R}^n$ ($n \in \mathbb{N}^*$) muni du produit scalaire canonique[3] défini par

$$\forall (\mathbf{x}, \mathbf{y}) \in \mathbb{R}^n \times \mathbb{R}^n \quad \mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^{n} x_i y_i,$$

[1] La conjecture de Poincaré pose la question de savoir si certaines propriétés caractérisent une sphère. Elle a été formulée par Henri Poincaré en 1904 et s'énonce précisément sous la forme suivante : est-ce qu'une variété compacte simplement connexe et sans bord de $\mathbb{R}^{n+1}$ est homéomorphe à la sphère de dimension n ?

[2] Voir le site `www.mathunion.org/General/Prizes/Fields/Prizewinners.html`.

[3] Voir la définition 8.15 p. 479.

où $(x_1, .., x_n)$ et $(y_1, ..., y_n)$ désignent les composantes de $\mathbf{x}$ et $\mathbf{y}$ dans la base canonique de $\mathbb{R}^n$. On note $\|\mathbf{x}\|$ la norme euclidienne d'un vecteur $\mathbf{x} \in \mathbb{R}^n$.

13.1.1 Intégrer le long d'une courbe

On se donne une application f d'un ouvert U de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^n$ continue (une telle application est qualifiée de *champ de vecteurs* sur U) et une courbe paramétrée[4] C de classe $\mathcal{C}^1$ de $\mathbb{R}^n$ dont un paramétrage est $\varphi \in \mathcal{C}^1([a,b], U)$. En chaque point $M = \varphi(t)$ de la courbe paramétrée C, on dispose du vecteur $f(\varphi(t)) \in \mathbb{R}^n$ correspondant à la valeur prise par l'application f en $M \in U$ et du vecteur $\varphi'(t) \in \mathbb{R}^n$ correspondant au vecteur tangent à la courbe paramétrée C au point M. On peut ainsi définir l'application

$$\xi : t \in [a,b] \longmapsto f(\varphi(t)) \cdot \varphi'(t) \in \mathbb{R}$$

qui à une valeur du paramètre t associe le produit scalaire des deux vecteurs $f(\varphi(t))$ et $\varphi'(t)$. D'après les hypothèses, l'application ξ de $[a,b]$ dans $\mathbb{R}$ est continue et par conséquent[5] Riemann-intégrable sur $[a,b]$. On appelle intégrale de f le long de la courbe C l'intégrale $\int_a^b \xi(t)\,\mathrm{d}t$, ce qui conduit à poser la définition suivante.

Définition 13.1 *Soient f une application continue sur un ouvert U de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^n$ et C une courbe paramétrée dont un paramétrage est $\varphi \in \mathcal{C}^1([a,b], U)$. On appelle* intégrale curviligne *de f le long de C et on note $\int_C f \cdot \mathrm{d}\mathbf{x}$ le réel défini par*

$$\int_C f \cdot \mathrm{d}\mathbf{x} = \int_a^b f(\varphi(t)) \cdot \varphi'(t)\,\mathrm{d}t.$$

La notion d'intégrale curviligne est utilisée par exemple pour définir le travail d'une force s'exerçant sur un objet qui se déplace le long d'une trajectoire dans l'espace. En physique, la quantité $\mathrm{d}\mathbf{x}$ peut être interprétée comme un déplacement infinitésimal et le travail d'une force f le long d'une courbe C est donné par

$$W = \int_C f \cdot \mathrm{d}\mathbf{x}\,.$$

- Ainsi, si la force f est parallèle au déplacement $\mathrm{d}\mathbf{x}$ et orientée dans le sens du déplacement alors le travail W fourni par la force est positif : du point de vue physique, la force a augmenté l'énergie cinétique du système, celui-ci se déplace donc plus rapidement. On appelle une telle force, une *force motrice.*

[4] Voir la définition 11.1 p. 714.

[5] Voir la proposition 18.2 p. 842 du *Cours de première année.*

- Si la force f est parallèle au déplacement $\mathbf{dx}$ mais orientée dans le sens opposé au déplacement alors le travail W fourni par la force est négatif : du point de vue physique, la force a diminué l'énergie cinétique du système, celui-ci se déplace donc plus lentement. On appelle une telle force, une *force résistante.*
- Si la force f est perpendiculaire au déplacement $\mathbf{dx}$, le travail W de la force est nul ; la force n'a pas modifié l'énergie cinétique du système.

Remarque L'intégrale curviligne $\int_C f \cdot \mathbf{dx}$ est dans certains contextes appelée la *circulation* du *champ de vecteurs* f le long de la courbe C. ◇

La notation $\int_C f \cdot \mathbf{dx}$ introduite dans la définition 13.1 peut sembler porter à confusion car elle ne fait pas apparaître explicitement la dépendance de cette quantité par rapport au paramétrage φ choisi pour la courbe. La proposition suivante permet de lever cette confusion.

PROPOSITION 13.1 *Soient f une application continue sur un ouvert U de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^n$ et C une courbe paramétrée de classe $\mathcal{C}^1$ contenue dans U. La valeur de l'intégrale curviligne $\int_C f \cdot \mathbf{dx}$ est invariante par changement de paramétrage $\mathcal{O}^1$-équivalent*[6] *de la courbe C.*

Démonstration Soient $\varphi \in \mathcal{C}^1([a,b],U)$ et $\psi \in \mathcal{C}^1([c,d],U)$ deux paramétrages $\mathcal{O}^1$-équivalents de la courbe paramétrée C. D'après la définition 11.8, il existe une application θ de $[a,b]$ vers $[c,d]$, de classe $\mathcal{C}^1$, strictement croissante et qui vérifie $\varphi = \psi \circ \theta$. Pour tout $t \in [a,b]$, on obtient en utilisant la formule de dérivation de la composée de fonctions : $\varphi'(t) = \psi'(\theta(t))\,\theta'(t)$. En utilisant la formule du changement de variable pour une intégrale[7], on obtient :

$$\int_c^d f(\psi(s)) \cdot \psi'(s)\ \mathrm{d}s = \int_{\theta(a)}^{\theta(b)} f(\psi(s)) \cdot \psi'(s)\ \mathrm{d}s$$
$$= \int_a^b f(\psi(\theta(t))) \cdot \psi'(\theta(t))\ \theta'(t)\ \mathrm{d}t = \int_a^b f(\varphi(t)) \cdot \varphi'(t)\ \mathrm{d}t.$$

Autrement dit, la valeur de l'intégrale curviligne $\int_C f \cdot \mathbf{dx}$ est la même que l'on choisisse le paramétrage φ ou le paramétrage ψ. □

Remarques

1. Si on dispose d'un paramétrage $\mathcal{C}^1$-équivalent mais non $\mathcal{O}^1$-équivalent alors[8] la valeur de l'intégrale curviligne change de signe. Géométriquement, cela signifie que si la courbe est parcourue dans le sens inverse, alors le signe de l'intégrale curviligne est modifié.

---

[6] Voir la définition 11.8 p. 739.

[7] Voir le *Cours de première année* p. 863.

[8] Il suffit de refaire la démonstration précédente en utilisant un changement de variable θ strictement décroissant.

2. Pour l'instant $\mathrm{d}\mathbf{x}$ est à prendre comme une notation ; on verra par la suite quel sens mathématique on peut lui donner. ◇

Exemples

1. Soient $f : (x_1, x_2) \in \mathbb{R}^2 \mapsto (x_1 x_2, x_1 + x_2)$ et C le segment de droite joignant les points $A = (0,1)$ et $B = (1,1)$, orienté de A vers B. Un paramétrage de C est[(9)]

$$\varphi : t \in [0,1] \longmapsto (t, 1) \in \mathbb{R}^2.$$

Pour $t \in [0,1]$ on a $\varphi'(t) = (1,0)$ et $f(\varphi(t)) = (t, 1+t)$ donc $f(\varphi(t)) \cdot \varphi'(t) = t$ et on en déduit que

$$\int_C f \cdot \mathrm{d}\mathbf{x} = \int_0^1 t \, \mathrm{d}t = \frac{1}{2}.$$

2. Considérons à nouveau l'application $f : (x_1, x_2) \in \mathbb{R}^2 \mapsto (x_1 x_2, x_1 + x_2)$ et calculons l'intégrale curviligne de f le long du segment de droite C' joignant les points $B = (1,1)$ et $A = (0,1)$, orienté de B vers A. Un paramétrage de C' est

$$\varphi : t \in [0,1] \longmapsto (1 - t, 1) \in \mathbb{R}^2.$$

Pour tout $t \in [0,1]$ on a $\varphi'(t) = (-1,0)$ et $f(\varphi(t)) = (1-t, 2-t)$ et par conséquent $f(\varphi(t)) \cdot \varphi'(t) = t - 1$. On en déduit que

$$\int_{C'} f \cdot \mathrm{d}\mathbf{x} = \int_0^1 (t-1) \, \mathrm{d}t = -\frac{1}{2}.$$

On illustre par ce calcul la remarque indiquant que si on dispose d'un paramétrage $\mathcal{C}^1$-équivalent mais non $\mathcal{O}^1$-équivalent, c'est-à-dire si l'on change l'orientation de la courbe, alors la valeur de l'intégrale change de signe.

3. Dans l'exemple précédent, le segment de droite C' joignant les points $B = (1,1)$ et $A = (0,1)$ admet également pour paramétrage

$$\varphi : t \in [-1,0] \longmapsto (-t, 1) \in \mathbb{R}^2.$$

Pour tout $t \in [-1,0]$ on a $\varphi'(t) = (-1,0)$ et $f(\varphi(t)) = (-t, 1-t)$ de sorte que $f(\varphi(t)) \cdot \varphi'(t) = t$ et

$$\int_{C'} f \cdot \mathrm{d}\mathbf{x} = \int_{-1}^0 t \, \mathrm{d}t = -\frac{1}{2}.$$

On illustre avec cet exemple la proposition 13.1 : la valeur de l'intégrale curviligne est invariante par changement de paramétrage.

4. Soient $f : (x_1, x_2) \in \mathbb{R}^2 \mapsto (x_1 x_2, x_1 + x_2)$ et C le demi-cercle joignant les points $A = (0,1)$ et $B = (1,1)$, orienté de A vers B dans le sens trigonométrique. Un paramétrage de C est[(9)]

$$\varphi : t \in [\pi, 2\pi] \longmapsto \left(\tfrac{1}{2} + \tfrac{1}{2}\cos(t), 1 + \tfrac{1}{2}\sin(t)\right) \in \mathbb{R}^2.$$

[(9)] Voir p. 724.

Pour tout $t \in [\pi, 2\pi]$ on a

$$\varphi'(t) = \left(-\tfrac{1}{2}\sin(t), \tfrac{1}{2}\cos(t)\right)$$
$$f(\varphi(t)) = \left(\tfrac{1}{4}(1+\cos(t))(2+\sin(t)), \tfrac{3}{2} + \tfrac{1}{2}(\cos(t)+\sin(t))\right)$$

et par conséquent $f(\varphi(t))\cdot\varphi'(t) = \frac{1}{8}\left(6\cos(t) - 2\sin(t) - \sin(t)^2\cos(t) - \sin(t)^2\right)$. En utilisant les résultats suivants

$$\int_\pi^{2\pi} \cos(t)\ \mathrm{d}t = 0, \qquad \int_\pi^{2\pi} \sin(t)\ \mathrm{d}t = -2,$$
$$\int_\pi^{2\pi} \sin(t)^2\cos(t)\ \mathrm{d}t = \left[\frac{1}{3}\sin(t)^3\right]_\pi^{2\pi} = 0 \quad \text{et}$$
$$\int_\pi^{2\pi} \sin(t)^2\ \mathrm{d}t = \int_\pi^{2\pi} \frac{1-\cos(2t)}{2}\ \mathrm{d}t = \frac{\pi}{2},$$

on en déduit que

$$\int_C f\cdot \mathrm{d}\mathbf{x} = \frac{1}{8}(4 - \frac{\pi}{2}) = \frac{1}{2} - \frac{\pi}{16}.$$

De par sa définition, l'intégrale curviligne hérite des propriétés de l'intégrale de Riemann. En particulier l'intégrale curviligne possède la propriété de linéarité suivante : pour tout $(\lambda, \mu) \in \mathbb{R}^2$ et tout $(f, g) \in \mathcal{C}(U, \mathbb{R}^n)^2$ on a

$$\int_C (\lambda f + \mu g)\cdot \mathrm{d}\mathbf{x} = \lambda \int_C f\cdot \mathrm{d}\mathbf{x} + \mu \int_C g\cdot \mathrm{d}\mathbf{x}\,.$$

La relation de Chasles[10] permet d'étendre la définition 13.1 à des courbes qui ne sont pas de classe $\mathcal{C}^1$. Lorsque l'application φ de $[a, b]$ dans U est continue sur $[a, b]$ et de classe $\mathcal{C}^1$ sur $[a, c]$ et sur $[c, b]$ avec $a < c < b$ sans être pour autant de classe $\mathcal{C}^1$ sur $[a, b]$, alors on peut définir deux sous-arcs de C de classe $\mathcal{C}^1$, le premier C_1 défini par la restriction de φ à l'intervalle $[a, c[$ et le second C_2 défini par la restriction de φ à l'intervalle $[c, b]$. La relation de Chasles appliquée à la définition de l'intégrale curviligne, voir p. 938, indique que

$$\int_C f\cdot \mathrm{d}\mathbf{x} = \int_{C_1} f\cdot \mathrm{d}\mathbf{x} + \int_{C_2} f\cdot \mathrm{d}\mathbf{x}.$$

Plus généralement, considérons une application φ continue sur $[a, b]$, à valeurs dans U et de classe $\mathcal{C}^1$ par morceaux, c'est-à-dire telle qu'il existe une subdivision $(c_k)_{k\in\{1,\dots,N\}}$ de l'intervalle $[a, b]$ telle que pour tout entier $k \in \{1, \dots, N-1\}$ l'application φ est de classe $\mathcal{C}^1$ sur $[c_k, c_{k+1}]$. Désignons par C_k la courbe de classe $\mathcal{C}^1$ de $\mathbb{R}^n$ dont un paramétrage est la restriction de l'application φ à l'intervalle $[c_k, c_{k+1}]$. L'intégrale curviligne le long de la courbe C d'une application f continue de U dans $\mathbb{R}^n$ est donnée par

$$\int_C f\cdot \mathrm{d}\mathbf{x} = \sum_{k=1}^{N-1} \int_{C_k} f\cdot \mathrm{d}\mathbf{x}.$$

[10] Voir le *Cours de première année* p. 844.

Exemple Calculons l'intégrale curviligne de $f : (x_1, x_2) \in \mathbb{R}^2 \mapsto (x_1 x_2, x_1 + x_2)$ le long du carré C de sommet O, $A = (0,1)$, $B = (1,1)$ et $D = (1,0)$ parcouru dans le sens $OABD$. Il n'est pas possible d'obtenir un paramétrage du carré qui soit de classe $\mathcal{C}^1$; on utilise donc la relation de Chasles :

$$\int_C f \cdot \mathrm{d}\mathbf{x} = \int_{C_1} f \cdot \mathrm{d}\mathbf{x} + \int_{C_2} f \cdot \mathrm{d}\mathbf{x} + \int_{C_3} f \cdot \mathrm{d}\mathbf{x} + \int_{C_4} f \cdot \mathrm{d}\mathbf{x}$$

où

- la courbe C_1 est le segment $[OA]$ dont un paramétrage est

$$\varphi_1 : t \in [0,1] \longmapsto (0,t) \in \mathbb{R}^2,$$

- la courbe C_2 est le segment $[AB]$; on a déjà établi dans les exemples précédents que

$$\int_{C_2} f \cdot \mathrm{d}\mathbf{x} = \frac{1}{2},$$

- la courbe C_3 est le segment $[BD]$ dont un paramétrage est

$$\varphi_3 : t \in [0,1] \longmapsto (1, 1-t) \in \mathbb{R}^2,$$

- la courbe C_4 est le segment $[DO]$ dont un paramétrage est

$$\varphi_4 : t \in [0,1] \longmapsto (1-t, 0) \in \mathbb{R}^2.$$

Pour tout $t \in [-1,0]$ on a $\varphi_1'(t) = (0,1)$ et $f(\varphi_1(t)) = (0,t)$ de sorte que $f(\varphi_1(t)) \cdot \varphi_1'(t) = t$. De même, $\varphi_3'(t) = (0,-1)$ et $f(\varphi_3(t)) = (1-t, 2-t)$ de sorte que $f(\varphi_3(t)) \cdot \varphi_3'(t) = t - 2$, et $\varphi_4'(t) = (-1,0)$ et $f(\varphi_4(t)) = (0, 1-t)$ de sorte que $f(\varphi_4(t)) \cdot \varphi_4'(t) = 0$. On en déduit que

$$\int_C f \cdot \mathrm{d}\mathbf{x} = \int_0^1 t \, \mathrm{d}t + \frac{1}{2} + \int_0^1 (t-2) \, \mathrm{d}t + \int_0^1 0 \, \mathrm{d}t = -\frac{1}{2}.$$

Calcul pratique d'une intégrale curviligne

En pratique, pour calculer l'intégrale curviligne $\int_C f \cdot \mathrm{d}\mathbf{x}$ on procède de la manière suivante :

1. Dans un premier temps, on détermine un paramétrage $\varphi \in \mathcal{C}^1([a,b], \mathbb{R}^n)$ de la courbe C. Dans le cas où seul existe un paramétrage de classe $\mathcal{C}^1$ par morceaux, on utilise la relation de Chasles sur chaque morceau.
2. Ensuite, on calcule le vecteur tangent à la courbe pour ce paramétrage, c'est-à-dire le vecteur $\varphi'(t)$ pour tout $t \in [a,b]$.
3. Enfin, on utilise la définition de l'intégrale curviligne pour se ramener au calcul de l'intégrale de Riemann d'une fonction réelle d'une variable réelle.
4. On calcule cette intégrale en utilisant les techniques de calcul de l'intégrale de Riemann, voir le *Cours de première année* chap. 18.

Exercice 13.1
1 - Calculer l'intégrale curviligne le long du carré de sommets $C = (1,1)$, $A = (-1,1)$, $S = (-1,-1)$, et $I = (1,-1)$ parcouru dans le sens trigonométrique de l'application $f : (x_1, x_2) \in \mathbb{R}^2 \mapsto (x_1 + x_2, x_1 x_2)$.
2 - Calculer l'intégrale curviligne de $f : (x_1, x_2) \in \mathbb{R}^2 \mapsto (-x_2, x_1) \in \mathbb{R}^2$ le long du cercle de centre 0 et de rayon $R \in \mathbb{R}_+^*$ parcouru dans le sens trigonométrique.

DÉFINITION 13.2 (Extrémités d'une courbe paramétrée)
Soient I un intervalle d'extrémités a et b ($(a,b) \in \overline{\mathbb{R}} \times \overline{\mathbb{R}}$, $a < b$) et C une courbe paramétrée de $\mathbb{R}^n$ dont un paramétrage est $\varphi \in \mathcal{C}(I, \mathbb{R}^n)$.

✗ *Si φ a une limite à droite $A \in \mathbb{R}^n$ en a et une limite à gauche $B \in \mathbb{R}^n$ en b alors A et B constituent les* extrémités de la courbe C.

✗ *Lorsque les deux extrémités d'une courbe C sont égales, on dit que la courbe C est* fermée.

Exemples

1. L'image de l'arc paramétré $\varphi : t \in [-\pi/2, \pi/2[\mapsto (\cos(t), \sin(t))$ est une courbe paramétrée qui a pour extrémités $A = (0,-1)$ et $B = (0,1)$. On notera que l'extrémité de la courbe n'appartient pas nécessairement à la courbe.

2. L'image de l'arc paramétré $\varphi : t \in]0, 2\pi[\mapsto (\cos(t), \sin(t))$ est une courbe paramétrée fermée.

Travail d'une force conservative

En physique, on qualifie de forces conservatives, des forces dont le travail le long d'une courbe quelconque ne dépend pas du chemin suivi mais uniquement des extrémités de la courbe. Nous allons montrer que le poids est un exemple de force conservative.

Considérons un corps de masse m se déplaçant dans l'espace d'un point A vers un point B dans un référentiel galiléen $(O; \mathbf{i}, \mathbf{j}, \mathbf{k})$, l'axe $(O; \mathbf{k})$ étant supposé vertical et dirigé dans le sens opposé à la gravité $\mathbf{g}$. Si la trajectoire suivie par le corps entre les points A et B est donnée par une courbe paramétrée C de classe $\mathcal{C}^1$ alors le travail du poids $\mathbf{P} = m\,\mathbf{g} = -mg\,\mathbf{k}$ vaut

$$W = \int_C \mathbf{P} \cdot \mathrm{d}\mathbf{x} \,.$$

Soit $\varphi : t \in [0,1] \mapsto (x(t), y(t), z(t)) \in \mathbb{R}^3$ un paramétrage de la courbe paramétrée C tel que $\varphi(0) = A$ et $\varphi(1) = B$. On a

$$W = \int_0^1 P(\varphi(t)) \cdot \varphi'(t)\ \mathrm{d}t = \int_0^1 mgz'(t)\ \mathrm{d}t = mg(z(1) - z(0)).$$

On remarque que le travail du poids d'un corps est indépendant du chemin suivi lors de son déplacement ; il ne dépend que de la variation d'altitude de ce corps. Par exemple, pour monter en haut du Mont Blanc (4808 m) en partant de la ville de Chamonix (1042 m), un alpiniste de masse 80 kg doit effectuer un travail correspondant à celui de son poids sur une hauteur de 3766 m, soit un travail W égal au produit mgh où la masse m vaut 80 kg, g est l'accélération due à la gravité ($9,81\,\mathrm{m}\cdot\mathrm{s}^{-2}$) et h vaut 3766 m. Le travail de son poids, correspondant à l'énergie mécanique dépensée, vaut par conséquent environ 2955 kJ (quel que soit l'itinéraire emprunté).

Intégrale d'un gradient le long d'une courbe

Intéressons-nous au calcul de l'intégrale curviligne le long d'une courbe paramétrée C de classe $\mathcal{C}^1$ d'une fonction f de $\mathbb{R}^n$ dans $\mathbb{R}^n$ qui s'exprime comme le gradient d'une fonction h de $\mathbb{R}^n$ dans $\mathbb{R}$ de classe $\mathcal{C}^1$, *i.e.* $f = \nabla h$. En considérant un paramétrage $\varphi \in \mathcal{C}^1([a,b],\mathbb{R}^n)$ de la courbe paramétrée C, on obtient

$$\int_C f\cdot \mathrm{d}\mathbf{x} = \int_C \nabla h\cdot \mathrm{d}\mathbf{x} = \int_a^b \nabla h(\varphi(t))\cdot\varphi'(t)\ \mathrm{d}t.$$

Désignons par $\varphi_1, \ldots, \varphi_n$ les composantes de la fonction φ. On a

$$\nabla h(\varphi(t))\cdot\varphi'(t) = \begin{pmatrix}\partial_1 h(\varphi(t))\\ \vdots \\ \partial_n h(\varphi(t))\end{pmatrix}\cdot\begin{pmatrix}\varphi_1'(t)\\ \vdots\\ \varphi_n'(t)\end{pmatrix} = \sum_{i=1}^{n}\partial_i h(\varphi_1(t),\ldots,\varphi_n(t))\ \varphi_i'(t).$$

Cette expression s'interprète d'après la formule de dérivation des fonctions composées[11] comme la dérivée de $h\circ\varphi$ évaluée en t. On en déduit que

$$\int_C f\cdot\mathrm{d}\mathbf{x} = \int_a^b (h\circ\varphi)'(t)\ \mathrm{d}t = h(\varphi(b)) - h(\varphi(a)).$$

On s'aperçoit que la valeur de l'intégrale curviligne de f le long de C ne dépend que des valeurs prises par h en $\varphi(a)$ et $\varphi(b)$. On peut énoncer le résultat suivant.

Proposition 13.2 *Soient h une application de classe $\mathcal{C}^1$ sur un ouvert U de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}$ et C une courbe paramétrée de U, de classe $\mathcal{C}^1$ et d'extrémités A et B. On a*

$$\int_C \nabla h\cdot\mathrm{d}\mathbf{x} = h(B) - h(A).$$

En particulier, si la courbe paramétrée C est fermée alors $\displaystyle\int_C \nabla h\cdot\mathrm{d}\mathbf{x} = 0.$

[11] Voir la proposition 10.6 p. 619.

En physique, lorsqu'une force f est le gradient d'une application h, on dit que cette force dérive d'un potentiel (h est appelé le potentiel). La proposition 13.2 indique que, dans ce contexte, le travail d'une force qui dérive d'un potentiel ne dépend que de la position initiale et de la position finale. Par exemple, la force correspondant au poids $\mathbf{P} = m\mathbf{g}$ est une force qui dérive du potentiel[(12)] $h : (x_1, x_2, x_3) \in \mathbb{R}^3 \mapsto (0, 0, -mgx_3)$. On dit que la force gravitationnelle dérive du potentiel de gravitation. De la même manière, en électricité, la force électrique dérive du potentiel électrique. Dans un gaz, les interactions entre les particules sont des interactions à courte portée (forces de Van Der Waals) qui dérivent elles aussi d'un potentiel.

13.1.2 Longueur d'une courbe

Un segment $[P_0P_1]$ de l'espace vectoriel normé $(\mathbb{R}^n, \|\cdot\|)$, où $\|\cdot\|$ désigne la norme euclidienne, correspond à l'ensemble

$$\left\{M \in \mathbb{R}^n \ ; \ \exists \lambda \in [0,1] \quad \overrightarrow{OM} = \lambda \overrightarrow{OP_0} + (1-\lambda)\overrightarrow{OP_1}\right\}$$

et a pour longueur $\mathrm{Long}([P_0P_1]) = \|\overrightarrow{P_0P_1}\|$. Une courbe formée par la réunion de plusieurs segments successifs $[P_0P_1]$, $[P_1P_2]$, $\ldots$, $[P_{N-1}P_N]$ ($N \in \mathbb{N}^*$) est appelée *ligne polygonale* et a pour longueur la somme des longueurs des segments qui la composent. Intuitivement, la longueur d'une courbe peut être introduite à partir de la longueur d'un segment de la façon suivante.

Considérons une courbe C paramétrée par une application φ de $[a, b]$ dans $\mathbb{R}^n$ de classe $\mathcal{C}^1$. Pour $N \in \mathbb{N}^*$, prenons une subdivision de l'intervalle $[a, b]$ en N parties :

$$a = t_0 < t_1 < \cdots < t_{N-1} < t_N = b.$$

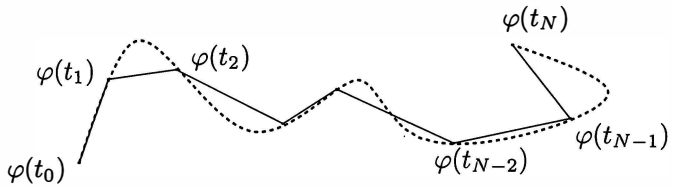


Fig. 1 Ligne polygonale inscrite dans une courbe.

Notons C_N la ligne polygonale passant par tous les points $P_i = \varphi(t_i)$, $i \in \{0, \ldots, N\}$ de la courbe C et notons $L_N = \mathrm{Long}(C_N)$ la longueur de cette ligne polygonale. Une telle ligne polygonale est appelée une ligne polygonale inscrite dans la courbe C. Montrons que si la courbe C est de classe $\mathcal{C}^1$ alors la borne supérieure de l'ensemble des longueurs des lignes polygonales inscrites dans la courbe C existe ; on peut alors définir la longueur de C comme étant cette borne supérieure. Intuitivement ce processus consiste à « approcher » la courbe par des lignes polygonale en prenant de plus en plus de points de passage, voir la figure 1.

(12) On pourra vérifier que le gradient de cette application h est bien égal à $m\mathbf{g} = (0, 0, -mg)$.

Les n fonctions coordonnées $\varphi_1, \ldots, \varphi_n$ de l'application φ sont des applications de classe $\mathcal{C}^1$ de $[a, b]$ dans $\mathbb{R}$ et pour tout $i \in \{1, \ldots, N\}$, on a

$$\varphi(t_i) - \varphi(t_{i-1}) = \int_{t_{i-1}}^{t_i} \varphi'(t)\ \mathrm{d}t,$$

où la quantité $\int_{t_{i-1}}^{t_i} \varphi'(t)\ \mathrm{d}t$ désigne le vecteur de $\mathbb{R}^n$ de coordonnées $(\int_{t_{i-1}}^{t_i} \varphi_1'(t)\ \mathrm{d}t, \ldots, \int_{t_{i-1}}^{t_i} \varphi_n'(t)\ \mathrm{d}t)$. En considérant la norme euclidienne sur $\mathbb{R}^n$ et en appliquant une version généralisée de l'inégalité triangulaire [13], on en déduit que

$$\|\varphi(t_i) - \varphi(t_{i-1})\| = \left\| \int_{t_{i-1}}^{t_i} \varphi'(t)\ \mathrm{d}t \right\| \leqslant \int_{t_{i-1}}^{t_i} \|\varphi'(t)\|\ \mathrm{d}t.$$

Le terme de gauche de cette inégalité correspond à la longueur du segment $[P_{i-1}P_i]$. En effectuant la somme sur tous les entiers $i \in \{1, \ldots, N\}$, on obtient :

$$\mathrm{Long}(C_N) = \sum_{i=1}^{N} \mathrm{Long}([P_{i-1}P_i]) \leqslant \sum_{i=1}^{N} \int_{t_{i-1}}^{t_i} \|\varphi'(t)\|\ \mathrm{d}t \leqslant \int_a^b \|\varphi'(t)\|\ \mathrm{d}t. \quad (1)$$

On en déduit que l'ensemble des longueurs des lignes polygonales inscrites dans la courbe C est un ensemble borné dont un majorant est $\int_a^b \|\varphi'(t)\|\ \mathrm{d}t$. Ceci permet de justifier que la borne supérieure des longueurs de lignes polygonales inscrites dans la courbe C existe [14]. La définition suivante a donc un sens.

Définition 13.3 *La longueur d'une courbe paramétrée C de classe $\mathcal{C}^1$ est la borne supérieure des longueurs des lignes polygonales inscrites dans C.*

Cette définition, bien qu'intuitive, n'est pas pratique puisque pour calculer la longueur d'une courbe paramétrée, il faut déterminer la borne supérieure de l'ensemble des longueurs de toutes les lignes polygonales qui sont inscrites dans la courbe. La proposition suivante indique que la longueur d'une courbe paramétrée peut être exprimée par une intégrale curviligne, ce qui rend généralement le calcul plus aisé.

---

[13] L'inégalité triangulaire (voir la définition 7.3 p. 358) peut se généraliser ainsi : pour tout $\mathbf{u} \in \mathcal{C}([a, b], \mathbb{R}^n)$ on a $\| \int_a^b \mathbf{u}(t)\ \mathrm{d}t\| \leqslant \int_a^b \|\mathbf{u}(t)\|\ \mathrm{d}t$ (où $a < b$). Ce résultat se démontre en utilisant les sommes de Riemann (voir le *Cours de première année*, p. 868) et la continuité de l'application $\mathbf{x} \in \mathbb{R}^n \mapsto \|\mathbf{x}\|$ de la manière suivante :

$$\left\| \int_a^b \mathbf{u}(t)\ \mathrm{d}t \right\| = \left\| \lim_{k \to +\infty} \frac{b-a}{k} \sum_{j=0}^{k} \mathbf{u}\Big(a + j\frac{b-a}{k}\Big) \right\| = \lim_{k \to +\infty} \frac{b-a}{k} \left\| \sum_{j=0}^{k} \mathbf{u}\Big(a + j\frac{b-a}{k}\Big) \right\|$$

$$\leqslant \lim_{k \to +\infty} \frac{b-a}{k} \sum_{j=0}^{k} \left\| \mathbf{u}\Big(a + j\frac{b-a}{k}\Big) \right\| = \int_a^b \|\mathbf{u}(t)\|\ \mathrm{d}t.$$

[14] Rappelons une propriété fondamentale de l'ensemble des réels : toute partie majorée non vide de $\mathbb{R}$ admet une borne supérieure, voir le *Cours de première année* p. 95.

Proposition 13.3 *La longueur d'une courbe paramétrée régulière est l'intégrale curviligne le long de cette courbe de son vecteur tangent unitaire. En d'autres termes, si $\varphi \in \mathcal{C}^1([a,b],\mathbb{R}^n)$ désigne un paramétrage d'une courbe paramétrée régulière*[15] *C alors la longueur de C est donnée par*

$$\mathrm{Long}(C) = \int_a^b \|\varphi'(t)\| \, \mathrm{d}t$$

où $\|\cdot\|$ désigne la norme euclidienne.

Démonstration Un vecteur tangent à la courbe paramétrée C en $\varphi(t)$ où $t \in [a,b]$, est donné par $\varphi'(t)$. Pour obtenir un vecteur unitaire, il suffit de diviser ce vecteur par sa norme[16]. On en déduit que l'intégrale curviligne le long de la courbe C de son vecteur tangent unitaire, que nous noterons provisoirement $L(C)$, vaut

$$L(C) = \int_a^b \frac{\varphi'(t)}{\|\varphi'(t)\|} \cdot \varphi'(t) \, \mathrm{d}t = \int_a^b \|\varphi'(t)\| \, \mathrm{d}t.$$

Démontrer la proposition 13.3 revient donc à montrer que $L(C) = \mathrm{Long}(C)$.

Pour tout $t \in [a,b]$, notons C_t la courbe paramétrée de $\mathbb{R}^n$ dont un paramétrage est la restriction de l'application φ à l'intervalle $[a,t]$ et considérons les deux applications $f : t \in [a,b] \mapsto \mathrm{Long}(C_t)$ et $g : t \in [a,b] \mapsto L(C_t)$. Avec ces notations, montrer que $L(C)$ est égal à $\mathrm{Long}(C)$ revient à montrer que $f(b) = g(b)$. Pour établir cette relation, nous allons prouver que les deux applications f et g ont la même dérivée ; puisque les deux applications f et g sont nulles en a et continues sur $[a,b]$, on en déduira qu'elles sont égales sur $[a,b]$ donc en particulier en b.

L'application g qui est donnée en $t \in [a,b]$ par $g(t) = \int_a^t \|\varphi'(s)\| \, \mathrm{d}s$ est l'intégrale indéfinie[17] associée à $t \mapsto \|\varphi'(t)\|$. Cela implique que g est dérivable sur $[a,b]$ et que pour tout $t \in [a,b]$ on a $g'(t) = \|\varphi'(t)\|$. Pour prouver que l'application f est dérivable sur $]a,b[$, il faut montrer que pour tout $t \in]a,b[$ le taux d'accroissement $(f(t+h) - f(t))/h$ admet une limite lorsque h tend vers zéro. Tout d'abord, puisque $f(t+h)$ est par définition la borne supérieure des longueurs des lignes polygonales inscrites dans C_{t+h}, on a[18]

$$f(t+h) \geqslant f(t) + \|M_t M_{t+h}\|, \tag{2}$$

[15] Voir la définition 11.11 p. 745.

[16] Puisque la courbe est par hypothèse régulière, le vecteur $\varphi'(\mathbf{t})$ est non nul, voir la définition 11.11 p. 745.

[17] Voir le *Cours de première année* p. 847.

[18] Le calcul présenté ici n'est valide que lorsque $h > 0$, mais lorsque $h < 0$ un raisonnement analogue s'applique. Pour ne pas compliquer inutilement la démonstration, nous omettons de traiter ce second cas. On obtiendrait en toute rigueur que f est dérivable à gauche et à droite en $t \in]a,b[$ avec des dérivées à gauche et à droite égales, ce qui permettrait de conclure à la dérivabilité de f en t.

où M_t et M_{t+h} sont les points de C paramétrés par t et $t+h$, voir la figure 2. Par ailleurs, si on s'intéresse à la courbe paramétrée C'_h dont un paramétrage est la restriction de l'application φ à l'intervalle $[t, t+h]$ alors, en utilisant l'inégalité (1) de la p. 946, on obtient

$$f(t+h) - f(t) \leqslant \int_t^{t+h} \|\varphi'(s)\| \, \mathrm{d}s. \tag{3}$$

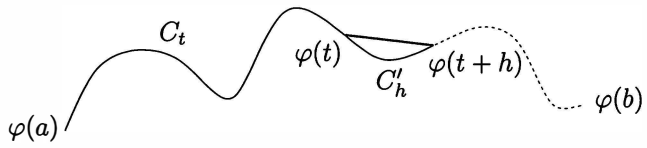


Fig. 2 Illustration de la construction.

Des inégalités (2) et (3), on déduit l'encadrement suivant :

$$\frac{\|\varphi(t+h) - \varphi(t)\|}{h} \leqslant \frac{f(t+h) - f(t)}{h} \leqslant \frac{1}{h} \int_t^{t+h} \|\varphi'(s)\| \, \mathrm{d}s.$$

Le membre de gauche de l'inégalité a pour limite $\|\varphi'\|$ car par définition de la dérivée[19] on a $\lim_{h\to 0} \frac{1}{h} (\varphi(t+h) - \varphi(t)) = \varphi'(t)$ et la norme est une application continue[20]. Le membre de droite de l'inégalité se réécrit :

$$\begin{aligned} \frac{1}{h} \int_t^{t+h} \|\varphi'(s)\| \, \mathrm{d}s &= \frac{1}{h} \left(\int_t^{t+h} \|\varphi'(s)\| \, \mathrm{d}s - \int_t^{t} \|\varphi'(s)\| \, \mathrm{d}s \right) \\ &= \frac{1}{h} \left(\Phi(t+h) - \Phi(t) \right) \end{aligned}$$

où Φ désigne l'intégrale indéfinie $x \mapsto \int_t^x \|\varphi'(s)\| \, \mathrm{d}s$. Lorsque h tend vers 0, le membre de droite de l'inégalité a donc pour limite la dérivée de Φ en t, c'est-à-dire $\|\varphi'(t)\|$. Comme le membre de droite et celui de gauche ont tous les deux pour limite $\|\varphi'(t)\|$ quand h tend vers 0, on en déduit que

$$\lim_{h\to 0} \frac{f(t+h) - f(t)}{h} = \|\varphi'(t)\|.$$

On en conclut que f est dérivable sur $[a,b]$ et que pour tout $t \in [a,b]$ sa dérivée en t vaut $\|\varphi'(t)\|$. Finalement, les deux applications f et g ont la même dérivée sur $[a,b]$. De plus, ces deux applications étant nulles en a, comme elles sont continues sur $[a,b]$, elles sont égales sur $[a,b]$ donc en particulier en b. On a $f(b) = g(b)$ ce qui conclut la preuve. □

[19] Voir la définition 16.1 p. 713 du *Cours de première année*.

[20] Voir la proposition 9.22 p. 572.

Exemples

1. Calculons la longueur du cercle C de centre $(0,0)$ et de rayon $R \in \mathbb{R}_+^*$ en considérant le paramétrage

$$\varphi : t \in [0, 2\pi] \longmapsto \big(R\cos(t), R\sin(t)\big) \in \mathbb{R}^2.$$

Pour tout $t \in [0, 2\pi]$, le vecteur tangent $\varphi'(t)$ est donné par $\big(-R\sin(t), R\cos(t)\big)$ et sa norme euclidienne vaut $\|\varphi'(t)\| = R$. On en déduit que :

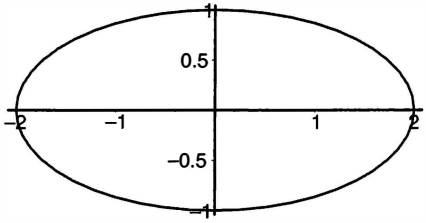
$$\text{Long(Cercle)} = \int_0^{2\pi} R \, \mathrm{d}t = 2\pi R.$$

2. L'arc paramétré $\varphi : t \in [0, 2\pi] \mapsto (R_1\cos(t), R_2\sin(t)) \in \mathbb{R}^2$ a pour image (lorsque R_1 et R_2 sont deux réels strictement positifs) une ellipse du plan de grand axe et de petit axe correspondant aux axes de coordonnées. C'est une courbe régulière dont la longueur vaut

$$\text{Long(Ellipse)} = \int_0^{2\pi} \sqrt{R_1^2\sin(t)^2 + R_2^2\cos(t)^2} \, \mathrm{d}t.$$

Ces intégrales ne peuvent pas s'exprimer à l'aide des fonctions usuelles. Elles sont appelées *intégrales elliptiques de seconde espèce.*

```
> R1:=2: R2:=1: phi:= t-> [R1*cos(t),R2*sin(t)]:
> plot([phi(t)[1],phi(t)[2],t=0..2*Pi],scaling=CONSTRAINED);
```

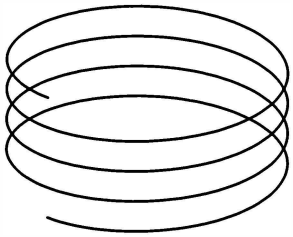



```
> dphi:=unapply(diff(phi(t),t),t):
> int(sqrt((dphi(t)[1])^2+(dphi(t)[2])^2),t=0..2*Pi);
                        8 EllipticE(1/2 sqrt(3))
> evalf(%);
                              9.688448216
```

3. L'hélice dont un paramétrage est $\varphi : t \in [0, \alpha] \mapsto (\cos(t), \sin(t), t) \in \mathbb{R}^3$ a pour vecteur tangent en $t \in [0, \alpha]$ le vecteur $\varphi'(t) = (-\sin(t), \cos(t), 1)$ qui a pour norme euclidienne $\|\varphi'(t)\| = \sqrt{2}$. On en déduit que

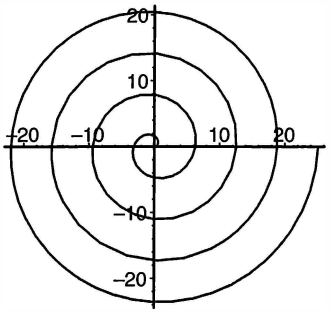
$$\text{Long(Hélice)} = \int_0^{\alpha} \sqrt{2} \, \mathrm{d}t = \sqrt{2}\,\alpha.$$

```
> with(plots): alpha:=8*Pi;
> phi:= t-> [cos(t),sin(t),t]:
> spacecurve([phi(t)[1],phi(t)[2],phi(t)[3],t=0..alpha]);
```



4. La spirale dont un paramétrage est $\varphi : t \in [0, \beta] \mapsto (t\cos(t), t\sin(t)) \in \mathbb{R}^2$ a pour vecteur tangent en $t \in [0, \beta]$ le vecteur $\varphi'(t) = (\cos(t) - t\sin(t), \sin(t) + t\cos(t))$. Il a pour norme euclidienne $\|\varphi'(t)\| = \sqrt{1+t^2}$.

```
> beta:=8*Pi:
> phi:= t-> [t*cos(t),t*sin(t)]:
> plot([phi(t)[1],phi(t)[2],t=0..beta],scaling=CONSTRAINED);
```



La longueur de la spirale est Long(Spirale) $= \int_0^\beta \sqrt{1+t^2}\,\mathrm{d}t$. Pour calculer cette intégrale utilisons une intégration par parties. En notant L l'intégrale, on a

$$\begin{aligned} L &= \left[t\sqrt{1+t^2}\right]_0^\beta - \int_0^\beta \frac{t^2}{\sqrt{1+t^2}}\,\mathrm{d}t = \beta\sqrt{1+\beta^2} - \int_0^\beta \frac{1+t^2}{\sqrt{1+t^2}} + \int_0^\beta \frac{1}{\sqrt{1+t^2}} \\ &= \beta\sqrt{1+\beta^2} - L + \int_0^\beta \frac{1}{\sqrt{1+t^2}} = \beta\sqrt{1+\beta^2} - L + \operatorname{argsh}(\beta). \end{aligned}$$

La longueur de la spirale est $L = \frac{1}{2}\left(\beta\sqrt{1+\beta^2} + \operatorname{argsh}(\beta)\right)$.

Pour obtenir au sens usuel la longueur d'une courbe, on prendra garde au fait que celle-ci ne doit être parcourue qu'une seule fois. Cette condition est satisfaite si par exemple le paramétrage choisi est injectif.

Exercice 13.2 Pour $a \in \mathbb{R}_+^*$, calculer la longueur Long(a) de la branche de parabole $\{(x,y) \in \mathbb{R}^2 \ ; \ y = x^2/2 \ \text{ et } \ x \in [0,a]\}$.

Intéressons-nous à une utilisation ludique de la notion de longueur d'une courbe dans le plan et dans l'espace. Un randonneur dispose d'une carte lui donnant le relief du terrain et en première approximation, la forme de la montagne est celle de la surface cartésienne d'équation $z = f(x, y)$ où

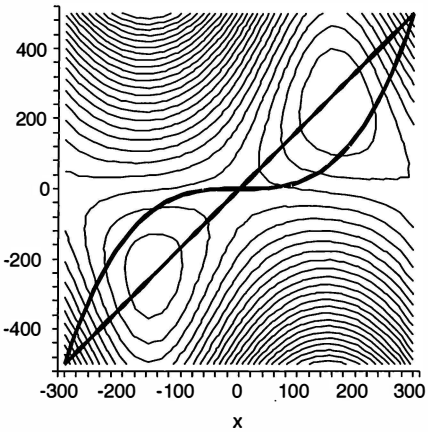
$$f : (x, y) \in [-300, 300] \times [-500, 500] \longmapsto 2500 + \frac{y}{20} \sin\left(\frac{x}{100}\right) - \left(\frac{y}{100}\right)^2 + \frac{x}{100}.$$

Entre deux points, il trace sur sa carte deux itinéraires possibles correspondant aux deux arcs paramétrés :

$$\varphi : s \in [-100, 100] \longmapsto \begin{pmatrix} 3s \\ 5s \end{pmatrix} \quad \text{et} \quad \psi : s \in [-100, 100] \longmapsto \begin{pmatrix} 3s \\ 5s^3/10000 \end{pmatrix}.$$

Pour visualiser les courbes de niveau de cette surface ainsi que les deux itinéraires, utilisons le logiciel MAPLE :

```
> with(plots):
> fig1:=contourplot(2500+y*sin(x/100)/20-(y/100)^2+x/100,
  x=-300..300, y=-500..500, contours=25):
> fig2:=plot([3*s,5*s,s=-100..100]):
> fig3:=plot([3*s,5*s^3/10000,s=-100..100]):
> display(fig1,fig2,fig3);
```



Dans un premier temps, le randonneur veut connaître la longueur de ces itinéraires sur la carte. En utilisant le logiciel MAPLE et la proposition 13.3 on obtient pour longueur des deux courbes sur la carte[21] :

$$\mathrm{Long}^{\mathrm{carte}}(\varphi) = \int_{-100}^{100} \sqrt{\varphi_1'(s)^2 + \varphi_2'(s)^2}\, ds = 1200 \text{ m},$$

$$\mathrm{Long}^{\mathrm{carte}}(\psi) = \int_{-100}^{100} \sqrt{\psi_1'(s)^2 + \psi_2'(s)^2}\, ds \approx 1272 \text{ m}.$$

[21] On suppose que l'unité de longueur est le mètre. Tous les résultats sont donnés avec une erreur inférieure à 1 mètre.

Sur la carte, le chemin le plus court est la ligne droite, c'est-à-dire l'itinéraire correspondant à l'arc paramétré φ.

Si le randonneur veut connaître la longueur sur le terrain des deux itinéraires, il doit calculer leur longueur dans l'espace. Les deux itinéraires précédents sont représentés dans l'espace par les deux arcs paramétrés :

$$\widetilde{\varphi} : s \in [-100, 100] \longmapsto \big(\varphi_1(s), \quad \varphi_2(s), \quad f(\varphi_1(s), \varphi_2(s))\big),$$

$$\widetilde{\psi} : s \in [-100, 100] \longmapsto \big(\psi_1(s), \quad \psi_2(s), \quad f(\psi_1(s), \psi_2(s))\big).$$

Les longueurs de ces itinéraires sur le terrain calculées avec MAPLE sont :

$$\mathrm{Long}^{\mathrm{terrain}}(\widetilde{\varphi}) = \int_{-100}^{100} \sqrt{\varphi_1'(s)^2 + \varphi_2'(s)^2 + (f \circ \varphi)'(s)^2}\ \mathrm{d}s \approx 1348 \text{ m},$$

$$\mathrm{Long}^{\mathrm{terrain}}(\widetilde{\psi}) = \int_{-100}^{100} \sqrt{\psi_1'(s)^2 + \psi_2'(s)^2 + (f \circ \psi)'(s)^2}\ \mathrm{d}s \approx 1303 \text{ m}.$$

Contrairement à la carte, sur le terrain le chemin le plus court est celui correspondant à l'arc paramétré $\widetilde{\psi}$. Cela n'est pas très surprenant car lors du premier calcul, le dénivelé n'a pas été pris en compte. Il faut parfois mieux faire le tour de la montagne que de passer par-dessus, mais la vue est souvent moins belle...

Intéressons-nous à présent à deux utilisations de la notion de longueur d'une courbe paramétrée. La première permet de majorer la valeur d'une intégrale le long d'une courbe C à l'aide de la longueur de cette courbe. La seconde permet de paramétrer toutes les courbes régulières en ayant toujours comme vecteur tangent un vecteur de norme 1.

PROPOSITION 13.4 *Soient f une application continue d'un ouvert U de $\mathbb{R}^n$ dans $\mathbb{R}^n$ et C une courbe paramétrée régulière de $\mathbb{R}^n$ contenue dans l'ouvert U. On a*

$$\left| \int_C f \cdot \mathrm{d}\mathbf{x} \right| \leqslant \mathrm{Long}(C) \times \sup_{\mathbf{x} \in C} \|f(\mathbf{x})\|.$$

Démonstration Soit $\varphi \in \mathcal{C}^1([a, b], \mathbb{R}^n)$ un paramétrage de la courbe C ; par définition de l'intégrale curviligne, on a

$$\left| \int_C f \cdot \mathrm{d}\mathbf{x} \right| = \left| \int_a^b f(\varphi(t)) \cdot \varphi'(t)\ \mathrm{d}t \right|.$$

En utilisant successivement l'inégalité triangulaire puis l'inégalité de Cauchy-Schwarz, on obtient :

$$\left| \int_C f \cdot \mathrm{d}\mathbf{x} \right| \leqslant \int_a^b \left| f(\varphi(t)) \cdot \varphi'(t) \right| \mathrm{d}t \leqslant \int_a^b \big\| f(\varphi(t)) \big\| \, \big\| \varphi'(t) \big\| \ \mathrm{d}t.$$

Puisque pour tout $t \in [a,b]$, on a $\|f(\varphi(t))\| \leqslant \sup_{\mathbf{x}\in C} \|f(\mathbf{x})\|$ on en déduit que

$$\left| \int_C f \cdot \mathrm{d}\mathbf{x} \right| \leqslant \left(\int_a^b \|\varphi'(t)\| \, \mathrm{d}t \right) \times \sup_{\mathbf{x}\in C} \|f(\mathbf{x})\| = \mathrm{Long}(C) \times \sup_{\mathbf{x}\in C} \|f(\mathbf{x})\|$$

ce qui conclut la preuve. □

Proposition 13.5 *Soit C une courbe paramétrée régulière dans $\mathbb{R}^n$. Il existe un paramétrage de C pour lequel la tangente à la courbe en tout point de la courbe est de norme 1.*

Démonstration Soit $\varphi \in \mathcal{C}^1([a,b],\mathbb{R}^n)$ un paramétrage de la courbe C. Considérons $\psi \in \mathcal{C}^1([c,d],\mathbb{R}^n)$ un paramétrage $\mathcal{O}^1$-équivalent à φ. D'après la définition 11.8 p. 739, il existe une application θ de $[a,b]$ vers $[c,d]$ de classe $\mathcal{C}^1$ et strictement croissante telle que $\varphi = \psi \circ \theta$. En dérivant la relation $\varphi = \psi \circ \theta$ on obtient pour tout $t \in [a,b]$ l'égalité $\varphi'(t) = \psi'(\theta(t))\theta'(t)$. Puisque l'application θ est strictement croissante, pour tout $t \in [a,b]$ on a $\theta'(t) > 0$ et par conséquent $\|\varphi'(t)\| = \theta'(t)\|\psi'(\theta(t))\|$. Si on souhaite que pour le paramétrage ψ, la tangente à la courbe en tout point de la courbe soit de norme 1, autrement dit que $\|\psi'(t)\| = 1$, alors il faut et il suffit que pour tout $t \in [a,b]$ on ait $\|\varphi'(t)\| = \theta'(t)$. En posant

$$\forall t \in [a,b] \quad \theta(t) = \int_a^t \|\varphi'(s)\| \, \mathrm{d}s,$$

on a pour tout $t \in [a,b]$ l'égalité $\|\psi'(\theta(t))\| = 1$. Notons que l'application θ ainsi définie est une application strictement croissante de $[a,b]$ vers $[0,L]$ où L est la longueur de la courbe C. Puisque θ est une bijection de $[a,b]$ vers $[0,L]$ on a $\|\psi'(\tau)\| = 1$ pour tout $\tau \in [0,L]$ ce qui prouve la proposition. □

Remarque La preuve donnée ci-dessus permet de justifier que toute courbe paramétrée régulière C dans $\mathbb{R}^n$ peut être « paramétrée par sa longueur d'arc ». Plus exactement, en notant $s(t) = \mathrm{Long}(C_t)$ la longueur de la courbe C_t paramétrée par la restriction de φ à l'intervalle $[a,t]$ et L la longueur de la courbe C, on a $s'(t) = \|\varphi'(t)\| > 0$ de sorte que l'application s est strictement croissante et donc bijective de $[a,b]$ dans $[0,L]$. On peut paramétrer la courbe C par l'application $\tau \in [0,L] \mapsto \varphi(s^{-1}(\tau))$. On dit que τ est l'*abscisse curviligne* de la courbe C.

Considérons par exemple le cercle de centre 0 et de rayon R paramétré par

$$\varphi : t \in [0,2\pi] \longmapsto (\cos(t), \sin(t)) \in \mathbb{R}^2.$$

Puisque $\mathrm{Long}(C_t) = Rt$, l'application s correspondant à ce cercle est l'application $t \in [0,2\pi] \mapsto Rt \in [0,2\pi R]$ dont la réciproque est l'application $\tau \in [0,2\pi R] \mapsto \tau/R \in [0,2\pi]$. En utilisant l'abscisse curviligne, le cercle est paramétré de la façon suivante

$$\tau \in [0,2\pi R] \longmapsto (\cos(\tau/R), \sin(\tau/R)).$$

◇

Intégrale d'une application à valeurs dans $\mathbb{R}$ le long d'une courbe

Dans les paragraphes précédents, nous avons introduit l'intégrale d'une application à valeurs dans $\mathbb{R}^n$ le long d'une courbe de $\mathbb{R}^n$. On définit l'intégrale d'une application à valeurs dans $\mathbb{R}$ le long d'une courbe de $\mathbb{R}^n$ de la façon suivante.

DÉFINITION 13.4 *Soient f une application continue sur un ouvert U de $\mathbb{R}^n$ et à valeurs dans $\mathbb{R}$ et C une courbe paramétrée dont un paramétrage est $\varphi \in \mathcal{C}^1([a,b],U)$. On appelle* intégrale curviligne *de f le long de C et on note $\int_C f \, \mathrm{d}\tau$ l'intégrale*

$$\int_C f \, \mathrm{d}\tau = \int_a^b f(\varphi(t)) \, \|\varphi'(t)\| \, \mathrm{d}t.$$

Remarquons que d'après la proposition 13.3 p. 947, la longueur d'une courbe paramétrée régulière C est donnée par :

$$\boxed{\mathrm{Long}(C) = \int_C 1 \, \mathrm{d}\tau.}$$

Donnons deux applications en physique de l'intégrale curviligne, l'une ayant trait à la géométrie des masses, l'autre à l'optique.

✗ Considérons un fil métallique dont la forme épouse celle d'une courbe paramétrée C de $\mathbb{R}^2$ et dont la densité linéique en un point $(x,y) \in C$ est $\rho(x,y) \in \mathbb{R}_+^*$. La masse M de ce fil et les coordonnées (x_G, y_G) du centre de gravité de ce fil sont données par des intégrales curvilignes suivantes :

$$\mathrm{M} = \int_C \rho \, \mathrm{d}\tau, \quad x_G = \frac{1}{M} \int_C f_x \, \mathrm{d}\tau \quad \text{et} \quad y_G = \frac{1}{M} \int_C f_y \, \mathrm{d}\tau.$$

où

$$f_x : (x,y) \in C \longmapsto x\,\rho(x,y) \quad \text{et} \quad f_y : (x,y) \in C \longmapsto y\,\rho(x,y).$$

Si le fil métallique est homogène (de densité constante $\rho = \rho_0$) et a la forme d'un arc de cercle paramétré par $\varphi : t \in [0,\pi] \mapsto (\cos(t), \sin(t))$ alors sa masse vaut :

$$\mathrm{M} = \int_C \rho_0 \, \mathrm{d}\tau = \rho_0 \int_0^\pi 1 \, \mathrm{d}t = \rho_0 \pi,$$

et les coordonnées de son centre de gravité valent :

$$x_G = \frac{1}{\rho_0 \pi} \int_0^\pi \rho_0 \cos(t) \, \mathrm{d}t = 0 \quad \text{et} \quad y_G = \frac{1}{\rho_0 \pi} \int_0^\pi \rho_0 \sin(t) \, \mathrm{d}t = \frac{2}{\pi}.$$

Exercice 13.3 Déterminer la masse et les coordonnées du centre de gravité d'un fil rectiligne, de longueur $L \in \mathbb{R}_+^*$, dont la densité linéique de masse en un point situé à une distance x de l'une des extrémités choisie comme origine vaut $1 + x^2$.

✗Selon les lois de l'optique, l'indice de réfraction n d'un rayon lumineux en un point dépend de la vitesse v de la lumière en ce point. On a la relation $n = c/v$ où c est la vitesse de la lumière dans le vide. Le temps que met la lumière pour aller d'un point A à un point B en suivant une courbe C vaut $T_C = \int_C 1/v \, \mathrm{d}\tau$ et est par conséquent proportionnel à $\int_C n \, \mathrm{d}\tau$. Selon le principe de Fermat[22], parmi tous les chemins possibles pour aller d'un point à un autre, la lumière choisit le plus court en temps. Autrement dit, pour aller du point A au point B, la lumière empreinte le chemin qui minimise la valeur de $\int_C n \, \mathrm{d}\tau$.

Par exemple, dans un milieu homogène (c'est-à-dire lorsque l'indice de réfraction $n = n_0$ est constant), l'intégrale curviligne $\int_C n \, \mathrm{d}\tau$ vaut $n_0 \mathrm{Long}(C)$ et la lumière empreinte le chemin qui minimise la longueur entre les points A et B, c'est-à-dire le segment $[AB]$: on en déduit que dans un milieu homogène, la lumière se propage de façon rectiligne. Lorsque n n'est pas constant, le problème de trouver le chemin C entre les points A et B qui minimise la quantité $\int_C n \, \mathrm{d}\tau$ est plus difficile[23]. Voici deux applications de ce phénomène.

- Pour transmettre une information sous forme d'onde on utilise des fibres optiques à gradient d'indice. Ce sont des fibres cylindriques dans lesquelles on fait varier l'indice de réfraction n pour « guider » les ondes (voir la figure 3 à gauche).
- La même raison physique explique l'effet des mirages. Au voisinage du sol chauffé par le soleil, l'indice de réfraction de l'air est plus faible qu'en altitude. Les rayons lumineux ont alors des trajectoires non rectilignes, ce qui explique qu'en regardant dans une direction, on puisse voir un objet qui est situé dans une autre direction (voir la figure 3 à droite).

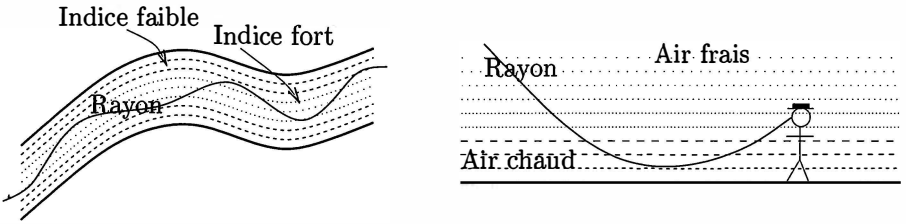


Fig. 3 Guide d'onde (à gauche) et phénomène de mirage (à droite).

[22] Fermat, Pierre, de (1601, Beaumont-de-Lomagne - 1665, Castres).

[23] Il existe une reformulation de ce problème sous la forme d'équations aux dérivées partielles qui dans certains cas permet d'obtenir une expression du chemin minimisant.

13.2 Circulation et formes différentielles

13.2.1 Formes différentielles de degré 1

Définition 13.5 *Soit U un ouvert de $\mathbb{R}^n$. On note $\Lambda^1(U)$ l'ensemble des applications définies sur U à valeurs dans*[24] $\mathcal{L}(\mathbb{R}^n, \mathbb{R})$. *De telles applications sont appelées des* formes différentielles de degré 1 *ou plus simplement des* 1-formes.

Nous avons déjà rencontré des 1-formes : par exemple[25] la différentielle $\mathrm{d}f$ d'une application f définie sur un ouvert $U \subset \mathbb{R}^n$ à valeurs dans $\mathbb{R}$ est une 1-forme puisque

$$\forall \mathbf{x} \in U \quad \mathrm{d}f_{\mathbf{x}} \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}).$$

En particulier, pour tout $i \in \{1, \ldots, n\}$, la i–ième projection canonique $p_i : (x_1, \ldots, x_n) \in \mathbb{R}^n \mapsto x_i \in \mathbb{R}$ est une application linéaire dont la différentielle ne dépend pas[26] de $\mathbf{x} = (x_1, \ldots, x_n)$; on a

$$\forall \mathbf{x} \in \mathbb{R}^n \qquad \mathrm{d}p_{i_{\mathbf{x}}} : (h_1, \ldots, h_n) \in \mathbb{R}^n \longmapsto h_i \in \mathbb{R}$$

On note $\mathrm{d}x_i$ cette 1-forme : $\mathrm{d}x_i = \mathrm{d}p_{i_{\mathbf{x}}}$. La proposition suivante, que nous admettrons, indique en quoi ces 1-formes $\mathrm{d}x_i$ sont particulièrement importantes.

Proposition 13.6 *Soit U un ouvert de $\mathbb{R}^n$; la famille $\{\mathrm{d}x_1, \ldots, \mathrm{d}x_n\}$ engendre $\Lambda^1(U)$ au sens suivant : pour toute 1-forme $\omega \in \Lambda^1(U)$ il existe n applications $\{\alpha_1, \ldots, \alpha_n\}$ de U dans $\mathbb{R}$ telles que*

$$\omega = \sum_{i=1}^{n} \alpha_i \mathrm{d}x_i.$$

De plus, cette décomposition est unique.

De façon plus condensée, on écrit $\omega = \alpha \cdot \mathrm{d}\mathbf{x}$ plutôt que $\omega = \sum_{i=1}^n \alpha_i \mathrm{d}x_i$. Cette notation est justifiée par le fait que le terme $\sum_{i=1}^n \alpha_i \mathrm{d}x_i$ peut s'obtenir en effectuant « formellement » le produit de $\alpha = (\alpha_1, \ldots, \alpha_n)$ par $\mathrm{d}\mathbf{x} = (\mathrm{d}x_1, \ldots, \mathrm{d}x_n)$ de la même manière qu'est calculé le produit scalaire canonique entre vecteurs de $\mathbb{R}^n$.

[24] L'ensemble $\mathcal{L}(\mathbb{R}^n, \mathbb{R})$ est l'ensemble des applications linéaires de $\mathbb{R}^n$ dans $\mathbb{R}$.

[25] Voir la définition 10.1 p. 609.

[26] Voir la proposition 10.5 p. 615.

La différentielle d'une application différentiable f d'un ouvert U de $\mathbb{R}^n$ dans $\mathbb{R}$ est donnée par $df : \mathbf{x} \in U \mapsto \mathrm{d}f_{\mathbf{x}} \in \mathcal{L}(\mathbb{R}^n, \mathbb{R})$ où pour tout $\mathbf{h} \in \mathbb{R}^n$ on a[27]

$$\mathrm{d}f_{\mathbf{x}}(\mathbf{h}) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(\mathbf{x})\ h_j.$$

Par définition des 1-formes $\mathrm{d}x_i$, on peut encore écrire cette relation de la façon suivante :

$$\forall \mathbf{x} \in U \quad \forall \mathbf{h} \in \mathbb{R}^n \quad \mathrm{d}f_{\mathbf{x}}(\mathbf{h}) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(\mathbf{x})\ \mathrm{d}x_j(\mathbf{h}).$$

Ainsi la décomposition de la 1-forme $\mathrm{d}f$ selon la famille $\{\mathrm{d}x_1, \ldots, \mathrm{d}x_n\}$ s'écrit

$$\boxed{\mathrm{d}f = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}\ \mathrm{d}x_j.} \tag{4}$$

On écrit aussi parfois cette relation sous la forme $\mathrm{d}f = \nabla f \cdot \mathrm{d}\mathbf{x}$.

Exemples

1. Considérons l'application f de $U = \{(x_1, x_2) \in \mathbb{R}^2\ ;\ x_1 \neq 0\}$ dans $\mathbb{R}$ définie par $f(x_1, x_2) = \arctan(x_2/x_1)$. Cette application est de classe $\mathcal{C}^\infty$ sur U et ses deux dérivées partielles en $(x_1, x_2) \in U$ sont :

$$\frac{\partial f}{\partial x_1}(x_1, x_2) = \frac{-x_2}{x_1^2 + x_2^2} \quad \text{et} \quad \frac{\partial f}{\partial x_2}(x_1, x_2) = \frac{x_1}{x_1^2 + x_2^2}.$$

La différentielle $\mathrm{d}f$ est une 1-forme qui s'exprime de la manière suivante :

$$\mathrm{d}f = \alpha_1 \mathrm{d}x_1 + \alpha_2 \mathrm{d}x_2,$$

où α_1 et α_2 sont les applications définies par

$$\alpha_1 : (x_1, x_2) \in U \longmapsto \frac{-x_2}{x_1^2 + x_2^2} \in \mathbb{R} \quad \text{et} \quad \alpha_2 : (x_1, x_2) \in U \longmapsto \frac{x_1}{x_1^2 + x_2^2} \in \mathbb{R}.$$

On écrit souvent de façon condensé, pour $\mathbf{x} = (x_1, x_2) \in U$,

$$\mathrm{d}f_{\mathbf{x}} = \frac{-x_2}{x_1^2 + x_2^2}\,\mathrm{d}x_1 + \frac{x_1}{x_1^2 + x_2^2}\,\mathrm{d}x_2.$$

2. Si la différentielle de $f : (x_1, x_2) \in U \mapsto \dfrac{x_1 x_2^2}{x_1^2 + x_2^2}$ est donnée par

$$\mathrm{d}f_{\mathbf{x}} = \frac{x_2^2(x_2^2 - x_1^2)}{(x_1^2 + x_2^2)^2}\,\mathrm{d}x_1 + \frac{2x_1^3 x_2}{(x_1^2 + x_2^2)^2}\,\mathrm{d}x_2,$$

[27] Voir la Proposition 10.9 p. 632.

cela signifie que

$$\mathrm{d}f_{\mathbf{x}} : (h_1, h_2) \in \mathbb{R}^2 \longmapsto \frac{x_2^2(x_2^2 - x_1^2)}{(x_1^2 + x_2^2)^2} h_1 + \frac{2x_1^3 x_2}{(x_1^2 + x_2^2)^2} h_2.$$

En particulier, on a $\mathrm{d}f_{(1,1)}(h_1, h_2) = \dfrac{h_2}{2}$.

Exercice 13.4 On considère l'application

$$f : (x_1, x_2, x_3) \in U \longmapsto x_1 \sqrt{\frac{x_2}{4x_1 - x_2 x_3^2}}$$

où $U = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \ ; \ x_1 > 0, \ x_2 > 0, \ x_3 > 0 \ \text{ et } \ 4x_1 - x_2 x_3^2 > 0\}$.
1 - Exprimer $h(x_1, x_2, x_3) = \ln(f(x_1, x_2, x_3))$ pour tout $(x_1, x_2, x_3) \in U$.
2 - Calculer la différentielle $\mathrm{d}h$ de l'application h sur U.
3 - En déduire l'expression de $\dfrac{\mathrm{d}f}{f}$ sur U, et donner l'expression de $\mathrm{d}f_{(1,1,1)}$.

DÉFINITION 13.6 *Soit $\omega = \sum_{i=1}^n \alpha_i \mathrm{d}x_i$ une 1-forme définie sur un ouvert U de $\mathbb{R}^n$.*

✗ *La 1-forme ω est* constante *si pour tout $i \in \{1, \ldots, n\}$ l'application α_i est une application constante sur U.*

✗ *La 1-forme ω est de classe $\mathcal{C}^k$ où $k \in \mathbb{N} \cup \{\infty\}$ si pour tout $i \in \{1, \ldots, n\}$ l'application α_i est une application de classe $\mathcal{C}^k$ sur U.*

Une application f de U dans $\mathbb{R}$ de classe $\mathcal{C}^k$ ($k \in \mathbb{N}$) admet des dérivées partielles $\frac{\partial f}{\partial x_i}$ qui sont des applications de classe $\mathcal{C}^{k-1}$. La relation (4) implique dans ce cas que la 1-forme $\mathrm{d}f$ est de classe $\mathcal{C}^{k-1}$. Si f est de classe $\mathcal{C}^\infty$ sur U alors la 1-forme $\mathrm{d}f$ est de classe $\mathcal{C}^\infty$.

Pour l'instant les seules 1-formes que nous avons rencontrées sont les différentielles d'applications. On peut se demander si toutes les 1-formes sont de ce type. La réponse est non et la proposition suivante donne une condition nécessaire pour qu'une 1-forme soit la différentielle d'une application.

PROPOSITION 13.7 *Soit $\omega = \sum_{i=1}^n \alpha_i \mathrm{d}x_i$ une 1-forme définie sur un ouvert U de $\mathbb{R}^n$. S'il existe $f \in \mathcal{C}^2(U, \mathbb{R})$ telle que $w = \mathrm{d}f$ alors ω est une 1-forme de classe $\mathcal{C}^1$ et*

$$\forall (i,j) \in \{1, \ldots, n\}^2 \qquad \frac{\partial \alpha_i}{\partial x_j} = \frac{\partial \alpha_j}{\partial x_i}. \tag{5}$$

Démonstration Supposons que $w = \mathrm{d}f$ avec $f \in \mathcal{C}^2(U, \mathbb{R})$; on a d'une part $\mathrm{d}f = \sum_{j=1}^n \frac{\partial f}{\partial x_j}\,\mathrm{d}x_j$ et d'autre part $\omega = \sum_{i=1}^n \alpha_i\,\mathrm{d}x_i$. Par unicité de la décomposition d'une 1-forme[28] selon la famille $\{\mathrm{d}x_1, \ldots, \mathrm{d}x_n\}$, on en déduit que pour tout $i \in \{1, \ldots, n\}$ on a $\alpha_i = \frac{\partial f}{\partial x_i}$. L'application f étant de classe $\mathcal{C}^2$ sur U, les applications α_i sont toutes de classe $\mathcal{C}^1$ sur U et on a

$$\forall (i,j) \in \{1, \ldots, n\}^2 \qquad \frac{\partial \alpha_i}{\partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i} \quad \text{et} \quad \frac{\partial \alpha_j}{\partial x_i} = \frac{\partial^2 f}{\partial x_i \partial x_j}.$$

Puisque l'application f est de classe $\mathcal{C}^2$, le théorème de Schwarz[29] indique que :

$$\forall (i,j) \in \{1, \ldots, n\}^2 \qquad \frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial^2 f}{\partial x_i \partial x_j}.$$

Cela implique que $\dfrac{\partial \alpha_i}{\partial x_j} = \dfrac{\partial \alpha_j}{\partial x_i}$; le résultat est démontré. □

Remarque La proposition 13.7 est le plus souvent utilisée sous sa forme contraposée. Si $\omega = \sum_{i=1}^n \alpha_i\,\mathrm{d}x_i$ est une 1-forme sur un ouvert U de $\mathbb{R}^n$ où les applications α_i sont toutes de classe $\mathcal{C}^1$ et s'il existe $(i,j) \in \{1, \ldots, n\}^2$ tel que $\frac{\partial \alpha_i}{\partial x_j} \neq \frac{\partial \alpha_j}{\partial x_i}$ alors il n'existe pas d'application $f \in \mathcal{C}^2(U, \mathbb{R})$ telle que $\omega = \mathrm{d}f$. ◇

Exemples

1. Considérons la forme différentielle définie sur $U = \{(x_1, x_2) \in \mathbb{R}^2 \ ; \ x_1 \neq 0\}$ par

$$\omega = \frac{-x_2}{x_1^2 + x_2^2}\,\mathrm{d}x_1 + \frac{x_1}{x_1^2 + x_2^2}\,\mathrm{d}x_2.$$

On a vu p. 957 que $\omega = \mathrm{d}f$ où $f : (x_1, x_2) \in U \mapsto \arctan(x_2/x_1)$. On peut vérifier sur cet exemple que la condition (5) est respectée par

$$\alpha_1 : (x_1, x_2) \in U \longmapsto \frac{-x_2}{x_1^2 + x_2^2} \in \mathbb{R} \quad \text{et} \quad \alpha_2 : (x_1, x_2) \in U \longmapsto \frac{x_1}{x_1^2 + x_2^2} \in \mathbb{R}.$$

2. Considérons la 1-forme $\omega : (x_1, x_2) \in \mathbb{R}^2 \mapsto 2x_2\mathrm{d}x_1 + x_1\mathrm{d}x_2$. Elle est du type $\omega = \alpha_1\mathrm{d}x_1 + \alpha_2\mathrm{d}x_2$ où les applications α_1 et α_2 sont définies par

$$\alpha_1 : (x_1, x_2) \in \mathbb{R}^2 \longmapsto 2x_2 \quad \text{et} \quad \alpha_2 : (x_1, x_2) \in \mathbb{R}^2 \longmapsto x_1.$$

Puisque pour tout $(x_1, x_2) \in \mathbb{R}^2$

$$\frac{\partial \alpha_1}{\partial x_2}(x_1, x_2) = 2 \quad \text{et} \quad \frac{\partial \alpha_2}{\partial x_1}(x_1, x_2) = 1,$$

on en déduit qu'il n'existe pas d'application $f \in \mathcal{C}^2(\mathbb{R}^2, \mathbb{R})$ telle que $\omega = \mathrm{d}f$.

[28] Voir la proposition 13.6.

[29] Voir le théorème de Schwarz à la p. 650.

3. La condition (5) est une condition nécessaire pour qu'une 1-forme soit la différentielle d'une application de classe $\mathcal{C}^2$ mais cette condition n'est pas suffisante. Ainsi, la 1-forme

$$\omega : (x_1, x_2) \in \mathbb{R}^2_* \longmapsto -\frac{x_2}{x_1^2 + x_2^2}\, \mathrm{d}x_1 + \frac{x_1}{x_1^2 + x_2^2}\, \mathrm{d}x_2$$

s'écrit $\omega = \alpha_1\, \mathrm{d}x_1 + \alpha_2\, \mathrm{d}x_2$ où

$$\alpha_1 : (x_1, x_2) \in \mathbb{R}^2_* \longmapsto -\frac{x_2}{x_1^2 + x_2^2} \quad \text{et} \quad \alpha_2 : (x_1, x_2) \in \mathbb{R}^2_* \longmapsto \frac{x_1}{x_1^2 + x_2^2}.$$

Ces deux applications sont de classe $\mathcal{C}^1$ sur l'ouvert $\mathbb{R}^2_*$ et on a, pour tout $(x_1, x_2) \in \mathbb{R}^2_*$

$$\frac{\partial \alpha_1}{\partial x_2}(x_1, x_2) = \frac{x_2^2 - x_1^2}{(x_1^2 + x_2^2)^2} \quad \text{et} \quad \frac{\partial \alpha_2}{\partial x_1}(x_1, x_2) = \frac{x_2^2 - x_1^2}{(x_1^2 + x_2^2)^2}. \tag{6}$$

Ces deux dérivées partielles étant égales, on en déduit que la 1-forme ω vérifie la condition (5). Nous verrons p. 971 qu'il n'existe toutefois pas d'application f de classe $\mathcal{C}^2$ sur $\mathbb{R}^2$ tel que $\omega = \mathrm{d}f$. La condition (5) n'est donc pas une condition suffisante pour que ω soit la différentielle d'une application.

Cette constatation conduit à introduire la définition suivante.

DÉFINITION 13.7 *Soient U un ouvert de $\mathbb{R}^n$ et $\omega = \sum_{i=1}^n \alpha_i\, \mathrm{d}x_i \in \Lambda^1(U)$.*

✗ *Si toutes les applications α_i, $i \in \{1, \ldots, n\}$ sont de classe $\mathcal{C}^1$ et si*

$$\forall (i,j) \in \{1, \ldots, n\}^2 \qquad \frac{\partial \alpha_i}{\partial x_j} = \frac{\partial \alpha_j}{\partial x_i}$$

alors ω est dite fermée.

✗ *S'il existe $f \in \mathcal{C}^2(U, \mathbb{R})$ telle que $w = \mathrm{d}f$ alors ω est dite* exacte.

Exemples

1. On a établi p. 959 que la 1-forme $\omega : (x_1, x_2) \in \mathbb{R}^2 \mapsto 2x_2 \mathrm{d}x_1 + x_1 \mathrm{d}x_2$ n'est pas une 1-forme fermée. On déduit de la proposition 13.7 que ω n'est pas une 1-forme exacte.

2. La relation (6) se traduit de la façon suivante : la 1-forme

$$\omega : (x_1, x_2) \in \mathbb{R}^2_* \longmapsto \frac{-x_2}{x_1^2 + x_2^2}\, \mathrm{d}x_1 + \frac{x_1}{x_1^2 + x_2^2}\, \mathrm{d}x_2$$

est une 1-forme fermée.

3. Considérons la 1-forme $\omega : (x_1, x_2) \in \mathbb{R}^2 \mapsto \alpha_1 \mathrm{d}x_1 + \alpha_2 \mathrm{d}x_2$ où

$$\alpha_1 : (x_1, x_2) \in \mathbb{R}^2 \longmapsto (2x_1 x_2^3 + 1) \quad \text{et} \quad \alpha_2 : (x_1, x_2) \in \mathbb{R}^2 \longmapsto (3x_1^2 x_2^2 - 2x_2)$$

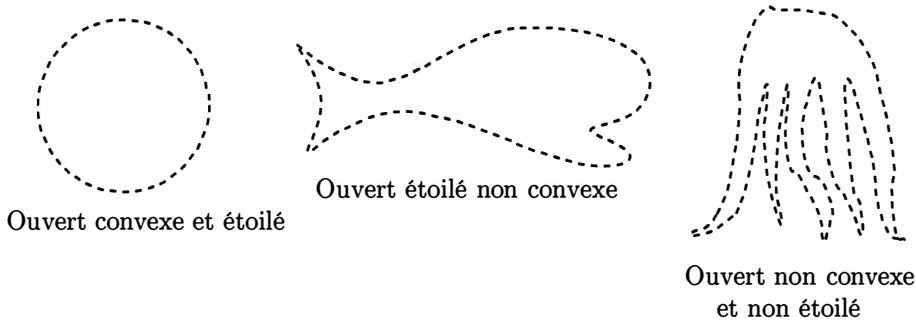


Fig. 4 Illustration de la notion d'ouvert étoilé.

Pour tout $(x_1, x_2) \in \mathbb{R}^2$, on a

$$\frac{\partial \alpha_1}{\partial x_2}(x_1, x_2) = 6x_1x_2^2 \quad \text{et} \quad \frac{\partial \alpha_2}{\partial x_1}(x_1, x_2) = 6x_1x_2^2.$$

Ces deux dérivées partielles étant égales, ω est une 1-forme fermée.

La proposition 13.7 peut se reformuler de la manière suivante : si une 1-forme est exacte alors elle est fermée. Comme nous l'avons mentionné, la réciproque est fausse. Néanmoins, moyennant certaines hypothèses sur l'ouvert U, pour une 1-forme sur U de classe $\mathcal{C}^1$ on a équivalence entre être exacte et être fermée. Afin de formuler une condition sur l'ouvert U, nous allons introduire la notion d'ouvert étoilé.

DÉFINITION 13.8 (Ouvert étoilé)
Un ouvert U de $\mathbb{R}^n$ est dit étoilé *si*

$$\exists \mathbf{a} \in U \quad \forall \mathbf{x} \in U \quad [\mathbf{a}, \mathbf{x}] \subset U,$$

où $[\mathbf{a}, \mathbf{x}] = \{\mathbf{z} \in \mathbb{R}^n \ ; \ \exists \lambda \in [0,1], \ \mathbf{z} = \mathbf{a} + \lambda(\mathbf{x} - \mathbf{a})\}$ *désigne le segment d'extrémités* $\mathbf{a}$ *et* $\mathbf{x}$ *dans* $\mathbb{R}^n$.

Remarques

1. On peut interpréter la définition 13.8 de la manière suivante : un ouvert U de $\mathbb{R}^n$ est étoilé lorsqu'il contient un point $\mathbf{a}$ depuis lequel on peut voir tous les points de l'ouvert (en supposant que les rayons lumineux se propagent en lignes droites, et qu'il n'y a pas de miroir !), voir la figure 4.

2. Les ouverts étoilés de $\mathbb{R}$ correspondent aux intervalles ouverts. Pour tout $n \in \mathbb{N}$, l'ouvert $\mathbb{R}^n$ est étoilé.

3. Tout ouvert convexe[30] est un ouvert étoilé. Par contre, la réciproque est fausse : il existe des ouverts qui sont étoilés sans être convexes (voir les exemples qui suivent). ◇

[30] Voir le *Cours de première année* p. 738 pour une définition d'ensemble convexe.

Exemples

1. Montrons que l'ouvert $U = \mathbb{R}^2 \setminus \{(x_1, x_2) \in \mathbb{R}^2 \ ; \ x_1 \leqslant 0 \text{ et } x_2 = 0\}$ est un ouvert étoilé. Considérons le point $\mathbf{a} = (1,0) \in U$ et montrons que pour tout $\mathbf{x} \in U$ on a $[\mathbf{a}, \mathbf{x}] \subset U$. Il s'agit de montrer que pour tout $\mathbf{x} \in U$ et pour tout $\lambda \in [0,1]$, l'élément $\mathbf{z} = \mathbf{a} + \lambda(\mathbf{x} - \mathbf{a})$ appartient à l'ouvert U, voir la figure 5. Notons $\mathbf{x} = (x_1, x_2)$ et $\mathbf{z} = (z_1, z_2)$; on a

$$z_1 = 1 + \lambda(x_1 - 1) \quad \text{et} \quad z_2 = \lambda x_2.$$

Montrer que $\mathbf{z} \in U$ revient à prouver que $\mathbf{z} \notin \{(x_1, x_2) \in \mathbb{R}^2 \ ; \ x_1 \leqslant 0 \text{ et } x_2 = 0\}$ c'est-à-dire que $z_1 > 0$ ou que $z_2 \neq 0$. Pour cela on va utiliser le fait que $\mathbf{x} \in U$ c'est-à-dire que $x_1 > 0$ ou que $x_2 \neq 0$.

- Soit $x_1 > 0$ et dans ce cas $z_1 = 1 + \lambda(x_1 - 1) > 1 + (\lambda - 1) \geqslant 0$ donc $z_1 > 0$.
- Soit $x_2 \neq 0$ et dans ce cas l'égalité $z_2 = \lambda x_2$ implique que $z_2 \neq 0$ ou bien que $\lambda = 0$ mais dans ce dernier cas, $z_1 = 1 > 0$.

Dans tous les cas, on en déduit que $z_1 > 0$ ou que $z_2 \neq 0$, c'est-à-dire que $\mathbf{z} \in U$ et par conséquent que l'ouvert U est étoilé.

Cet ouvert U fournit un exemple d'ouvert étoilé qui n'est pas convexe. Pour montrer que U n'est pas convexe, il suffit de déterminer deux éléments $\mathbf{b}$ et $\mathbf{c}$ de U tels que le segment $[\mathbf{b}, \mathbf{c}]$ ne soit pas inclus dans U. En s'inspirant de la figure 5, on pose $\mathbf{b} = (-1, 1)$ et $\mathbf{c} = (-1, -1)$ (on vérifie aisément que $(\mathbf{b}, \mathbf{c}) \in U^2$). Le point $\mathbf{b} + \frac{1}{2}(\mathbf{c} - \mathbf{b}) = (-1, 0)$ est un élément du segment $[\mathbf{b}, \mathbf{c}]$ qui n'est pas dans l'ouvert U, ce qui montre que U n'est pas convexe.

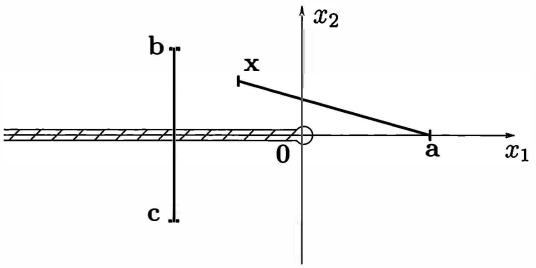


Fig. 5 Illustration de la situation considérée.

2. Montrons que l'ouvert $U = \mathbb{R}^2 \setminus \{\mathbf{0}\}$ n'est pas un ouvert étoilé. Il s'agit de montrer que[31]

$$\forall \mathbf{a} \in U \quad \exists \mathbf{x} \in U \quad [\mathbf{a}, \mathbf{x}] \not\subset U.$$

Pour $\mathbf{a} \in U$, en choisissant $\mathbf{x} = -\mathbf{a}$, on a $\mathbf{x} \in U$ et $\mathbf{0} \in [\mathbf{a}, \mathbf{x}]$ mais $\mathbf{0} \notin U$. Par conséquent $[\mathbf{a}, \mathbf{x}] \not\subset U$, ce qui prouve que l'ouvert $\mathbb{R}^2 \setminus \{\mathbf{0}\}$ n'est pas étoilé.

[31] Cette assertion est la négation de l'assertion définissant un ouvert étoilé.

THÉORÈME 13.1 (Théorème de Poincaré)
Si U est un ouvert étoilé alors toute 1-forme fermée sur U est une 1-forme exacte.

Démonstration Soit $\omega = \sum_{i=1}^{n} \alpha_i \, \mathrm{d}x_i$ une 1-forme fermée sur un ouvert étoilé U de $\mathbb{R}^n$. Les applications α_i sont de classe $\mathcal{C}^1$ et vérifient[32]

$$\forall (i,j) \in \{1,\dots,n\}^2 \qquad \frac{\partial \alpha_i}{\partial x_j} = \frac{\partial \alpha_j}{\partial x_i}.$$

Pour montrer que ω est une 1-forme exacte, il faut exhiber une fonction f telle que $\mathrm{d}f = \omega$. L'unicité de la décomposition[33] d'une 1-forme selon la famille $\{\mathrm{d}x_1,\dots,\mathrm{d}x_n\}$, indique que la fonction f recherchée est telle que

$$\forall i \in \{1,\dots,n\} \quad \frac{\partial f}{\partial x_i} = \alpha_i. \tag{7}$$

Par hypothèse, l'ouvert U est étoilé donc il existe $\mathbf{a} \in U$ tel que pour tout élément $\mathbf{x} \in U$ on ait $[\mathbf{a},\mathbf{x}] \subset U$. Cela implique que pour tout $t \in [0,1]$ on a $\mathbf{a} + t(\mathbf{x}-\mathbf{a}) \in U$. Par conséquent la fonction

$$f : \mathbf{x} \in U \longmapsto \int_0^1 \sum_{j=1}^{n} (x_j - a_j)\ \alpha_j(\mathbf{a} + t(\mathbf{x}-\mathbf{a}))\ \mathrm{d}t$$

est bien définie sur U. Nous allons vérifier que cette application satisfait $\mathrm{d}f = \omega$ en montrant que f vérifie la relation (7).

Soit $i \in \{1,\dots,n\}$ fixé ; pour tout $\mathbf{x} \in U$ on a

$$\frac{\partial f}{\partial x_i}(\mathbf{x}) = \int_0^1 \sum_{j=1}^{n} \frac{\partial}{\partial x_i}\Big((x_j - a_j)\ \alpha_j(\mathbf{a} + t(\mathbf{x}-\mathbf{a}))\Big)\ \mathrm{d}t.$$

Pour $(i,j) \in \{1,\dots,n\}^2$ et $t \in [0,1]$, calculons

$$A_{i,j}(t) = \frac{\partial}{\partial x_i}\Big((x_j - a_j)\ \alpha_j(\mathbf{a} + t(\mathbf{x}-\mathbf{a}))\Big).$$

Si $i \neq j$, on a $A_{i,j}(t) = (x_j - a_j)\ \dfrac{\partial}{\partial x_i}\Big(\alpha_j(\mathbf{a} + t(\mathbf{x}-\mathbf{a}))\Big)$ et la formule de dérivation des applications composées indique que

$$A_{i,j}(t) = (x_j - a_j)\ t\ \frac{\partial \alpha_j}{\partial x_i}(\mathbf{a} + t(\mathbf{x}-\mathbf{a})).$$

Si $i = j$ alors en utilisant la formule de dérivation d'un produit, on obtient

$$A_{i,i}(t) = \alpha_i(\mathbf{a} + t(\mathbf{x}-\mathbf{a})) + (x_i - a_i)\ t\ \frac{\partial \alpha_j}{\partial x_i}(\mathbf{a} + t(\mathbf{x}-\mathbf{a})).$$

[32] Voir la définition 13.7 p. 960.
[33] Voir la proposition 13.6 p. 956.

POINCARÉ, Henri (1854, Nancy - 1912, Paris).



Henri Poincaré est connu pour ses travaux sur les formes différentielles mais également pour ses études sur la stabilité du système solaire et pour ses réflexions sur les fondements des sciences. En 30 ans, il a publié une trentaine d'ouvrages, et près de 500 notes, articles ou longs mémoires. Ses travaux ont changé totalement le paysage mathématique de son époque. Il a révolutionné l'étude des équations différentielles par ses études qualitatives de solutions. Son cousin Raymond Poincaré fut Président de la République Française de 1913 à 1920.

On en déduit que

$$\frac{\partial f}{\partial x_i}(\mathbf{x}) = \int_0^1 \Big(\alpha_i(\mathbf{a}+t(\mathbf{x}-\mathbf{a})) + \sum_{j=1}^{n}(x_j-a_j)\,t\,\frac{\partial \alpha_j}{\partial x_i}(\mathbf{a}+t(\mathbf{x}-\mathbf{a}))\Big)\,\mathrm{d}t. \quad (8)$$

Par ailleurs, l'application $g : t \in [0,1] \mapsto t\,\alpha_i(\mathbf{a}+t(\mathbf{x}-\mathbf{a}))$ est différentiable et on a

$$g'(t) = \alpha_i(\mathbf{a}+t(\mathbf{x}-\mathbf{a})) + \sum_{j=1}^{n}(x_j-a_j)\,t\,\frac{\partial \alpha_i}{\partial x_j}(\mathbf{a}+t(\mathbf{x}-\mathbf{a})). \quad (9)$$

Puisque ω est une 1-forme fermée, on a $\frac{\partial \alpha_i}{\partial x_j} = \frac{\partial \alpha_j}{\partial x_i}$ pour tout $(i,j) \in \{1,\dots,n\}^2$ et les relations (8) et (9) indiquent que pour tout $i \in \{1,\dots n\}$

$$\frac{\partial f}{\partial x_i}(\mathbf{x}) = \int_0^1 g'(t)\,\mathrm{d}t = g(1) - g(0) = \alpha_i(\mathbf{x}).$$

L'application f satisfait donc la relation (7), ce qui permet de conclure que la forme différentielle ω est exacte. □

Remarque Plus généralement, le théorème de Poincaré est vrai pour des ouverts appelés ouverts *contractiles*. Un ouvert U est dit contractile s'il existe un élément $\mathbf{p}$ de U et une application continue $r : U \times [0,1] \to U$ vérifiant pour tout $\mathbf{x} \in U$ les relations $r(\mathbf{x},0) = \mathbf{x}$ et $r(\mathbf{x},1) = \mathbf{p}$. La démonstration du théorème de Poincaré sous cette hypothèse est plus difficile mais l'idée reste la même : trouver une « primitive » de la 1-forme ω. ◇

Exemple Nous avons vu p. 960 que la 1-forme

$$\omega : (x_1,x_2) \in \mathbb{R}^2_* \longmapsto \frac{-x_2}{x_1^2+x_2^2}\,\mathrm{d}x_1 + \frac{x_1}{x_1^2+x_2^2}\,\mathrm{d}x_2$$

est une 1-forme fermée mais nous avons vu p. 962 que l'ouvert $\mathbb{R}^2_*$ n'est pas un ouvert étoilé. On ne peut donc pas utiliser le théorème de Poincaré. Nous verrons ultérieurement que la 1-forme ω sur $\mathbb{R}^2_*$ n'est pas exacte. Par contre, nous avons vu p. 962 que l'ouvert $U = \mathbb{R}^2 \setminus D$ où $D = \{(x_1, 0)\ ;\ x_1 \leqslant 0\}$ est un ouvert étoilé. Par conséquent, si on s'intéresse à la restriction de la 1-forme ω à U alors, d'après le théorème de Poincaré, il existe une application f de U dans $\mathbb{R}$ telle que $\mathrm{d}f = \omega$ sur U.

On notera que le théorème de Poincaré assure l'existence d'une application f telle que $\omega = \mathrm{d}f$ mais ne donne pas l'application f de façon explicite[34]. Dans le cas considéré ici, on peut vérifier que[35] l'application

$$f : (x_1, x_2) \in U \longmapsto \begin{cases} \arctan(x_2/x_1) & \text{si } x_1 \neq 0 \\ \pi/2 & \text{si } x_1 = 0 \text{ et } x_2 > 0 \\ -\pi/2 & \text{si } x_1 = 0 \text{ et } x_2 < 0 \end{cases}$$

satisfait $\omega = \mathrm{d}f$.

Application à l'étude de solutions d'équations différentielles

L'introduction des 1-formes différentielles et du théorème de Poincaré permet d'étudier certaines équations différentielles non linéaires. Considérons deux applications α_1 et α_2 de classe $\mathcal{C}^1$ de $\mathbb{R}^2$ dans $\mathbb{R}$; résoudre l'équation différentielle

$$\alpha_1(x, y)\, y' + \alpha_2(x, y) = 0 \tag{10}$$

sur $\mathbb{R}$ consiste à déterminer les applications $y : x \in \mathbb{R} \mapsto y(x) \in \mathbb{R}$ de classe $\mathcal{C}^1$ telles que pour tout $x \in \mathbb{R}$ on ait $\alpha_1(x, y(x))\, y'(x) + \alpha_2(x, y(x)) = 0$. Remarquons que s'il existe une application $f \in \mathcal{C}^1(\mathbb{R}^2, \mathbb{R})$ telle que

$$\alpha_1 = \frac{\partial f}{\partial x} \quad \text{et} \quad \alpha_2 = \frac{\partial f}{\partial y}$$

alors l'équation différentielle (10) est équivalente à $\frac{d}{dx}(f(x, y(x))) = 0$, c'est-à-dire équivalente à $f(x, y(x)) = C$ où C est une constante réelle. Autrement dit, si la 1-forme définie sur $\mathbb{R}^2$ par $\omega = \alpha_1 \mathrm{d}y + \alpha_2 \mathrm{d}x$ est une 1-forme exacte, alors les solutions de l'équation (10) vérifient une relation du type $f(x, y(x)) = C$ où C est une constante réelle. Illustrons ce résultat par un exemple.

Considérons l'équation différentielle suivante :

$$(3x^2y^2 - 2y)y' + 2xy^3 + 1 = 0 \qquad \forall x \in \mathbb{R} \tag{11}$$

et introduisons la 1-forme ω définie sur $\mathbb{R}^2$ par

$$\omega(x, y) = (2xy^3 + 1)\, \mathrm{d}x + (3x^2y^2 - 2y)\, \mathrm{d}y.$$

---

[34] Lors de la démonstration du théorème de Poincaré, l'application f est donnée par une intégrale mais cette expression n'est pas souvent exploitable en pratique.

[35] Voir l'exemple de la p. 957 pour un calcul similaire.

Cette 1-forme ω est fermée[36] et, puisque $\mathbb{R}^2$ est un ouvert étoilé, le théorème de Poincaré assure que la 1-forme ω est une 1-forme exacte : il existe donc une application $f \in \mathcal{C}^1(\mathbb{R}^2, \mathbb{R})$ telle que $\omega = \mathrm{d}f$. Pour tout $(x, y) \in \mathbb{R}^2$ on a

$$\begin{cases} \dfrac{\partial f}{\partial x}(x, y) = 2xy^3 + 1, \\ \dfrac{\partial f}{\partial x}(x, y) = 3x^2y^2 - 2y. \end{cases} \tag{12}$$

Pour que f vérifie la première de ces deux relations il faut et il suffit qu'il existe une application $g \in \mathcal{C}^1(\mathbb{R}, \mathbb{R})$ telle que pour tout $(x, y) \in \mathbb{R}^2$ on ait $f(x, y) = x^2y^3 + x + g(y)$. En reportant cette relation dans la seconde relation du système (12), on en déduit que pour tout $y \in \mathbb{R}$ on a $g'(y) = -2y$, c'est-à-dire $g(y) = k - y^2$ où k est une constante réelle. Les applications f telles que $\mathrm{d}f = \omega$ sont donc les applications $f : (x, y) \in \mathbb{R}^2 \mapsto x^2y^3 + x + k - y^2$ où k est une constante réelle. Si y est une solution de l'équation différentielle (11) alors on a $\mathrm{d}f = 0$. Cela implique que f est constante et donc que pour tout $x \in \mathbb{R}$ on a

$$x^2y(x)^3 + x + y(x)^2 = C \quad \text{où } C \text{ est une constante réelle.} \tag{13}$$

Cette méthode ne fournit pas de solution de l'équation différentielle mais permet d'obtenir des informations qualitatives sur d'éventuelles solutions. Ainsi, considérons une solution y de l'équation différentielle (11) telle que $y(0) = 0$. D'après la relation (13), pour tout $x \in \mathbb{R}$ on a $x^2y(x)^3 + x + y(x)^2 = 0$. On peut aussi établir que cette solution y ne s'annulera jamais ailleurs qu'en $x = 0$ puisque si $y(x) = 0$ alors $x^2y(x)^3 + x + y(x)^2 = x = 0$.

Exercice 13.5 Montrer que les solutions sur $\mathbb{R}$ de l'équation différentielle

$$3x^2y + (x^3 - \sin(y))y' = 0$$

vérifient, pour tout $x \in \mathbb{R}$, la relation $x^3y(x) + \cos(y(x)) = C$ où C est une constante réelle. En déduire que toute solution sur $\mathbb{R}$ de cette équation a pour limite 0 en $+\infty$ et en $-\infty$.

13.2.2 Utilisation du formalisme des 1-formes en physique

Étant donné une application f de $\mathbb{R}^n$ dans $\mathbb{R}$ de classe $\mathcal{C}^1$ et un vecteur $\mathbf{h} = (h_1, \ldots, h_n) \in \mathbb{R}^n$, notons $\delta f = f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x})$ un accroissement des valeurs prises par f autour d'un élément $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$. Il est important de remarquer que δf est un nombre réel qui, malgré sa notation, dépend de $\mathbf{x}$ et de $\mathbf{h}$. Considérons aussi, pour ces valeurs de $\mathbf{x}$ et de $\mathbf{h}$, les accroissements $\delta x_i = (x_i + h_i) - x_i = h_i$ pour $i \in \{1, \ldots n\}$.

[36] Voir l'exemple de la p. 960.

La formule de Taylor[37] à l'ordre 1 pour l'application f au voisinage de $\mathbf{x}$ s'écrit :

$$f(\mathbf{x}+\mathbf{h}) - f(\mathbf{x}) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(\mathbf{x})\, h_j + o_{\mathbf{0}}(\|\mathbf{h}\|).$$

Avec les notations précédentes, on en déduit que l'on peut approcher δf par

$$\boxed{\delta f \approx \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(\mathbf{x})\, \delta x_j.} \tag{14}$$

Dans cette approximation, l'erreur commise est de l'ordre de $\max_{i\in\{1,\dots,n\}} (\delta x_i^2)$.

On ne confondra pas la relation (14) avec la relation (4) obtenue à la p. 957. La relation (14) est une valeur approchée de l'écart entre deux nombres réels alors que la relation (4) est une égalité entre deux 1-formes.

Illustrons l'utilisation possible de la relation (14) par quelques exemples.

✗ Considérons un circuit électrique ayant une résistance ohmique R, une inductance propre L et un condensateur de capacité C. Lorsque la résistance R est suffisamment petite, ce circuit est responsable d'oscillations électriques amorties. Plus précisément, lorsque $R < 2\sqrt{L/C}$ la pseudo-période des oscillations s'exprime par :

$$T = 4\pi L \sqrt{\frac{C}{4L - CR^2}}.$$

Par exemple, pour $R = 120\ \Omega$, $L = 0,1\ H$ et $C = 2 \times 10^{-5}\ F$, des oscillations amorties de période $T \approx 16,79 \times 10^{-3}$ s se développent.

Un expérimentateur sait que la résistance, l'inductance et la capacité de son système électrique subissent les variations relatives suivantes :

$$\frac{\delta R}{R} = 10^{-2}, \quad \frac{\delta L}{L} = -10^{-2} \quad \text{et} \quad \frac{\delta C}{C} = 10^{-2},$$

et il veut en déduire la variation relative de la pseudo-période T. Définissons l'application

$$T : (L, C, R) \in U \longmapsto 4\pi L \sqrt{\frac{C}{4L - CR^2}}$$

où U est l'ouvert de $\mathbb{R}^3$ défini par

$$U = \Big\{ (L, C, R) \in \mathbb{R}^3 \ ; \ L > 0,\ C > 0,\ R > 0 \ \text{ et } \ 4L - CR^2 > 0 \Big\}.$$

Lors de l'exercice 13.4, il a été montré que pour tout $\mathbf{x} = (L, C, R) \in U$ on a

$$\frac{\mathrm{d}T_{\mathbf{x}}}{T(\mathbf{x})} = \left(\frac{1}{L} - \frac{2}{4L - CR^2}\right) \mathrm{d}L + \left(\frac{1}{2C} + \frac{R^2}{2(4L - CR^2)}\right) \mathrm{d}C + \frac{CR}{4L - CR^2} \mathrm{d}R.$$

[37] Voir le théorème 10.8 p. 684.

En utilisant les notations introduites au début de cette partie, on en déduit

$$\frac{\delta T}{T} \approx \left(1 - \frac{2L}{4L - CR^2}\right)\frac{\delta L}{L} + \left(\frac{1}{2} + \frac{CR^2}{2(4L - CR^2)}\right)\frac{\delta C}{C} + \frac{CR^2}{4L - CR^2}\frac{\delta R}{R},$$

ce qui fournit pour valeur approchée : $\dfrac{\delta T}{T} \approx 5,14 \times 10^{-2}$.

Cet exemple montre comment, à partir d'une loi physique (donnant ici la pseudo-période des oscillations en fonction des paramètres R, C et L), on peut obtenir des informations sur des petites variations. Inversement, à partir d'observations ou de mesures physiques, on peut obtenir une loi décrivant un phénomène physique. Illustrons cette approche par un exemple.

✗ Un physicien sait que pour de petites variations δV du volume et $\delta\theta$ de la température d'un gaz, la variation de pression δP obéit à la loi suivante ($\mathcal{R}$ étant une constante fixée) :

$$\delta P = -\frac{\mathcal{R}\theta}{V^2}\left(1 + \frac{4}{V}\right)\delta V + \frac{\mathcal{R}}{V}\left(1 + \frac{2}{V}\right)\delta\theta. \tag{15}$$

Il sait aussi que lorsque le volume devient très grand, la pression tend à devenir nulle. À partir de ces informations, il voudrait établir une relation $P = f(V, \theta)$ exprimant la valeur de la pression P en fonction du volume V et de la température θ. La relation (14) indique que la relation (15) exprime le fait que l'application f doit vérifier

$$\frac{\partial f}{\partial V}(V,\theta) = -\frac{\mathcal{R}\theta}{V^2}\left(1 + \frac{4}{V}\right) \quad \text{et} \quad \frac{\partial f}{\partial \theta}(V,\theta) = \frac{\mathcal{R}}{V}\left(1 + \frac{2}{V}\right). \tag{16}$$

De la deuxième relation dans (16), on déduit que f est de la forme

$$f(V,\theta) = \frac{\mathcal{R}\theta}{V}\left(1 + \frac{2}{V}\right) + g(V),$$

où g est une fonction de classe $\mathcal{C}^1$ sur $\mathbb{R}$. En reportant dans la première relation de (16), on en déduit que pour tout $V \in \mathbb{R}$ on a $g'(V) = 0$, donc que g est une constante. La valeur de cette constante C se calcule en utilisant l'hypothèse $\lim_{V\to+\infty} f(V,\theta) = 0$. On obtient $C = 0$ et on en conclut que la pression du gaz dépend du volume et de la température de la manière suivante :

$$f(V,\theta) = \frac{\mathcal{R}\theta}{V}\left(1 + \frac{2}{V}\right).$$

13.2.3 Circulation d'une 1-forme

Intéressons-nous maintenant à la façon d'intégrer une 1-forme. Considérons une application continue α d'un ouvert U de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^n$ et une courbe paramétrée C de $\mathbb{R}^n$ incluse dans U et paramétrée par une application $\varphi \in \mathcal{C}^1([a,b], U)$. Dans la section 13.1.1, nous avons défini l'intégrale curviligne $\int_C \alpha \cdot \mathrm{d}\mathbf{x}$ de l'application α le long de C de la façon suivante :

$$\int_C \alpha \cdot \mathrm{d}\mathbf{x} = \int_a^b \alpha(\varphi(t)) \cdot \varphi'(t)\ \mathrm{d}t.$$

À cette application $\alpha : \mathbf{x} \in U \mapsto (\alpha_1(\mathbf{x}), \ldots, \alpha_n(\mathbf{x})) \in \mathbb{R}^n$ on peut associer la 1-forme $\omega = \sum_{i=1}^n \alpha_i \mathrm{d}x_i$ et la proposition 13.6 indique que l'application qui à α associe ω est bijective. Cette observation conduit à définir l'intégrale d'une 1-forme de la manière suivante.

DÉFINITION 13.9 *Soient U un ouvert de $\mathbb{R}^n$, C une courbe paramétrée de $\mathbb{R}^n$ de classe $\mathcal{C}^1$ incluse dans U et $\omega = \sum_{i=1}^n \alpha_i \mathrm{d}x_i$ une 1-forme continue sur U. On appelle* circulation *de la 1-forme ω le long de la courbe C, le réel*

$$\int_C \omega = \int_C \alpha \cdot \mathrm{d}\mathbf{x},$$

où $\int_C \alpha \cdot \mathrm{d}\mathbf{x}$ désigne l'intégrale curviligne de $\alpha : \mathbf{x} \in U \mapsto (\alpha_1(\mathbf{x}), \ldots, \alpha_n(\mathbf{x}))$ le long de la courbe C.

Exemples

1. Considérons la 1-forme $\omega : (x_1, x_2) \in \mathbb{R}^2 \mapsto x_1 \, \mathrm{d}x_1 + x_2 \, \mathrm{d}x_2$, et calculons la circulation de cette 1-forme le long de l'ellipse C paramétrée par l'application $\varphi : t \in [0, 2\pi] \mapsto (2\cos(t), \sin(t))$ qui est de classe $\mathcal{C}^1$. La 1-forme ω s'écrit $\omega = \alpha_1 \, \mathrm{d}x_1 + \alpha_2 \, \mathrm{d}x_2$ où $\alpha_1 : (x_1, x_2) \in \mathbb{R}^2 \mapsto x_1$ et $\alpha_2 : (x_1, x_2) \in \mathbb{R}^2 \mapsto x_2$ et en notant α l'application de $\mathbb{R}^2$ dans $\mathbb{R}^2$ dont les composantes sont α_1 et α_2, on obtient :

$$\begin{aligned}\int_C \omega = \int_C \alpha \cdot \mathrm{d}\mathbf{x} &= \int_0^{2\pi} \alpha(\varphi(t)) \cdot \varphi'(t) \, \mathrm{d}t = \int_0^{2\pi} \begin{pmatrix} \alpha_1(\varphi(t)) \\ \alpha_2(\varphi(t)) \end{pmatrix} \cdot \begin{pmatrix} -2\sin(t) \\ \cos(t) \end{pmatrix} \mathrm{d}t \\ &= \int_0^{2\pi} \begin{pmatrix} 2\cos(t) \\ \sin(t) \end{pmatrix} \cdot \begin{pmatrix} -2\sin(t) \\ \cos(t) \end{pmatrix} \mathrm{d}t = \int_0^{2\pi} -3\cos(t)\sin(t) \, \mathrm{d}t \\ &= \int_0^{2\pi} -\frac{3}{2}\sin(2t) \, \mathrm{d}t = \left[\frac{3}{4}\cos(2t)\right]_0^{2\pi} = 0.\end{aligned}$$

2. Considérons la 1-forme définie sur $\mathbb{R}^2_*$ par

$$\omega : (x_1, x_2) \in \mathbb{R}^2_* \longmapsto \frac{-x_2}{x_1^2 + x_2^2} \, \mathrm{d}x_1 + \frac{x_1}{x_1^2 + x_2^2} \, \mathrm{d}x_2,$$

et calculons la circulation de cette 1-forme le long du cercle C de centre $\mathbf{0}$ et de rayon 1. Notons que le cercle C est une courbe incluse dans l'ouvert $\mathbb{R}^2_*$ qui peut être paramétrée par l'application $\varphi : t \in [0, 2\pi] \mapsto (\cos(t), \sin(t))$ qui est de classe $\mathcal{C}^1$. La 1-forme ω s'écrit $\omega = \alpha_1 \, \mathrm{d}x_1 + \alpha_2 \, \mathrm{d}x_2$ où

$$\alpha_1 : (x_1, x_2) \in \mathbb{R}^2_* \longmapsto -\frac{x_2}{x_1^2 + x_2^2} \quad \text{et} \quad \alpha_2 : (x_1, x_2) \in \mathbb{R}^2_* \longmapsto \frac{x_1}{x_1^2 + x_2^2}$$

et en notant α l'application de $\mathbb{R}^2_*$ dans $\mathbb{R}^2$ dont les composantes sont α_1 et α_2, on obtient

$$\int_C \omega = \int_C \alpha \cdot \mathrm{d}\mathbf{x} = \int_0^{2\pi} \alpha(\varphi(t)) \cdot \varphi'(t) \, \mathrm{d}t = \int_0^{2\pi} \begin{pmatrix} \alpha_1(\varphi(t)) \\ \alpha_2(\varphi(t)) \end{pmatrix} \cdot \begin{pmatrix} -\sin(t) \\ \cos(t) \end{pmatrix} \mathrm{d}t.$$

Pour tout $t \in [0, 2\pi]$ on a $\alpha_1(\varphi(t)) = -\sin(t)$ et $\alpha_2(\varphi(t)) = \cos(t)$ de sorte que

$$\int_C \omega = \int_0^{2\pi} \left(\sin(t)^2 + \cos(t)^2\right)\, \mathrm{d}t = \int_0^{2\pi} 1\, \mathrm{d}t = 2\pi.$$

Les notions de formes différentielles fermées et exactes, et la distinction faite entre elles peuvent sembler assez abstraites. Le rôle particulier joué par les 1-formes exactes et fermées s'éclaire lorsque l'on s'intéresse à la circulation de 1-formes le long d'une courbe de $\mathbb{R}^n$. Si ω est une 1-forme exacte sur un ouvert U de $\mathbb{R}^n$ alors $\omega = \mathrm{d}f$ où f est une application de U dans $\mathbb{R}$. Cette 1-forme a pour décomposition dans la famille $\{\mathrm{d}x_1, \ldots, \mathrm{d}x_n\}$:

$$\mathrm{d}f = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}\, \mathrm{d}x_j = \nabla f \cdot \mathrm{d}\mathbf{x},$$

de sorte que la circulation de $\mathrm{d}f$ le long d'une courbe C correspond à l'intégrale de l'application ∇f le long de C. On déduit de la proposition 13.2 le résultat suivant.

PROPOSITION 13.8 *Soient U un ouvert de $\mathbb{R}^n$, C une courbe paramétrée de $\mathbb{R}^n$ de classe $\mathcal{C}^1$, incluse dans U, d'extrémités A et B, et f une application de U dans $\mathbb{R}^n$ de classe $\mathcal{C}^1$. On a*

$$\int_C \mathrm{d}f = f(B) - f(A).$$

La proposition 13.8 indique que si ω est une 1-forme exacte alors sa circulation le long d'une courbe paramétrée C de classe $\mathcal{C}^1$ ne dépend que des extrémités de la courbe. En particulier la circulation de cette 1-forme exacte ω sur une autre courbe paramétrée de classe $\mathcal{C}^1$ ayant les mêmes extrémités est égale à la circulation de ω le long de C.

Les lois de la physique affirment qu'une particule de charge élémentaire e immergée dans un champ électrique $\mathbf{E}$ subit une force $\mathbf{F}$ (force dite électrostatique ou force de Coulomb) donnée par la relation $\mathbf{F} = e\mathbf{E}$. Le champ électrique $\mathbf{E}$ « dérive d'un potentiel » ce qui signifie qu'il existe une application V telle que $\mathbf{E} = \nabla V$. L'application V est appelé le potentiel électrique et s'exprime en volts. En vertu de la proposition 13.8, on en déduit que le travail de la force électrostatique ne dépend pas du trajet de la particule.

Remarques

1. On peut symboliquement écrire la relation $\int_C \mathrm{d}f = f(B) - f(A)$ sous la forme

$$\int_C \mathrm{d}f = \int_{\partial C} f \tag{17}$$

où ∂C désigne les extrémités de C qui correspondent ici aux deux points A et B. On verra ultérieurement qu'une telle notation est justifiée par des résultats plus généraux.

2. Si U est un ouvert de $\mathbb{R}$ et C le segment $[a, b]$ alors la relation (17) s'écrit

$$\int_a^b f'(x)\,\mathrm{d}x = f(b) - f(a).$$

Cette relation permet de relier une donnée sur un intervalle à une donnée sur le bord de cet intervalle. On peut en donner différentes interprétations en physique : si f représente la température le long d'un fil métallique alors f' représente la variation de température ; si f' représente la densité d'une barre en fonction de la position alors $f(b)-f(a)$ représente le poids total de la barre, etc. Il est intéressant de remarquer que dans ces exemples, f et f' sont exprimées dans des unités différentes ; ainsi un changement de paramétrage modifiera la valeur de f' mais pas celle de f. Un changement de paramétrage fréquent est un changement d'unité. Par exemple, le poids d'une barre ne dépend pas de l'unité de longueur choisie mais il n'en est pas de même pour la densité : on n'obtient pas le même résultat selon qu'on l'exprime en $\mathrm{g}\cdot\mathrm{m}^{-1}$ ou en $\mathrm{g}\cdot\mathrm{cm}^{-1}$. La raison profonde de ces différences est que ce n'est pas l'objet f' qui est adapté à ce type de calcul mais l'objet $f'\,\mathrm{d}x$ qui est une forme différentielle.

3. Si on change l'orientation de la courbe (par exemple en choisissant un paramétrage équivalent mais non $\mathcal{O}$-équivalent) alors la valeur de $\int_C \mathrm{d}f$ change de signe : $\int_C \mathrm{d}f = f(\mathbf{A}) - f(\mathbf{B})$. Rigoureusement, dans la formule (17) on doit orienter le bord ∂C en fonction de l'orientation de C : orienter ∂C revient ici à affecter les valeurs $+1$ ou -1 aux deux extrémités. ◇

La proposition 13.8 fournit un moyen d'établir qu'une 1-forme n'est pas exacte. En effet, si une 1-forme $\omega \in \Lambda^1(U)$ est exacte alors il existe une application $f \in \mathcal{C}(U, \mathbb{R})$ telle que $\omega = \mathrm{d}f$ et par conséquent la circulation de ω le long de toute courbe fermée incluse dans U doit être nulle.

Corollaire 13.1 *Soient U un ouvert de $\mathbb{R}^n$ et ω une 1-forme sur U. Si il existe une courbe fermée C incluse dans U pour laquelle $\int_C \omega \neq 0$ alors ω n'est pas exacte.*

Exemples

1. Considérons sur l'ouvert $U = \mathbb{R}^2_*$ la 1-forme définie par

$$\omega(x_1, x_2) = \frac{-x_2}{x_1^2 + x_2^2}\,\mathrm{d}x_1 + \frac{x_1}{x_1^2 + x_2^2}\,\mathrm{d}x_2.$$

D'après l'exemple de la p. 969, la circulation de ω le long du cercle C de centre $(0, 0)$ et de rayon 1 (qui est une courbe fermée incluse dans U) n'est pas nulle. Par conséquent la 1-forme ω n'est pas exacte[38].

2. La 1-forme $\omega : (x_1, x_2) \in \mathbb{R}^2 \mapsto x_1\,\mathrm{d}x_1 + x_2\,\mathrm{d}x_2$ est une 1-forme exacte puisqu'on a $\omega = \mathrm{d}f$ où f est l'application $f : (x_1, x_2) \in \mathbb{R}^2 \mapsto \frac{1}{2}(x_1^2 + x_2^2)$. D'après la réciproque du corollaire 13.1, on en déduit que pour toute courbe fermée de $\mathbb{R}^2$ on aura $\int_C \omega = 0$. Dans le cas où la courbe C est l'ellipse dont un paramétrage est $\varphi : t \in [0, 2\pi] \mapsto (2\cos(t), \sin(t))$, on retrouve le résultat obtenu p. 969.

La proposition 13.8 admet également le corollaire suivant.

Corollaire 13.2 *Soient U un ouvert* étoilé *de $\mathbb{R}^n$, C une courbe* fermée *de $\mathbb{R}^n$ incluse dans U et ω une 1-forme fermée sur U. On a*

$$\int_C \omega = 0.$$

Démonstration D'après le théorème de Poincaré, U étant un ouvert étoilé, on sait que la forme fermée ω est une forme exacte. Par conséquent il existe une application f telle que $\omega = \mathrm{d}f$. La courbe C est fermée donc d'après la proposition 13.8, on en déduit que $\int_C \omega = \int_C \mathrm{d}f = 0$. □

Exemple Considérons à nouveau sur l'ouvert $U = \mathbb{R}^2_*$ la 1-forme

$$\omega(x_1, x_2) = \frac{-x_2}{x_1^2 + x_2^2}\,\mathrm{d}x_1 + \frac{x_1}{x_1^2 + x_2^2}\,\mathrm{d}x_2.$$

On a vu p. 960 que cette 1-forme est fermée. L'ouvert U n'étant pas un ouvert étoilé, on ne peut pas appliquer le corollaire 13.2. Par contre si on restreint la 1-forme ω à un ouvert V étoilé inclus dans U alors le corollaire 13.2 indique que $\int_C \omega = 0$ pour toute courbe C incluse dans V. Considérons par exemple le cercle de centre $(2, 0)$ et de rayon 1. Ce cercle est contenu dans l'ouvert $V = \{(x_1, x_2) \in \mathbb{R}^2 \ ; \ x_1 \neq 0\}$ qui est un ouvert étoilé inclus dans U. Ainsi $\int_C \omega = 0$.

(38) Il existe de très nombreux travaux en mathématiques, connus sous le nom de cohomologie de De Rham (Georges De Rham (1903-1990) était un mathématicien suisse), dont l'un des buts est de déterminer quelles sont les 1-formes fermées qui ne sont pas exactes. Ainsi, le théorème de Poincaré est le premier ingrédient de cette théorie puisqu'il indique que toutes les 1-formes sur un ouvert U étoilé qui sont fermées sont exactes. La théorie de la cohomologie de De Rham montre par exemple que sur l'ouvert $U = \mathbb{R}^2_*$ les 1-formes fermées et non exactes sont toutes du type $\lambda\omega + \widetilde{\omega}$, où λ est un réel quelconque, $\omega(x_1, x_2) = \frac{-x_2}{x_1^2+x_2^2}\mathrm{d}x_1 + \frac{x_1}{x_1^2+x_2^2}\mathrm{d}x_2$ et $\widetilde{\omega}$ une 1-forme exacte quelconque.

13.3 Intégrale de surface

13.3.1 Intégrer sur une surface

On se donne une application continue f d'un ouvert U de $\mathbb{R}^n$ ($n \geqslant 2$) à valeurs dans $\mathbb{R}^n$ (une telle application est qualifiée de champ de vecteurs sur U) et S une surface paramétrée de classe $\mathcal{C}^1$ de $\mathbb{R}^n$ représentée par une nappe φ de classe $\mathcal{C}^1$ sur un sous-ensemble K de $\mathbb{R}^2$ à valeurs dans U.

En chaque point $M = \varphi(u,v)$ de la surface S, le vecteur $f(\varphi(u,v)) \in \mathbb{R}^n$ correspond à la valeur prise par l'application f en $M \in U$. Lorsque le produit vectoriel des deux vecteurs $\frac{\partial \varphi}{\partial u}(u,v)$ et $\frac{\partial \varphi}{\partial v}(u,v)$ est non nul, il permet de définir[39] un vecteur normal $\mathbf{N}(u,v)$ à la surface S au point M. On peut ainsi considérer l'application

$$\xi : (u,v) \in K \longmapsto f(\varphi(u,v)) \cdot \mathbf{N}(u,v) \in \mathbb{R},$$

qui à un couple de paramètres (u,v) associe le produit scalaire des deux vecteurs $f(\varphi(u,v))$ et $\mathbf{N}(u,v)$. Si on suppose de plus que la surface S est orientable alors l'application ξ de K dans $\mathbb{R}$ est continue[40] et par conséquent Riemann-intégrable sur K. Ces considérations conduisent à introduire la définition suivante.

Définition 13.10 (Intégrale de surface)
Soient U un ouvert de $\mathbb{R}^n$, f une application de U dans $\mathbb{R}^n$ continue et S une surface paramétrée, orientée, de classe $\mathcal{C}^1$, dont un paramétrage est $\varphi \in \mathcal{C}^1(K,U)$ où K désigne un sous-ensemble de $\mathbb{R}^2$. On appelle intégrale de f sur la surface S *et on note $\int_S f \cdot \mathrm{d}\mathbf{s}$ l'intégrale*

$$\int_K f(\varphi(u,v)) \cdot \mathbf{N}(u,v)\, \mathrm{d}u\mathrm{d}v$$

où $\mathbf{N}(u,v) = \dfrac{\partial \varphi}{\partial u}(u,v) \wedge \dfrac{\partial \varphi}{\partial v}(u,v)$.

En physique, la notion d'intégrale de surface est par exemple utilisée pour définir le flux d'une force à travers une surface. Par exemple, la puissance thermique est une notion attachée au flux de chaleur à travers une surface. Plus précisément, le transfert de chaleur d'une région de température élevée vers une région de température plus basse obéit à la loi dite de Fourier[41]. Cette loi exprime le fait que la densité J de flux de chaleur est proportionnelle au gradient de température T, c'est-à-dire $J = \lambda \nabla T$ où la constante λ est appelée la conductivité thermique du milieu. La puissance thermique P à travers une surface S est définie par le flux du vecteur J à travers la surface S, c'est-à-dire : $P = \int_S J \cdot \mathrm{d}\mathbf{s}$.

---

[39] Voir la p. 731.

[40] En effet, lorsqu'une surface est orientable on peut définir un vecteur normal $\mathbf{N}$ en chaque point M de façon à ce que l'application qui à M associe $\mathbf{N}$ soit continue, voir p. 800.

La proposition suivante indique que la valeur de l'intégrale de f sur une surface ne dépend pas du choix du paramétrage de la surface choisi pour effectuer le calcul.

PROPOSITION 13.9 *Soient f une application continue d'un ouvert U de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^n$ et S une surface paramétrée orientable de classe $\mathcal{C}^1$ contenue dans U. La valeur de l'intégrale de surface $\int_S f \cdot \mathrm{d}\mathbf{s}$ est invariante par changement de paramétrage $\mathcal{O}^1$-équivalent de la surface S.*

Démonstration Soient $\varphi \in \mathcal{C}^1(K, \mathbb{R}^n)$ et $\psi \in \mathcal{C}^1(H, \mathbb{R}^n)$, où H et K désignent deux ensembles de $\mathbb{R}^2$, des paramétrages $\mathcal{O}^1$-équivalents de la surface S. D'après la définition 11.18 p. 790, il existe un $\mathcal{C}^1$-difféomorphisme θ de K dans H dont le jacobien $\mathcal{J}_\theta$ est strictement positif et tel que $\varphi = \psi \circ \theta$.
Pour $(u, v) \in K$ notons $\mathbf{N}_\varphi(u, v)$ le vecteur normal à la surface au point $\varphi(u, v)$ correspondant au paramétrage φ, et pour $(s, t) \in H$ notons $\mathbf{N}_\psi(s, t)$ le vecteur normal au point $\psi(s, t)$ correspondant au paramétrage ψ ; on a

$$\mathbf{N}_\varphi(u, v) = \frac{\partial \varphi}{\partial u}(u, v) \wedge \frac{\partial \varphi}{\partial v}(u, v) \tag{18}$$

et

$$\mathbf{N}_\psi(s, t) = \frac{\partial \psi}{\partial s}(s, t) \wedge \frac{\partial \psi}{\partial t}(s, t). \tag{19}$$

Notons θ_1 et θ_2 les deux composantes de l'application θ. Pour tout $(u, v) \in K$ on a $\varphi(u, v) = \psi(\theta(u, v)) = \psi(\theta_1(u, v), \theta_2(u, v))$ et en utilisant la formule de dérivation des applications composées[42], on obtient

$$\frac{\partial \varphi}{\partial u}(u, v) = \frac{\partial \theta_1}{\partial u}(u, v) \frac{\partial \psi}{\partial s}(\theta(u, v)) + \frac{\partial \theta_2}{\partial u}(u, v) \frac{\partial \psi}{\partial t}(\theta(u, v)),$$

$$\frac{\partial \varphi}{\partial v}(u, v) = \frac{\partial \theta_1}{\partial v}(u, v) \frac{\partial \psi}{\partial s}(\theta(u, v)) + \frac{\partial \theta_2}{\partial v}(u, v) \frac{\partial \psi}{\partial t}(\theta(u, v)).$$

En substituant ces relations dans la relation (18), on obtient

$$\mathbf{N}_\varphi(u, v) = \left(\frac{\partial \theta_1}{\partial u}(u, v) \frac{\partial \theta_2}{\partial v}(u, v) - \frac{\partial \theta_2}{\partial u}(u, v) \frac{\partial \theta_1}{\partial v}(u, v) \right) \times \left(\frac{\partial \psi}{\partial s}(\theta(u, v)) \wedge \frac{\partial \psi}{\partial t}(\theta(u, v)) \right)$$

On peut réécrire cette égalité en utilisant la relation (19) de la façon suivante :

$$\mathbf{N}_\varphi(u, v) = \det \begin{pmatrix} \dfrac{\partial \theta_1}{\partial u}(u, v) & \dfrac{\partial \theta_1}{\partial v}(u, v) \\ \dfrac{\partial \theta_2}{\partial u}(u, v) & \dfrac{\partial \theta_2}{\partial v}(u, v) \end{pmatrix} \mathbf{N}_\psi(\theta(u, v)) = \mathcal{J}_\theta(u, v)\, \mathbf{N}_\psi(\theta(u, v)).$$

[41] Cette loi a été établie mathématiquement en 1804 par Jean-Baptiste Biot (1774 - 1862) puis expérimentalement par Joseph Fourier en 1822.

[42] Voir la proposition 10.10 p. 638.

On a donc

$$\int_K f(\varphi(u,v)) \cdot \mathbf{N}_\varphi(u,v)\ du dv = \int_K f(\psi(\theta(u,v))) \cdot \mathbf{N}_\psi(\theta(u,v))\, \mathcal{J}_\theta(u,v)\ du dv$$

et en appliquant la formule du changement de variable (voir le théorème 12.2 p. 894), on obtient :

$$\int_K f(\varphi(u,v)) \cdot \mathbf{N}_\varphi(u,v)\ du dv = \int_H f(\psi(s,t)) \cdot \mathbf{N}_\psi(s,t)\ ds dt.$$

La valeur de l'intégrale de f sur la surface S est donc la même quel que soit le paramétrage $\mathcal{O}^1$-équivalent choisi pour S. □

Remarque Comme lors de l'étude des intégrales curvilignes, l'orientation intervient implicitement dans la définition de l'intégrale de surface. Pour une surface orientable S donnée, si on considère deux paramétrages φ et ψ qui sont $\mathcal{C}^1$-équivalent mais non $\mathcal{O}^1$-équivalent (autrement dit, ils représentent la même surface mais l'orientent de façon différente) alors[43] la valeur de l'intégrale sur S d'une fonction f donnée aura un signe opposé selon qu'elle est calculée avec le paramétrage φ ou avec le paramétrage ψ.

Cela implique en particulier qu'on ne peut pas parler de l'intégrale d'une fonction f sur une surface orientable sans préciser l'orientation choisie. Pour préciser une orientation, soit on indique explicitement le paramétrage choisi, soit on donne le sens du vecteur normal à la surface[44], par exemple en imposant que le vecteur normal soit dirigé vers l'extérieur. Compte tenu des difficultés inhérentes à l'orientation des surfaces, la notion de normale extérieure pour une surface sera, au niveau de ce cours, purement intuitive.

Dans les applications physiques, l'orientation de la surface est généralement liée au problème considéré : on cherche par exemple un flux « entrant » à travers une surface ou un flux « sortant ». En toute généralité, seul le cas où la surface sépare l'espace en exactement deux composantes connexes (comme le fait par exemple la sphère) permet de définir assez simplement la notion de normale extérieure. En pratique, on se contente d'une vision intuitive de cette notion de normale extérieure pour les applications visées. ◇

Exemples

1. Calculons le flux de $f : (x_1, x_2, x_3) \in \mathbb{R}^3 \mapsto (x_1^2 + x_2^2, x_3, x_1)$ à travers le demi-cylindre S dont un paramétrage est

$$\varphi : (u,v) \in [0, \pi/2] \times [0,h] \longmapsto \big(a\cos(u), a\sin(u), v\big)\,,$$

a et h étant deux nombres réels strictement positifs décrivant respectivement le rayon et la hauteur du cylindre. Ce flux est donné par $F = \int_S f \cdot \mathbf{ds}$.

---

[43] Pour s'en convaincre, il suffit de refaire la démonstration précédente en utilisant un changement de variables de jacobien négatif.

[44] Voir à ce sujet la remarque 11.4.3 p. 800.

Le vecteur normal à S en un point $M = \varphi(u,v)$ de S est

$$\mathbf{N}(u,v) = \frac{\partial\varphi}{\partial u}(u,v) \wedge \frac{\partial\varphi}{\partial u}(u,v) = \begin{pmatrix} -a\sin(u) \\ a\cos(u) \\ 0 \end{pmatrix} \wedge \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} a\cos(u) \\ a\sin(u) \\ 0 \end{pmatrix}.$$

En un point de la surface paramétré par $(u,v) \in [0,\pi/2] \times [0,h]$, on a

$$f(\varphi(u,v)) = \big(a^2, v, a\cos(u)\big)$$

et le produit scalaire $f(\varphi(u,v))\cdot\mathbf{N}(u,v)$ vaut $a^3\cos(u)+av\sin(u)$. On en déduit que le flux de f à travers S vaut

$$\begin{aligned} F &= \int_0^{\pi/2} \left(\int_0^h (a^3\cos(u) + av\sin(u))\, \mathrm{d}v \right) \mathrm{d}u \\ &= \int_0^{\pi/2} \Big(a^3 h\cos(u) + \frac{a}{2}h^2\sin(u) \Big) \mathrm{d}u = a^3 h + \frac{a}{2}h^2. \end{aligned}$$

2. Soit S la demi-sphère de $\mathbb{R}^3$ de centre $\mathbf{0}$ et de rayon $R \in \mathbb{R}_+^*$ dont un paramétrage est[45]

$$\varphi : (u,v) \in [0,2\pi] \times [0,\pi/2] \longmapsto \big(R\cos(u)\sin(v), R\sin(u)\sin(v), R\cos(v)\big).$$

Calculons le flux sortant (c'est-à-dire dirigé vers l'extérieur de la sphère) de S de l'application (k désigne un réel strictement positif fixé)

$$f : (x_1,x_2,x_3) \in \mathbb{R}^3 \longmapsto (kx_1, kx_2, kx_3).$$

Pour ce paramétrage de la demi-sphère, un vecteur normal à S en un point $\varphi(u,v)$, $(u,v) \in [0,2\pi] \times [0,\pi/2]$ est donné par

$$\begin{aligned} \mathbf{N}(u,v) &= \frac{\partial\varphi}{\partial u}(u,v) \wedge \frac{\partial\varphi}{\partial u}(u,v) \\ &= \begin{pmatrix} -R\sin(u)\sin(v) \\ R\cos(u)\sin(v) \\ 0 \end{pmatrix} \wedge \begin{pmatrix} R\cos(u)\cos(v) \\ R\sin(u)\cos(v) \\ -R\sin(v) \end{pmatrix} = \begin{pmatrix} -R^2\cos(u)\sin(v)^2 \\ -R^2\sin(u)\sin(v)^2 \\ -R^2\sin(v)\cos(v) \end{pmatrix}. \end{aligned}$$

Notons que ce vecteur normal est dirigé vers l'intérieur de la sphère. On le voit en considérant par exemple le vecteur normal au point $\varphi(0,\pi/2) = (R,0,0)$ qui vaut $\mathbf{N}(0,\pi/2) = (-R^2,0,0)$. Pour calculer le flux sortant on utilisera par conséquent le vecteur normal $-\mathbf{N}$.

[45] Voir la partie 11.1.5 p. 724.

On a $f(\varphi(u,v)) = \big(kR\cos(u)\sin(v), kR\sin(u)\sin(v), kR\cos(v)\big)$ et par conséquent le flux de f à travers la demi-sphère S vaut

$$\begin{aligned}\int_S f\cdot \mathrm{d}\mathbf{s} &= -\int_{[0,2\pi]\times[0,\pi/2]} f(\varphi(u,v))\cdot \mathbf{N}(u,v)\ \mathrm{d}u\mathrm{d}v\\ &= \int_0^{2\pi}\left(\int_0^{\pi/2} kR^3(\sin(v)^3+\sin(v)\cos(v)^2)\ \mathrm{d}v\right)\mathrm{d}u\\ &= 2\pi kR^3\int_0^{\pi/2}\sin(v)\ \mathrm{d}v = 2\pi kR^3.\end{aligned}$$

Exercice 13.6 Considérons un morceau S de l'escalier hélicoïdal de hauteur h dont un paramétrage est[46]

$$\varphi : (t,\lambda)\in[0,h]\times[0,1]\longmapsto \big(\lambda\cos(t),\lambda\sin(t),t\big)\in\mathbb{R}^3.$$

Pour $(t,\lambda)\in[0,h]\times[0,1]$, on note $\mathbf{N}(t,\lambda)$ le vecteur $\frac{\partial\varphi}{\partial\lambda}(t,\lambda)\wedge\frac{\partial\varphi}{\partial t}(t,\lambda)$. Vérifier que ce vecteur n'est jamais le vecteur nul, puis calculer l'intégrale de $\frac{\mathbf{N}}{\|\mathbf{N}\|}$ à travers la surface S.

Propriétés de l'intégrale de surface

L'intégrale de surface hérite des propriétés de l'intégrale multiple, en particulier celle de linéarité. Considérons une surface paramétrée S, orientable, de classe $\mathcal{C}^1$ contenue dans un ouvert U de $\mathbb{R}^n$; pour tout $(\lambda,\mu)\in\mathbb{R}^2$ et pour tout $(f,g)\in\mathcal{C}(U,\mathbb{R}^n)^2$, on a

$$\int_S(\lambda f+\mu g)\cdot\mathrm{d}\mathbf{s} = \lambda\int_S f\cdot\mathrm{d}\mathbf{s}+\mu\int_S g\cdot\mathrm{d}\mathbf{s}.$$

Par ailleurs, soit $\{S_1,\dots,S_N\}$ où $N\in\mathbb{N}^*$ une partition[47] de la surface S et f une fonction de U dans $\mathbb{R}^n$ continue. On a la relation

$$\int_S f\cdot\mathrm{d}\mathbf{s} = \int_{S_1} f\cdot\mathrm{d}\mathbf{s}+\cdots+\int_{S_N} f\cdot\mathrm{d}\mathbf{s}$$

appelée *relation de Chasles*. Dans le cas où f est seulement continue par morceaux sur S ou dans le cas où la surface S est de classe $\mathcal{C}^1$ par morceaux, la relation de Chasles permet de donner un sens à l'intégrale $\int_S f\cdot\mathrm{d}\mathbf{s}$. Il suffit pour cela de considérer une partition de S telle que f soit continue sur chaque morceau S_i de la partition et telle que chaque morceau S_i soit de classe $\mathcal{C}^1$.

[46] Cette surface a été paramétrée à la p. 788 et étudiée à la p. 793.

[47] La définition d'une partition d'un ensemble est donnée à la p. 873.

Intégrale de surface d'une fonction de $\mathbb{R}^n$ dans $\mathbb{R}$

Nous avons défini l'intégrale de surface pour des fonctions de $\mathbb{R}^n$ dans $\mathbb{R}^n$. Comme pour l'intégrale curviligne, nous allons étendre la notion d'intégrale de surface aux fonctions à valeurs dans $\mathbb{R}$.

DÉFINITION 13.11 *Soient f une application continue d'un ouvert U de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}$ et S une surface paramétrée, orientable, de classe C^1, dont un paramétrage est $\varphi \in C^1(K,U)$ où K désigne un sous-ensemble de $\mathbb{R}^2$. On appelle* intégrale *de f à travers la surface S et on note $\int_S f \, \mathrm{d}\sigma$ l'intégrale*

$$\int_S f \, \mathrm{d}\sigma = \int_K f(\varphi(u,v)) \, \|\mathbf{N}(u,v)\| \, \mathrm{d}u\mathrm{d}v,$$

où $\mathbf{N}(u,v) = \dfrac{\partial \varphi}{\partial u}(u,v) \wedge \dfrac{\partial \varphi}{\partial v}(u,v)$ désigne le vecteur normal à la surface S au point $\varphi(u,v)$.

Cette notion d'intégrale surfacique permet par exemple de calculer la masse d'une plaque supposée extrêmement fine épousant la forme d'une surface S. Si cette plaque a pour densité surfacique de masse $\rho(x_1,x_2,x_3)$ en un point (x_1,x_2,x_3) de la surface alors la masse totale de la plaque est donnée par

$$\mathrm{M} = \int_S \rho \, \mathrm{d}\sigma.$$

De la même façon, les coordonnées $(x_{G,1},x_{G,2},x_{G,3})$ du centre de gravité $\mathbf{G}$ d'une telle plaque sont données pour $i \in \{1,2,3\}$ par[48]

$$x_{G,i} = \frac{1}{M} \int_S x_i \, \rho(x_1,x_2,x_3) \, \mathrm{d}\sigma, \quad i \in \{1,2,3\}$$

Exemples

1. Calculons l'intégrale de $f : (x_1,x_2,x_3) \in \mathbb{R}^3 \mapsto x_1^2 + x_2^2 + x_3^2$ sur la surface S qui est la paroi latérale du cylindre de génératrice l'axe $(0,\mathbf{e}_3)$, de hauteur $H \in \mathbb{R}_+^*$ et de rayon $R \in \mathbb{R}_+^*$. Un paramétrage de la surface S est

$$\varphi : (\theta,z) \in [0,2\pi] \times [0,H] \longmapsto (R\cos(\theta), R\sin(\theta), z).$$

Pour $(\theta,z) \in [0,2\pi] \times [0,H]$, le vecteur normal en $\varphi(\theta,z)$ est donné par

$$\mathbf{N}(\theta,z) = \begin{pmatrix} -R\sin(\theta) \\ R\cos(\theta) \\ 0 \end{pmatrix} \wedge \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} R\cos(\theta) \\ R\sin(\theta) \\ 0 \end{pmatrix}$$

[48] Par abus de notation, on note ici $\int_S x_i \, \rho(x_1,x_2,x_3) \, \mathrm{d}\sigma$ l'intégrale de l'application $(x_1,x_2,x_3) \mapsto x_i \, \rho(x_1,x_2,x_3)$ à travers la surface S.

et a pour norme $\|\mathbf{N}(\theta,z)\| = \sqrt{R^2\cos(\theta)^2 + R^2\sin(\theta)^2} = R$. On en déduit que

$$\begin{aligned}\int_S f\,\mathrm{d}\sigma &= \int_0^H \int_0^{2\pi} (R^2\cos(\theta)^2 + R^2\sin(\theta)^2 + z^2)R\,\mathrm{d}\theta\mathrm{d}z \\ &= R\int_0^H \int_0^{2\pi} (R^2+z^2)\,\mathrm{d}\theta\mathrm{d}z \\ &= 2\pi R\int_0^H (R^2+z^2)\,\mathrm{d}z = 2\pi R\left(R^2H + \frac{1}{3}H^3\right).\end{aligned}$$

2. Calculons la masse d'une plaque métallique de forme circulaire de rayon R qui épouse la forme de la surface dont une représentation paramétrique est

$$\varphi : (r,\theta) \in [0,R]\times[0,2\pi] \longmapsto (r\cos(\theta), r\sin(\theta), 0)$$

et dont la densité surfacique de masse ρ est proportionnelle à la distance au centre de la plaque :

$$\rho : (x_1,x_2,x_3) \in \mathbb{R}^3 \longmapsto k\sqrt{x_1^2+x_2^2+x_3^2} \qquad \text{avec } k \in \mathbb{R}_+^*.$$

Pour $(r,\theta) \in [0,R]\times[0,2\pi]$, le vecteur normal à la plaque au point $\varphi(r,\theta)$ vaut $\mathbf{N}(r,\theta) = (0,0,r)$. On en déduit que la masse M cherchée vaut

$$\begin{aligned}M &= \int_S \rho\,\mathrm{d}\sigma = \int_0^R \int_0^{2\pi} \rho(r\cos(\theta), r\sin(\theta), 0) r\,\mathrm{d}r\mathrm{d}\theta \\ &= \int_0^R \int_0^{2\pi} kr^2\,\mathrm{d}r\mathrm{d}\theta = \frac{2\pi}{3}kR^3.\end{aligned}$$

Cet exemple illustre le fait que lorsqu'une surface S appartient à un plan (ce qui est le cas du disque représentant la plaque métallique dans cet exemple), l'intégrale de surface d'une application f coïncide avec l'intégrale double de f.

13.3.2 Aire d'une surface

Nous avons vu à la p. 874 que l'aire d'un ensemble D du plan $\mathbb{R}^2$ peut être obtenue en intégrant la fonction constante égale à 1 sur D. Autrement dit, on a $\mathrm{Aire}(D) = \int_D 1\,\mathrm{d}x_1\mathrm{d}x_2$. Pour calculer l'aire d'une surface paramétrée, on peut étendre cette définition de la manière suivante

Définition 13.12 *Soit S une surface régulière. L'*aire *de S est égale à l'intégrale de l'application constante égale à 1 sur la surface S :*

$$\mathrm{Aire}(S) = \int_S 1\,\mathrm{d}\sigma.$$

L'aire d'une surface paramétrée S, orientable, de classe $\mathcal{C}^1$, dont un paramétrage est $\varphi \in \mathcal{C}^1(K,S)$ où $K \subset \mathbb{R}^2$ vaut donc

$$\mathrm{Aire}(S) = \int_K \|\mathbf{N}(u,v)\|\,\mathrm{d}u\mathrm{d}v \quad \text{où } \mathbf{N}(u,v) = \frac{\partial\varphi}{\partial u}(u,v) \wedge \frac{\partial\varphi}{\partial u}(u,v).$$

Remarque La relation précédente indique que l'*aire* d'une surface régulière S est égale au flux du vecteur normal unitaire à travers S. En effet, supposons que $\mathbf{N}$ désigne un vecteur normal à la surface correspondant à un paramétrage $\varphi \in \mathcal{C}^1(K, \mathbb{R}^n)$. Cette normale n'est pas nécessairement unitaire[49]. Un vecteur normal unitaire est $\mathbf{n} = \mathbf{N}/\|\mathbf{N}\|$ de sorte que le flux de $\mathbf{n}$ à travers la surface est égal à

$$\int_S \mathbf{n} \cdot \mathrm{d}\mathbf{s} = \int_K \frac{\mathbf{N}(u,v)}{\|\mathbf{N}(u,v)\|} \cdot \mathbf{N}(u,v)\ \mathrm{d}u\mathrm{d}v = \int_K \|\mathbf{N}(u,v)\|\ \mathrm{d}u\mathrm{d}v,$$

ce qui correspond à l'aire de la surface S. ◇

Exemple Calculons l'aire du morceau S de l'escalier hélicoïdal de hauteur h dont un paramétrage est[50]

$$\varphi : (u,v) \in [0,h] \times [0,1] \longmapsto \big(v\cos(u), v\sin(u), u\big) \in \mathbb{R}^3.$$

Nous avons établi au cours de l'exercice 13.6 p. 977, que cette surface est régulière et qu'un vecteur normal au point $\varphi(u,v)$, pour $(u,v) \in K$ où $K = [0,h] \times [0,1]$, est

$$\mathbf{N}(u,v) = \frac{\partial \varphi}{\partial v}(u,v) \wedge \frac{\partial \varphi}{\partial u}(u,v) = \big(\sin(u), -\cos(u), v\big)\,.$$

L'aire de cette surface vaut donc

$$\mathrm{Aire}(S) = \int_K \|\mathbf{N}(u,v)\|\ \mathrm{d}u\mathrm{d}v = \int_0^h \int_0^1 \sqrt{1+v^2}\ \mathrm{d}u\mathrm{d}v = h \int_0^1 \sqrt{1+v^2}\ \mathrm{d}v.$$

Cette intégrale a été calculée au cours de l'exercice 13.6 ; elle vaut $\frac{1}{2}\left(\sqrt{2} + \mathrm{argsh}(1)\right)$. Finalement, on obtient

$$\mathrm{Aire}(S) = \frac{h}{2}\left(\sqrt{2} + \mathrm{argsh}(1)\right).$$

Les lampions de Schwarz

Nous avons introduit[51] la longueur d'une courbe comme la borne supérieure des longueurs des lignes polygonales inscrites dans cette courbe. Pour la notion d'aire, ce raisonnement intuitif est inexact : par exemple, les lampions de Schwarz (détaillés ci-dessous) fournissent des suites de polyèdres tous inscrits dans un cylindre donné, qui s'approchent du cylindre, mais dont la limite des aires peut prendre n'importe quelle valeur supérieure à l'aire du cylindre.

---

[49] Par contre, puisque la surface est régulière, on est assuré que cette normale est bien définie et ne s'annule pas.

[50] Cette surface a été étudiée p. 788 et p. 793.

[51] Voir la définition 13.3.

Précisons cette affirmation en considérant le cylindre de rayon 1 et de hauteur 1. En le découpant le long d'une de ses génératrices et en le déployant, on obtient un rectangle de longueur 2π et de largeur 1. On réalise un quadrillage de ce rectangle par mn petits rectangles avec m rectangles par colonne et n rectangles par ligne. On trace ensuite les diagonales de chacun de ces mn rectangles (voir la figure 6, à gauche).

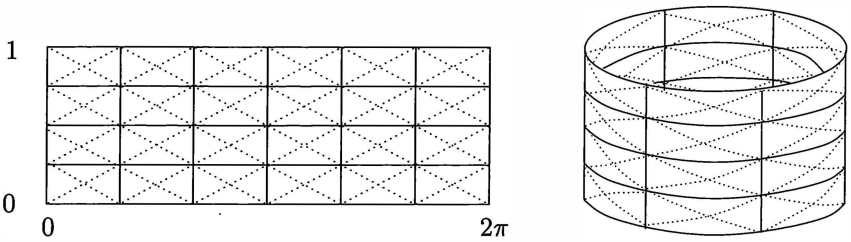


Fig. 6 Quadrillage avec $m = 4$ et $n = 6$.

On reforme le cylindre et on remplace chaque triangle dessiné sur le cylindre par un triangle plat ayant les mêmes sommets. La réunion des petits triangles plats forme un polyèdre $\mathcal{A}_{mn}$ inscrit dans le cylindre (voir la figure 6, à droite) appelé lampion de Schwarz.

L'aire du polyèdre $\mathcal{A}_{mn}$ est la somme des aires des triangles. Le nombre de triangles est égal à $4mn$. La moitié de ces triangles a pour base $1/m$ (base « verticale »), la hauteur de ces triangles étant $2\sin(\pi/2n)$. L'autre moitié des triangles est formée par des triangles dont une base « horizontale » vaut $2\sin(\pi/n)$, la hauteur correspondante valant $\sqrt{1/(4m^2) + (1-\cos(\pi/n))^2}$. Finalement, on obtient

$$\mathcal{A}_{mn} = mn\left(\frac{1}{m} \times 2\sin\left(\frac{\pi}{2n}\right) + 2\sin\left(\frac{\pi}{n}\right) \times \sqrt{\frac{1}{4m^2} + \left(1-\cos\left(\frac{\pi}{n}\right)\right)^2}\right).$$

Cette expression peut se réécrire en utilisant les formules de trigonométrie sous la forme suivante :

$$\mathcal{A}_{mn} = 2n\sin\left(\frac{\pi}{2n}\right)\left(1 + 2\cos\left(\frac{\pi}{2n}\right)\sqrt{\frac{1}{4} + 4m^2\sin^4\left(\frac{\pi}{2n}\right)}\right).$$

Nous allons montrer que selon les valeurs de m et de n, la limite lorsque m et n tendent vers $+\infty$ de $\mathcal{A}_{mn}$ prend des valeurs différentes. Tout d'abord, rappelons que pour k au voisinage de $+\infty$ on a les équivalents suivants :

$$2k\sin\left(\frac{\pi}{2k}\right) \underset{+\infty}{\sim} \pi \quad \text{et} \quad \cos\left(\frac{\pi}{2k}\right) \underset{+\infty}{\sim} 1.$$

Ainsi en prenant $m = n$ et en faisant tendre n vers $+\infty$ on obtient :

$$\lim_{n\to+\infty} \mathcal{A}_{nn} = 2\pi.$$

Par contre, en prenant $m = n^3$ on obtient :

$$\lim_{n\to+\infty} \mathcal{A}_{n^3 n} = +\infty.$$

Plus généralement, montrons que selon les valeurs de m et n, la quantité $\mathcal{A}_{mn}$ peut tendre vers tout réel $A \in [2\pi, +\infty[$. Soit $x \in \mathbb{R}$; comme[52] $\mathbb{Q}$ est dense dans $\mathbb{R}$ le réel x est la limite d'une suite de rationnels $(q_p)_p$ avec $q_p = \alpha_p/\beta_p$ où $\alpha_p \in \mathbb{Z}$ et $\beta_p \in \mathbb{N}^*$. Quitte à multiplier α_p et β_p par un même nombre (dépendant de p), on peut supposer que les suites $(\alpha_p)_p$ et $(\beta_p)_p$ tendent vers $+\infty$ lorsque p tend vers $+\infty$. Pour $p \in \mathbb{N}$ fixé, prenons $n = \beta_p$ et $m = \alpha_p\beta_p$; on a alors

$$\mathcal{A}_{mn} = 2\beta_p \sin\left(\frac{\pi}{2\beta_p}\right)\left(1 + 2\cos\left(\frac{\pi}{2\beta_p}\right)\sqrt{\frac{1}{4} + 4\alpha_p^2\beta_p^2 \sin^4\left(\frac{\pi}{2\beta_p}\right)}\right)$$

$$\mathcal{A}_{mn} \underset{p\to+\infty}{\sim} \pi\left(1 + 2\sqrt{\frac{1}{4} + \frac{\alpha_p^2\pi^4}{4\beta_p^2}}\right) = \pi\left(1 + \sqrt{1 + q_p^2\pi^4}\right).$$

En passant à la limite lorsque p tend vers $+\infty$, on obtient

$$\lim_{p\to+\infty} \mathcal{A}_{mn} = \pi\left(1 + \sqrt{1 + x^2\pi^4}\right).$$

Lorsque x décrit $\mathbb{R}$, la quantité $\pi\left(1 + \sqrt{1 + x^2\pi^4}\right)$ décrit $[2\pi, +\infty[$. Ainsi pour tout réel y supérieur ou égal à 2π, il existe une suite de lampions dont l'aire tend vers y.

Aire d'une surface cartésienne

Considérons une surface cartésienne S définie par l'équation $x_3 = g(x_1, x_2)$ où g est une application de classe $\mathcal{C}^1$ sur un ouvert K de $\mathbb{R}^2$ à valeurs dans $\mathbb{R}$. Cette surface admet pour paramétrage

$$\varphi : (x_1, x_2) \in K \longmapsto (x_1, x_2, g(x_1, x_2))$$

de sorte qu'on peut utiliser la définition 13.12 de l'aire d'une surface paramétrée pour obtenir l'aire de S.

Pour $(x_1, x_2) \in K$, le vecteur normal en $(x_1, x_2, g(x_1, x_2))$ à S vaut[53] :

$$\mathbf{N}(x_1, x_2) = \left(-\frac{\partial g}{\partial x_2}(x_1, x_2), \frac{\partial g}{\partial x_1}(x_1, x_2), 1\right).$$

On déduit que l'aire de cette surface cartésienne vaut

$$\mathcal{A} = \int_K \sqrt{1 + \left(\frac{\partial g}{\partial x_1}(x_1, x_2)\right)^2 + \left(\frac{\partial g}{\partial x_2}(x_1, x_2)\right)^2}\, \mathrm{d}x_1\mathrm{d}x_2.$$

[52] Voir le *Cours de première année*, proposition 3.13 p. 108.

[53] Voir la remarque p. 799.

Aire d'une surface de révolution

Dans un repère orthonormé direct $(O;\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3)$ de $\mathbb{R}^3$, intéressons-nous à une surface de révolution d'axe $(O;\mathbf{e}_3)$ obtenue de la façon suivante. Soit $[a,b]$ un intervalle de $\mathbb{R}$ et Γ la courbe de $\mathbb{R}^3$ paramétrée par

$$\gamma : r \in [a,b] \longmapsto (\rho(r), 0, \varphi(r)) \in \mathbb{R}^3,$$

où ρ et φ désignent deux applications de classe $\mathcal{C}^1$ de $[a,b]$ dans $\mathbb{R}$ avec ρ positive ; cette condition implique que la courbe Γ est incluse dans le demi-plan $\{x_2 = 0, x_1 \geqslant 0\}$. Une surface de révolution Σ est obtenue en faisant « tourner » la courbe Γ autour de l'axe $(O;\mathbf{e}_3)$. Elle admet pour paramétrage

$$\sigma : (r,\theta) \in [a,b] \times [0,2\pi] \longmapsto \big(\rho(r)\cos(\theta), \rho(r)\sin(\theta), \varphi(r)\big) \in \mathbb{R}^3.$$

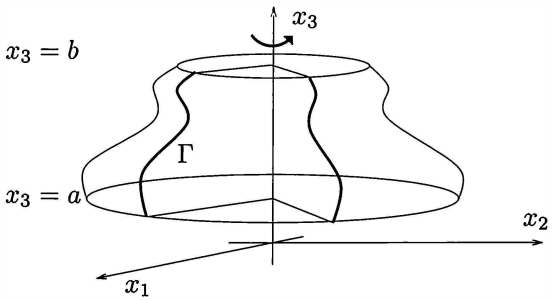


Fig. 7 Une surface de révolution engendrée par la rotation de la courbe Γ autour de l'axe $(O;\mathbf{e}_3)$.

Un vecteur normal à cette surface Σ au point $\sigma(r,\theta)$ est donné par

$$\mathbf{N}(r,\theta) = \frac{\partial\sigma}{\partial r} \wedge \frac{\partial\sigma}{\partial\theta} = \big(-\rho(r)\varphi'(r)\cos(\theta), -\rho(r)\varphi'(r)\sin(\theta), \varphi(r)\big)\,.$$

D'après la définition 13.12, l'aire de Σ s'exprime sous la forme de l'intégrale double sur le domaine $K = [a,b] \times [0,2\pi]$ suivante :

$$\mathcal{A} = \int_K |\rho(r)|\sqrt{\varphi'(r)^2 + \rho'(r)^2}\;\mathrm{d}r\mathrm{d}\theta = 2\pi \int_a^b |\rho(r)|\sqrt{\varphi'(r)^2 + \rho'(r)^2}\;\mathrm{d}r.$$

Exemple La sphère de $\mathbb{R}^3$ de rayon $R \in \mathbb{R}_+^*$ et de centre $(0,0,0)$ est une surface de révolution d'axe $(O;\mathbf{e}_3)$ obtenue en faisant tourner le demi-cercle de centre $(0,0,0)$ de rayon R situé dans le plan $x_2 = 0$ autour de l'axe $(O;\mathbf{e}_3)$. Ce demi-cercle peut être paramétré par

$$\gamma : r \in [-R,R] \longmapsto \Big(R\sqrt{1-(r/R)^2}, 0, r\Big).$$

D'après ce qui précède, un paramétrage de la sphère est

$$\sigma : (r,\theta) \in [-R,R] \times [0,2\pi] \longmapsto \big(\rho(r)\cos(\theta), \rho(r)\sin(\theta), \varphi(r)\big) \in \mathbb{R}^3$$

avec $\rho(r) = R\sqrt{1-(r/R)^2}$ et $\varphi(r) = r$. Puisque pour $r \in [-R, R]$ on a

$$\varphi'(r) = 1 \quad \text{et} \quad \rho'(r) = \frac{-(r/R)}{\sqrt{1-(r/R)^2}},$$

on en déduit que

$$\sqrt{\varphi'(r)^2 + \rho'(r)^2} = \sqrt{1 + \frac{(r/R)^2}{1-(r/R)^2}} = \sqrt{\frac{1}{1-(r/R)^2}}.$$

Par conséquent, l'aire de la sphère vaut

$$\mathcal{A} = 2\pi \int_{-R}^{R} R\sqrt{1-(r/R)^2}\sqrt{\frac{1}{1-(r/R)^2}}\,\mathrm{d}r = 2\pi \int_{-R}^{R} R\,\mathrm{d}r = 4\pi R^2.$$

Exercice 13.7 Déterminer l'aire $\mathcal{A}(\alpha, h)$ d'un tronçon de cône d'axe $(O; \mathbf{e}_3)$ et d'angle au sommet $\alpha \in]0, \pi/2[$ compris entre les plans horizontaux d'équation $x_3 = 0$ et $x_3 = h$ $(h \in \mathbb{R})$.

Proposition 13.10 *Soient f une application continue sur un ouvert U de $\mathbb{R}^n$, à valeurs dans $\mathbb{R}^n$ et S une surface paramétrée régulière de $\mathbb{R}^n$ contenue dans l'ouvert U. On a*

$$\left|\int_S f \cdot \mathrm{d}\mathbf{s}\right| \leqslant \mathrm{Aire}(S) \times \sup_{\mathbf{x}\in S} \|f(\mathbf{x})\|.$$

Démonstration Soit $\varphi \in \mathcal{C}^1(K, \mathbb{R}^n)$ où $K \subset \mathbb{R}^2$, un paramétrage de la surface S. On a

$$\left|\int_S f \cdot \mathrm{d}\mathbf{s}\right| = \left|\int_K f(\varphi(u,v)) \cdot \mathbf{N}(u,v)\,\mathrm{d}u\mathrm{d}v\right| \leqslant \int_K \left|f(\varphi(u,v)) \cdot \mathbf{N}(u,v)\right| \mathrm{d}u\mathrm{d}v,$$

où $\mathbf{N}(u,v) = \frac{\partial\varphi}{\partial u}(u,v) \wedge \frac{\partial\varphi}{\partial u}(u,v)$. D'après l'inégalité de Cauchy-Schwarz[(54)], pour tout $(u,v) \in K$, on a $|f(\varphi(u,v)) \cdot \mathbf{N}(u,v)| \leqslant \|f(\varphi(u,v))\|\,\|\mathbf{N}(u,v)\|$. On en déduit que

$$\left|\int_S f \cdot \mathrm{d}\mathbf{s}\right| \leqslant \int_K \left\|f(\varphi(u,v))\right\| \left\|\mathbf{N}(u,v)\right\| \mathrm{d}u\mathrm{d}v.$$

Puisque pour tout $(u,v) \in K$, on a $\left\|f(\varphi(u,v))\right\| \leqslant \sup_{\mathbf{x}\in S} \|f(\mathbf{x})\|$, on obtient finalement

$$\left|\int_S f \cdot \mathrm{d}\mathbf{s}\right| \leqslant \left(\int_K \|\mathbf{N}(u,v)\|\,\mathrm{d}u\mathrm{d}v\right) \times \sup_{\mathbf{x}\in S} \|f(\mathbf{x})\|.$$

D'après la définition 13.12, l'intégrale $\int_K \|\mathbf{N}(u,v)\|\,\mathrm{d}u\mathrm{d}v$ s'interprète comme l'aire de la surface S, ce qui conclut la preuve. □

(54) Voir le théorème 8.5 p. 481.

13.4 Formule de Stokes

13.4.1 Généralités sur les formes de degré p

Rappelons[55] qu'une application f de $\mathbb{R}^n \times \mathbb{R}^n$ à valeurs dans $\mathbb{R}$ est une forme bilinéaire si elle est linéaire par rapport à chacune des deux variables, c'est-à-dire si

$$\forall (\mathbf{x},\mathbf{y},\mathbf{z}) \in (\mathbb{R}^n)^3 \quad \forall (\lambda,\mu) \in \mathbb{R}^2 \quad \big(\varphi(\lambda\mathbf{x}+\mu\mathbf{y},\mathbf{z}) = \lambda\varphi(\mathbf{x},\mathbf{z}) + \mu\varphi(\mathbf{y},\mathbf{z}) \text{ et } \varphi(\mathbf{x},\lambda\mathbf{y}+\mu\mathbf{z}) = \lambda\varphi(\mathbf{x},\mathbf{y}) + \mu\varphi(\mathbf{x},\mathbf{z}) \big).$$

Cette notion se généralise à des applications définies sur $(\mathbb{R}^n)^p = \mathbb{R}^n \times \cdots \times \mathbb{R}^n$ (produit cartésien p fois de $\mathbb{R}^n$ avec lui-même) de la façon suivante :

Définition 13.13 *Soit f une application de $(\mathbb{R}^n)^p$, $p \in \mathbb{N}^*$, à valeurs dans $\mathbb{R}^q$, $q \in \mathbb{N}^*$. On dit que l'application f est une application* p-linéaire *si elle est « linéaire par rapport à chaque variable », c'est-à-dire si*

$$\forall i \in \{1,\dots,p\} \quad \forall \mathbf{y}_i \in \mathbb{R}^n \quad \forall (\mathbf{x}_1,\dots,\mathbf{x}_n) \in (\mathbb{R}^n)^p \quad \forall (\lambda,\mu) \in \mathbb{R}^2$$
$$f(\mathbf{x}_1,\dots,\mathbf{x}_{i-1},\lambda\mathbf{x}_i+\mu\mathbf{y}_i,\mathbf{x}_{i+1},\dots,\mathbf{x}_n) = \lambda f(\mathbf{x}_1,\dots,\mathbf{x}_{i-1},\mathbf{x}_i,\mathbf{x}_{i+1},\dots,\mathbf{x}_n) + \mu f(\mathbf{x}_1,\dots,\mathbf{x}_{i-1},\mathbf{y}_i,\mathbf{x}_{i+1},\dots,\mathbf{x}_n).$$

Remarque Une application 2-linéaire est une application bilinéaire. Une application 3-linéaire est aussi appelée une *application trilinéaire.* ◇

Soit f une application bilinéaire de $\mathbb{R}^n \times \mathbb{R}^n$ dans $\mathbb{R}$ antisymétrique[56].
– Pour tout $(\mathbf{x},\mathbf{y}) \in \mathbb{R}^n \times \mathbb{R}^n$ on a $f(\mathbf{x},\mathbf{y}) = -f(\mathbf{y},\mathbf{x})$ et en particulier lorsque $\mathbf{x} = \mathbf{y}$, on obtient $f(\mathbf{x},\mathbf{x}) = 0$.
– Réciproquement, considérons une application bilinéaire f de $\mathbb{R}^n \times \mathbb{R}^n$ dans $\mathbb{R}$ vérifiant $f(\mathbf{z},\mathbf{z}) = 0$ pour tout $\mathbf{z} \in \mathbb{R}^n$. Pour tout $(\mathbf{x},\mathbf{y}) \in (\mathbb{R}^n)^2$ on a $f(\mathbf{x}+\mathbf{y},\mathbf{x}+\mathbf{y}) = 0$ et la bilinéarité de l'application f implique que pour tout $(\mathbf{x},\mathbf{y}) \in (\mathbb{R}^n)^2$ on a $f(\mathbf{x},\mathbf{y}) = -f(\mathbf{y},\mathbf{x})$, ce qui signifie que l'application f est antisymétrique.
On a établi l'équivalence suivante : une application bilinéaire de $\mathbb{R}^n \times \mathbb{R}^n$ à valeurs dans $\mathbb{R}$ est antisymétrique si et seulement si elle est vérifie

$$\forall \mathbf{x} \in \mathbb{R}^n \quad f(\mathbf{x},\mathbf{x}) = 0.$$

On dit dans ce cas que f est alternée. C'est sous cette forme que l'on généralise la notion d'« antisymétrie » pour une application p-linéaire définie sur $(\mathbb{R}^n)^p$.

---

[55] Voir la définition 8.1 p. 438.

[56] Rappelons, voir la définition 8.5, p. 446, qu'une application bilinéaire f est dite antisymétrique si pour tout $(\mathbf{x},\mathbf{y}) \in U^2$ on a $f(\mathbf{x},\mathbf{y}) = -f(\mathbf{y},\mathbf{x})$.

DÉFINITION 13.14 *Soit f une application p-linéaire de $(\mathbb{R}^n)^p$ dans $\mathbb{R}$. On dit que l'application f est* alternée *lorsqu'elle vérifie $f(\mathbf{x}_1, \ldots, \mathbf{x}_p) = 0$ dès que deux des vecteurs $\mathbf{x}_i \in \mathbb{R}^n$ sont identiques ; en d'autres termes, f est dite alternée si pour tout $(\mathbf{x}_1, \ldots, \mathbf{x}_n) \in (\mathbb{R}^n)^p$,*

$$\Big(\exists (i,j) \in \{1, \ldots, p\}^2 \quad i \neq j \quad \text{et} \quad \mathbf{x}_i = \mathbf{x}_j\Big) \Longrightarrow f(\mathbf{x}_1, \ldots, \mathbf{x}_n) = 0.$$

Comme nous l'avons établi en préambule, on a l'équivalence suivante : une application bilinéaire de $\mathbb{R}^n \times \mathbb{R}^n$ à valeurs dans $\mathbb{R}$ est antisymétrique si et seulement si elle est alternée. La terminologie employée est celle d'application « alternée », le terme antisymétrique étant réservé aux applications bilinéaires.

Remarque De la même manière que l'on a montré l'équivalence entre être antisymétrique et être alternée pour une application bilinéaire de $\mathbb{R}^n \times \mathbb{R}^n$ dans $\mathbb{R}$, on peut caractériser une application p-linéaire alternée f de $(\mathbb{R}^n)^p$ dans $\mathbb{R}$ par la condition

$$\forall (i,j) \in \{(k,\ell) \in \mathbb{N}^2 \ ; \ 1 \leqslant k < \ell \leqslant n\} \quad \forall (\mathbf{x}_1, \ldots, \mathbf{x}_n) \in (\mathbb{R}^n)^p$$
$$f(\mathbf{x}_1, \ldots, \mathbf{x}_i, \ldots, \mathbf{x}_j, \ldots, \mathbf{x}_n) = -f(\mathbf{x}_1, \ldots, \mathbf{x}_j, \ldots, \mathbf{x}_i, \ldots, \mathbf{x}_n).$$

Autrement dit, le signe de f change si l'on permute deux des arguments parmi les p arguments de f. ◇

Exemples

1. Considérons deux vecteurs $\mathbf{u}$ et $\mathbf{v}$ de $\mathbb{R}^2$ et notons $\mathrm{A}(\mathbf{u}, \mathbf{v})$ l'aire algébrique[57] du parallélogramme construit à partir des vecteurs $\mathbf{u}$ et $\mathbf{v}$, voir la figure 8. On a

$$\mathrm{A}(\mathbf{u}, \mathbf{v}) = \det(\mathbf{u}, \mathbf{v}) = \det \begin{pmatrix} u_1 & v_1 \\ u_2 & v_2 \end{pmatrix} = u_1 v_2 - u_2 v_1$$

où (u_1, u_2) et (v_1, v_2) désignent respectivement les composantes des vecteurs $\mathbf{u}$ et $\mathbf{v}$. L'application $\mathrm{A} : (\mathbf{u}, \mathbf{v}) \in \mathbb{R}^3 \times \mathbb{R}^3 \mapsto \det(\mathbf{u}, \mathbf{v})$ ainsi définie vérifie les propriétés suivantes[58] :

- l'application A est une application bilinéaire,
- pour tout $\mathbf{u} \in \mathbb{R}^2$ on a $\mathrm{A}(\mathbf{u}, \mathbf{u}) = 0$.

Autrement dit, l'application A est une application 2-linéaire alternée.

---

[57] Une aire (resp. une longueur, un volume) algébrique est une aire (resp. une longueur, un volume) affectée d'un signe, alors qu'une aire (resp. une longueur, un volume) géométrique qui est toujours positive. La mesure géométrique est la valeur absolue de la mesure algébrique.

[58] Ces propriétés résultent des propriétés du déterminant, voir le *Cours de première année* p. 465.

2. Étant donné trois vecteurs $\mathbf{u}$, $\mathbf{v}$ et $\mathbf{w}$ de $\mathbb{R}^3$, le volume algébrique $\mathrm{V}(\mathbf{u},\mathbf{v},\mathbf{w})$ du parallélépipède construit à partir de ces vecteurs, voir la figure 8, vaut

$$\mathrm{V}(\mathbf{u},\mathbf{v},\mathbf{w}) = \det(\mathbf{u},\mathbf{v},\mathbf{w}) = \det\begin{pmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{pmatrix}$$

où (u_1,u_2,u_3), (v_1,v_2,v_3) et (w_1,w_2,w_3) désignent respectivement les composantes des vecteurs $\mathbf{u}$, $\mathbf{v}$ et $\mathbf{w}$. Il résulte des propriétés du déterminant que l'application

$$\mathrm{V} : (\mathbf{u},\mathbf{v},\mathbf{w}) \in \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \mapsto \det(\mathbf{u},\mathbf{v},\mathbf{w})$$

est trilinéaire et vérifie $\mathrm{A}(\mathbf{u},\mathbf{v},\mathbf{w}) = 0$ dès que deux des trois vecteurs $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ sont égaux. Autrement dit, l'application V est une application trilinéaire alternée.

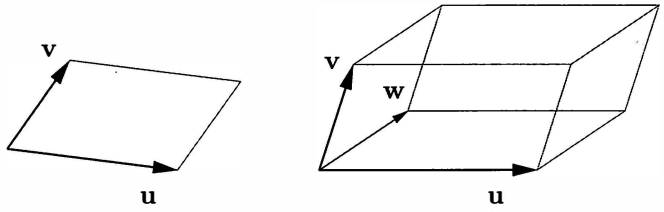


Fig. 8 Parallélogramme et parallélépipède construits à partir des vecteurs $\mathbf{u}$, $\mathbf{v}$ et $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$.

Exercice 13.8 ***Produit vectoriel***
Considérons l'application

$$f : (\mathbf{u},\mathbf{v}) \in \mathbb{R}^3 \times \mathbb{R}^3 \longmapsto (u_2v_3 - u_3v_2\ ,\ u_3v_1 - u_1v_3\ ,\ u_1v_2 - u_2v_1) \in \mathbb{R}^3,$$

où (u_1,u_2,u_3) et (v_1,v_2,v_3) désignent respectivement les composantes de $\mathbf{u}$ et $\mathbf{v}$ dans la base canonique de $\mathbb{R}^3$.
1 - Montrer que f est une application bilinéaire alternée.

Étant donné deux vecteurs $\mathbf{u}$ *et* $\mathbf{v}$ *de* $\mathbb{R}^3$, *le vecteur* $f(\mathbf{u},\mathbf{v})$ *est appelé le* produit vectoriel *de* $\mathbf{u}$ *par* $\mathbf{v}$, *et on le note* $\mathbf{u}\wedge\mathbf{v}$.

2 - Montrer les propriétés suivantes, pour tout $(\mathbf{u},\mathbf{v},\mathbf{w}) \in (\mathbb{R}^3)^3$:

a) $\mathbf{u}\wedge\mathbf{v} = -\mathbf{v}\wedge\mathbf{u}$; b) $\mathbf{u}\wedge\mathbf{u} = 0$;

c) $(\mathbf{u}+\mathbf{v})\wedge\mathbf{w} = \mathbf{u}\wedge\mathbf{w} + \mathbf{v}\wedge\mathbf{w}$; d) $\mathbf{w}\wedge(\mathbf{u}+\mathbf{v}) = \mathbf{w}\wedge\mathbf{u} + \mathbf{w}\wedge\mathbf{v}$.

Exercice 13.9 ***Produit mixte***
Considérons l'application

$$g : (\mathbf{u}, \mathbf{v}, \mathbf{w}) \in \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \longmapsto (\mathbf{u} \wedge \mathbf{v}) \cdot \mathbf{w} \in \mathbb{R},$$

où $\wedge$ désigne le produit vectoriel et $\cdot$ désigne le produit scalaire canonique entre deux vecteurs de $\mathbb{R}^3$.
1 - Montrer que g est une application trilinéaire alternée.
2 - En déduire que pour tout $(\mathbf{u}, \mathbf{v}) \in \mathbb{R}^3 \times \mathbb{R}^3$ le produit vectoriel $\mathbf{u} \wedge \mathbf{v}$ est un vecteur de $\mathbb{R}^3$ orthogonal aux vecteurs $\mathbf{u}$ et $\mathbf{v}$.

Étant donné trois vecteurs $\mathbf{u}$, $\mathbf{v}$ *et* $\mathbf{w}$ *de* $\mathbb{R}^3$, *le réel* $g(\mathbf{u}, \mathbf{v}, \mathbf{w})$ *est appelé le* produit mixte *de* $\mathbf{u}$, $\mathbf{v}$ *et* $\mathbf{w}$, *et on le note* $\langle\!\langle \mathbf{u}, \mathbf{v}, \mathbf{w} \rangle\!\rangle$.

3 - Comparer le produit mixte $\langle\!\langle \mathbf{u}, \mathbf{v}, \mathbf{w} \rangle\!\rangle$ et le volume du parallélépipède construit à partir des trois vecteurs $\mathbf{u}$, $\mathbf{v}$ et $\mathbf{w}$.
4 - Montrer que pour tout $(\mathbf{u}, \mathbf{v}, \mathbf{w}) \in \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3$ on a

$$\mathbf{u} \wedge (\mathbf{v} \wedge \mathbf{w}) = (\mathbf{w} \cdot \mathbf{u})\mathbf{v} - (\mathbf{u} \cdot \mathbf{v})\mathbf{w}.$$

DÉFINITION 13.15 *Soient U un ouvert de $\mathbb{R}^n$ et $p \in \mathbb{N}^*$. On note $\Lambda^p(U)$ l'ensemble des applications définies sur U et à valeurs dans l'ensemble des applications p-linéaires alternées de $(\mathbb{R}^n)^p$ dans $\mathbb{R}$. De telles applications sont appelées* formes différentielles de degré p *ou plus simplement* p-formes.

Par convention, on appelle 0-forme (ou forme différentielle de degré 0) sur un ouvert U de $\mathbb{R}^n$ toute application de U dans $\mathbb{R}$. Bien entendu, la définition des p-formes avec $p = 1$ coïncide avec la définition des 1-formes donnée p. 956.

Exemples

1. L'application $\mathbf{x} \in \mathbb{R}^2 \mapsto \mathrm{A}$ où A est l'application 2-linéaire alternée définie par[59] $\mathrm{A} : (\mathbf{u}, \mathbf{v}) \in (\mathbb{R}^2)^2 \mapsto \det(\mathbf{u}, \mathbf{v})$ est une 2-forme de $\mathbb{R}^2$. On dira[60] que c'est une 2-forme constante puisqu'à tout élément $\mathbf{x} \in \mathbb{R}^2$ elle associe l'application 2-linéaire alternée A qui ne dépend pas de $\mathbf{x}$.

2. Plus généralement, l'application $\mathbf{x} \in \mathbb{R}^n \mapsto \det$ où det est l'application déterminant (qui est une application n-linéaire alternée[61]) définie par

$$\det : (\mathbf{u}_1, \ldots, \mathbf{u}_n) \in (\mathbb{R}^n)^n \longmapsto \det(\mathbf{u}_1, \ldots, \mathbf{u}_n)$$

est une n-forme constante de $\mathbb{R}^n$.

[59] Voir l'exemple de la p. 986.
[60] Voir la définition 13.16 p. 990.
[61] Voir le *Cours de première année* p. 483.

Cas particulier des 2-formes

On a vu[62] qu'il existe des 1-formes qui ont un rôle particulier : elles « engendrent » tous les autres 1-formes. On va construire des 2-formes qui ont un rôle similaire pour l'ensemble des 2-formes. Pour tout $(i,j) \in \{1,\dots,n\}^2$, l'application notée $\mathrm{d}x_i \wedge \mathrm{d}x_j$ et définie par

$$\mathrm{d}x_i \wedge \mathrm{d}x_j : (\mathbf{u},\mathbf{v}) \in (\mathbb{R}^n)^2 \longmapsto \det \begin{pmatrix} u_i & v_i \\ u_j & v_j \end{pmatrix}$$

où $(u_1,\dots,u_n)$ et $(v_1,\dots,v_n)$ désignent respectivement les composantes de $\mathbf{u}$ et $\mathbf{v}$ dans la base canonique de $\mathbb{R}^n$ est une 2-forme constante sur $\mathbb{R}^n$. On peut montrer que ces 2-formes « engendrent » tous les autres 2-formes. Il résulte des propriétés du déterminant que pour $(i,j) \in \{1,\dots,n\}^2$ on a

$$\mathrm{d}x_i \wedge \mathrm{d}x_j = -\mathrm{d}x_j \wedge \mathrm{d}x_i$$

et qu'en particulier pour $i \in \{1,\dots,n\}$ on a $\mathrm{d}x_i \wedge \mathrm{d}x_i = 0$. Cela permet d'affirmer que pour décrire l'ensemble des 2-formes, il suffit de considérer uniquement les 2-formes $\mathrm{d}x_i \wedge \mathrm{d}x_j$ avec $1 \leqslant i < j \leqslant n$. Plus précisément, on a la proposition suivante que nous admettrons.

Proposition 13.11 *Soit U un ouvert de $\mathbb{R}^n$; la famille $\{\mathrm{d}x_i \wedge \mathrm{d}x_j\}_{1\leqslant i<j\leqslant n}$ engendre $\Lambda^2(U)$ au sens suivant : pour toute 2-forme $\omega \in \Lambda^2(U)$ il existe une unique famille d'applications $\{\alpha_{ij}\}_{1\leqslant i<j\leqslant n}$ de U dans $\mathbb{R}$ telle que*

$$\omega = \sum_{i=1}^{n} \sum_{j=1}^{i-1} \alpha_{ij}\, \mathrm{d}x_i \wedge \mathrm{d}x_j.$$

Remarque La somme « double » $\sum_{i=1}^{n} \sum_{j=1}^{i-1}$ s'écrit aussi $\sum_{1\leqslant i<j\leqslant n}$. Elle correspond à la somme sur tous les couples d'entiers $(i,j) \in \{1,\dots,n\}^2$ tels que $i < j$. ◇

Le cardinal de la famille $\{\mathrm{d}x_i \wedge \mathrm{d}x_j\}_{1\leqslant i<j\leqslant n}$ est égal à $\frac{1}{2}n(n-1)$.

- Cette famille est vide pour $n = 1$, autrement dit la seule 2-forme sur un ouvert de $\mathbb{R}$ est la 2-forme nulle : $\Lambda^2(\mathbb{R}) = \{0\}$.
- Lorsque $n = 2$, cette famille contient le seul élément $\mathrm{d}x_1 \wedge \mathrm{d}x_2$, ce qui montre que toutes les 2-formes de $\mathbb{R}^2$ sont du type $\omega = \alpha\, \mathrm{d}x_1 \wedge \mathrm{d}x_2$ où $\alpha \in \mathcal{A}(U,\mathbb{R})$. On pourra remarquer que cette 2-forme $\mathrm{d}x_1 \wedge \mathrm{d}x_2$ est la 2-forme notée A dans les exemples des pp. 986 et 988.
- Pour $n = 3$, il existe trois 2-formes qui engendrent toutes les 2-formes de $\mathbb{R}^3$; ce sont les 2-formes $\mathrm{d}x_2 \wedge \mathrm{d}x_3$, $\mathrm{d}x_3 \wedge \mathrm{d}x_1$ et $\mathrm{d}x_1 \wedge \mathrm{d}x_2$.

[62] Voir la proposition 13.6 p. 956.

DÉFINITION 13.16 *Soit* $\omega = \sum_{i=1}^{n} \sum_{j=1}^{i-1} \alpha_{ij}\, \mathrm{d}x_i \wedge \mathrm{d}x_j$ *une 2-forme définie sur un ouvert* U *de* $\mathbb{R}^n$.

✗ *La 2-forme* ω *est* constante *si pour tout* $(i,j) \in \{1,\dots,n\}^2$ *l'application* α_{ij} *est une application constante.*

✗ *La 2-forme* ω *est de classe* $\mathcal{C}^k$ *où* $k \in \mathbb{N} \cup \{\infty\}$ *si pour tout* $(i,j) \in \{1,\dots,n\}^2$ *l'application* α_{ij} *est une application de classe* $\mathcal{C}^k$.

Cas particulier des 3-formes

Pour ce qui est des 3-formes d'un ouvert U de $\mathbb{R}^n$, on peut effectuer le même raisonnement que celui effectué avec les 2-formes. Pour tout $(i,j,k) \in \{1,\dots,n\}^3$, l'application notée $\mathrm{d}x_i \wedge \mathrm{d}x_j \wedge \mathrm{d}x_k$ et définie par

$$\mathrm{d}x_i \wedge \mathrm{d}x_j \wedge \mathrm{d}x_k : (\mathbf{u},\mathbf{v},\mathbf{w}) \in (\mathbb{R}^n)^3 \longmapsto \det \begin{pmatrix} u_i & v_i & w_i \\ u_j & v_j & w_j \\ u_k & v_k & w_k \end{pmatrix}.$$

où $(u_1,\dots,u_n)$ et $(v_1,\dots,v_n)$ désignent respectivement les composantes de $\mathbf{u}$ et $\mathbf{v}$ dans la base canonique de $\mathbb{R}^n$, est une 3-forme constante sur $\mathbb{R}^n$. Les 3-formes d'un ouvert U de $\mathbb{R}^n$ sont toutes du type

$$\omega = \sum_{i=1}^{n} \sum_{j=1}^{i-1} \sum_{k=1}^{j-1} \alpha_{ijk}\, \mathrm{d}x_i \wedge \mathrm{d}x_j \wedge \mathrm{d}x_k,$$

où α_{ijk} désigne une application de U dans $\mathbb{R}$. Notons que cette somme « triple » $\sum_{i=1}^{n} \sum_{j=1}^{i-1} \sum_{k=1}^{j-1}$ s'écrit aussi $\sum_{1 \leqslant i<j<k \leqslant n}$. Le cardinal de la famille de 3-formes engendrant toutes les 3-formes de $\mathbb{R}^n$ est donc le cardinal de l'ensemble

$$\{(i,j,k) \in \{1,\dots,n\}^3 \ ;\ 1 \leqslant i < j < k \leqslant n\}.$$

On en déduit en particulier que

$$\Lambda^3(\mathbb{R}) = \Lambda^3(\mathbb{R}^2) = \{0\} \quad \text{et} \quad \Lambda^3(\mathbb{R}^3) = \{\alpha\, \mathrm{d}x_1 \wedge \mathrm{d}x_2 \wedge \mathrm{d}x_3 \ ;\ \alpha \in \mathcal{A}(\mathbb{R}^3,\mathbb{R})\}.$$

On notera que la 3-forme $\mathrm{d}x_1 \wedge \mathrm{d}x_2 \wedge \mathrm{d}x_3$ est la 3-forme notée V à la p. 986.

Décomposition des p-formes

On montre que plus généralement, une p-forme ω sur un ouvert U de $\mathbb{R}^n$ est toujours du type

$$\omega = \sum_{1 \leqslant i_1 < \dots < i_p \leqslant n} \alpha_{i_1 \dots i_p}\, \mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_p}$$

où chaque $\alpha_{i_1 \dots i_p}$ $(1 \leqslant i_1 < \cdots < i_p \leqslant n)$ est une application de U dans $\mathbb{R}$ et où les p-formes constantes $\mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_p}$ désignent des applications de $\mathbb{R}^p$ dans $\mathbb{R}$ définies de façon analogue aux 2-formes $\mathrm{d}x_i \wedge \mathrm{d}x_j$ et aux 3-formes $\mathrm{d}x_i \wedge \mathrm{d}x_j \wedge \mathrm{d}x_k$ données précédemment : la p-forme $\mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_p}$ est l'application qui aux p vecteurs $\mathbf{u}_1, \dots, \mathbf{u}_p$ de $\mathbb{R}^n$ associe le déterminant

$$\det \begin{pmatrix} u_{1i_1} & u_{2i_1} & & u_{pi_1} \\ u_{1i_2} & u_{2i_2} & & u_{pi_2} \\ \vdots & \vdots & & \vdots \\ u_{1i_p} & u_{2i_p} & & u_{pi_p} \end{pmatrix}$$

où $(u_1, \dots, u_n)$ et $(v_1, \dots, v_n)$ désignent respectivement les composantes de $\mathbf{u}$ et $\mathbf{v}$ dans la base canonique de $\mathbb{R}^n$.

Remarque Il résulte des propriétés du déterminant que si deux indices parmi les p indices $i_1, \dots, i_p$ sont égaux alors la p-forme $\mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_p}$ est nulle. De même, si on échange deux indices alors on obtient la p-forme opposée. Ainsi, pour $p = 2$, on obtient

$$\forall (i,j) \in \{1, \dots, n\}^2 \quad \mathrm{d}x_i \wedge \mathrm{d}x_i = 0 \quad \text{et} \quad \mathrm{d}x_i \wedge \mathrm{d}x_j = -\mathrm{d}x_j \wedge \mathrm{d}x_i,$$

alors que lorsque $p = 3$ on a par exemple

$$\mathrm{d}x_1 \wedge \mathrm{d}x_2 \wedge \mathrm{d}x_1 = 0 \quad \text{et} \quad \mathrm{d}x_1 \wedge \mathrm{d}x_2 \wedge \mathrm{d}x_3 = -\mathrm{d}x_1 \wedge \mathrm{d}x_3 \wedge \mathrm{d}x_2.$$

◇

À l'aide de cette décomposition des p-formes, on construit une loi somme sur l'ensemble $\Lambda^p(U)$: si $\omega = \sum_{1 \leqslant i_1 < \dots < i_p \leqslant n} \alpha_{i_1 \dots i_p} \ \mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_p}$ et $\widetilde{\omega} = \sum_{1 \leqslant i_1 < \dots < i_p \leqslant n} \beta_{i_1 \dots i_p} \ \mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_p}$ sont deux p-formes alors on définit la p-forme somme de ω et $\widetilde{\omega}$ de la manière suivante :

$$\omega + \widetilde{\omega} = \sum_{1 \leqslant i_1 < \dots < i_p \leqslant n} (\alpha_{i_1 \dots i_p} + \beta_{i_1 \dots i_p}) \, \mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_p}.$$

On définit le produit d'une p-forme $\omega = \sum_{1 \leqslant i_1 < \dots < i_p \leqslant n} \alpha_{i_1 \dots i_p} \ \mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_p}$ par une application f de U dans $\mathbb{R}$ de la façon suivante :

$$f\omega = \sum_{1 \leqslant i_1 < \dots < i_p \leqslant n} (f \alpha_{i_1 \dots i_p}) \, \mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_p}.$$

Produit extérieur

Nous avons introduit, sans le mentionner explicitement, une loi sur l'ensemble des formes différentielles appelée *produit extérieur* et noté $\wedge$. Le produit des deux 1-formes $\mathrm{d}x_i$ et $\mathrm{d}x_j$ est la 2-forme $\mathrm{d}x_i \wedge \mathrm{d}x_j$, le produit de la 1-forme $\mathrm{d}x_i$ et de la 2-forme $\mathrm{d}x_j \wedge \mathrm{d}x_k$ est la 3-forme $\mathrm{d}x_i \wedge \mathrm{d}x_j \wedge \mathrm{d}x_k$, etc. Puisque chaque forme différentielle se décompose selon une famille de formes différentielles $\{\mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_p}\}_{1 \leqslant i_1 \leqslant \cdots \leqslant i_p \leqslant n}$, on prolonge par linéarité le produit extérieur à toute p-forme sur le principe suivant : le produit extérieur des deux 1-formes $\sum_{i=1}^n \alpha_i \,\mathrm{d}x_i$ et $\sum_{j=1}^n \beta_j \,\mathrm{d}x_j$ est la 2-forme

$$\left(\sum_{i=1}^n \alpha_i \,\mathrm{d}x_i\right) \wedge \left(\sum_{j=1}^n \beta_j \,\mathrm{d}x_j\right) = \sum_{i=1}^n \sum_{j=1}^n \alpha_i \beta_j \,\mathrm{d}x_i \wedge \mathrm{d}x_j.$$

Malgré leur notation commune, on prendra garde à ne pas confondre le produit extérieur entre deux formes différentielles ω et $\widetilde{\omega}$, noté $\omega \wedge \widetilde{\omega}$ et le produit vectoriel entre deux vecteurs $\mathbf{u}$ et $\mathbf{v}$, noté $\mathbf{u} \wedge \mathbf{v}$.

Exemple En utilisant la remarque de la p. 991, on établit que le produit extérieur des deux 1-formes $\omega = \sum_{i=1}^n \alpha_i \mathrm{d}x_i$ et $\widetilde{\omega} = \sum_{j=1}^n \beta_j \mathrm{d}x_j$ s'écrit

$$\begin{aligned}\left(\sum_{i=1}^n \alpha_i \,\mathrm{d}x_i\right) \wedge \left(\sum_{j=1}^n \beta_j \,\mathrm{d}x_j\right) &= \sum_{i=1}^n \sum_{j=1}^{i-1} \alpha_i \beta_j \,\mathrm{d}x_i \wedge \mathrm{d}x_j \\ &\quad + \sum_{i=1}^n \alpha_i \beta_i \underbrace{\mathrm{d}x_i \wedge \mathrm{d}x_i}_{=0} + \sum_{i=1}^n \sum_{j=i+1}^n \alpha_i \beta_j \underbrace{\mathrm{d}x_i \wedge \mathrm{d}x_j}_{=-\mathrm{d}x_j \wedge \mathrm{d}x_i} \\ &= \sum_{i=1}^n \sum_{j=1}^{i-1} (\alpha_i \beta_j - \alpha_j \beta_i) \,\mathrm{d}x_i \wedge \mathrm{d}x_j.\end{aligned}$$

En prenant $\omega = \widetilde{\omega}$ on en déduit que pour toute 1-forme ω on a $\omega \wedge \omega = 0$.

Exercice 13.10 Soit $\omega = \mathrm{d}x_1 \wedge \mathrm{d}x_2 + \mathrm{d}x_3 \wedge \mathrm{d}x_4$ une 2-forme de $\mathbb{R}^n$ où n est un entier supérieur ou égal à 4. Calculer $\omega \wedge \omega$.

13.4.2 Intégrale d'une 2-forme sur une surface de $\mathbb{R}^3$

Considérons une application continue α d'un ouvert U de $\mathbb{R}^3$ à valeurs dans $\mathbb{R}^3$ et une surface paramétrée S de $\mathbb{R}^3$ incluse dans U dont un paramétrage est l'application $\varphi \in \mathcal{C}^1(K, U)$. Dans la section 13.3, nous avons défini l'intégrale $\int_S \alpha \cdot \mathrm{d}\mathbf{s}$ de α à travers la surface S de la façon suivante :

$$\int_S \alpha \cdot \mathrm{d}\mathbf{s} = \int_K \alpha(\varphi(u,v)) \cdot \mathbf{N}(u,v) \,\mathrm{d}u\mathrm{d}v,$$

où $\mathbf{N}(u,v) = \dfrac{\partial\varphi}{\partial u}(u,v) \wedge \dfrac{\partial\varphi}{\partial u}(u,v)$.

On peut associer à l'application $\alpha : \mathbf{x} \in U \mapsto (\alpha_1(\mathbf{x}), \alpha_2(\mathbf{x}), \alpha_3(\mathbf{x})) \in \mathbb{R}^3$ la 2-forme $\omega = \alpha_1\ \mathrm{d}x_2 \wedge \mathrm{d}x_3 + \alpha_2\ \mathrm{d}x_3 \wedge \mathrm{d}x_1 + \alpha_3\ \mathrm{d}x_1 \wedge \mathrm{d}x_2$. En utilisant la proposition 13.11, on montre que l'application qui à α associe ω est bijective. Cela conduit à définir l'intégrale d'une 2-forme de la manière suivante.

DÉFINITION 13.17 *Soient U un ouvert de $\mathbb{R}^3$, S une surface de $\mathbb{R}^3$ paramétrée, orientable, de classe $\mathcal{C}^1$ incluse dans U et*

$$\omega = \alpha_1\ \mathrm{d}x_2 \wedge \mathrm{d}x_3 + \alpha_2\ \mathrm{d}x_3 \wedge \mathrm{d}x_1 + \alpha_3\ \mathrm{d}x_1 \wedge \mathrm{d}x_2$$

une 2-forme continue sur U.

On appelle intégrale de la 2-forme ω sur la surface S l'intégrale de l'application $\alpha : \mathbf{x} \in U \mapsto (\alpha_1(\mathbf{x}), \alpha_2(\mathbf{x}), \alpha_3(\mathbf{x}))$ sur S ; en d'autres termes

$$\int_S \omega = \int_S \alpha \cdot \mathrm{d}\mathbf{s}.$$

Exemple Considérons sur $\mathbb{R}^3$ la 2-forme $\omega = x_1\ \mathrm{d}x_2 \wedge \mathrm{d}x_3$ et la surface paramétrée[63] S de $\mathbb{R}^3$ dont un paramétrage est

$$\varphi : (u,v) \in K \longmapsto (u, v, u^2 + v^2) \in \mathbb{R}^3$$

où K est le disque de $\mathbb{R}^2$ de centre $(0,0)$ et de rayon 1. Intégrer ω sur S revient à intégrer l'application $\alpha : (x_1, x_2, x_3) \in \mathbb{R}^3 \mapsto (x_1, 0, 0) \in \mathbb{R}^3$ sur la surface S ; on a :

$$\int_S x_1 \mathrm{d}x_2 \wedge \mathrm{d}x_3 = \int_S \alpha \cdot \mathrm{d}\mathbf{s}.$$

Pour $(u,v) \in K$, un vecteur normal à la surface au point $\varphi(u,v)$ est

$$\mathbf{N}(u,v) = \frac{\partial\varphi}{\partial u}(u,v) \wedge \frac{\partial\varphi}{\partial u}(u,v) = \big(2u, 2v, 1\big)$$

et $\alpha(\varphi(u,v)) = (u,0,0)$. L'intégrale de la 2-forme ω sur la surface S est donc

$$\int_S \omega = \int_S x_1 \mathrm{d}x_2 \wedge \mathrm{d}x_3 = \int_K \alpha(\varphi(u,v)) \cdot \mathbf{N}(u,v)\ \mathrm{d}u\mathrm{d}v = \int_K 2u^2\ \mathrm{d}u\mathrm{d}v.$$

En utilisant un changement de variable en coordonnées polaires, on obtient

$$\int_S \omega = \int_0^{2\pi} \int_0^1 2r^2 \cos(\theta)^2 r\ \mathrm{d}r\mathrm{d}\theta = \left(\int_0^{2\pi} \cos(\theta)^2\ \mathrm{d}\theta\right)\left(\int_0^1 2r^3\ \mathrm{d}r\right) = \frac{\pi}{2}.$$

[63] Cette surface est un « morceau » du paraboloïde de révolution d'équation implicite $x_3 = x_1^2 + x_2^2$.

Exercice 13.11 Calculer l'intégrale de la 2-forme

$$\omega : (x_1, x_2, x_3) \in \mathbb{R}^3 \longmapsto x_1 \mathrm{d}x_2 \wedge \mathrm{d}x_3$$

sur la surface $S = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \ ; \ x_3 = 1 \text{ et } (x_1, x_2) \in K\}$ où K est un sous-ensemble borné quelconque de $\mathbb{R}^2$.

On a vu que la quantité $\mathbf{dx}$ pouvait être interprétée comme le « vecteur » de $\mathbb{R}^n$ dont les n composantes correspondent aux 1-formes qui engendrent toutes les 1-formes de $\mathbb{R}^n$, c'est-à-dire $\mathbf{dx} = (\mathrm{d}x_1, \ldots, \mathrm{d}x_n)$. Chacune de ces 1-formes pouvant être assimilée à un déplacement infinitésimal dans une direction de la base canonique de $\mathbb{R}^n$, la quantité $\mathbf{dx}$ peut être interprétée comme un déplacement infinitésimal de $\mathbb{R}^n$. De même, d'après la définition 13.17, la quantité $\mathbf{ds}$ peut être interprétée comme le « vecteur » de $\mathbb{R}^3$ dont les 3 composantes correspondent aux 2-formes qui engendrent toutes les 2-formes de $\mathbb{R}^3$, c'est-à-dire que l'on peut considérer que $\mathbf{ds} = (\mathrm{d}x_2 \wedge \mathrm{d}x_3, \mathrm{d}x_3 \wedge \mathrm{d}x_1, \mathrm{d}x_1 \wedge \mathrm{d}x_2)$. Chacune de ces 2-formes étant assimilable à une mesure d'aire infinitésimale[(64)], la quantité $\mathbf{ds}$ peut être interprétée comme un élément d'aire infinitésimal.

13.4.3 Intégrale d'une p-forme sur un sous-ensemble de dimension p de $\mathbb{R}^n$

Nous avons défini p. 968 l'intégrale d'une 1-forme de $\mathbb{R}^n$ sur une courbe paramétrée de $\mathbb{R}^n$. La définition 13.17 permet d'intégrer une 2-forme de $\mathbb{R}^n$ sur une surface paramétrée de $\mathbb{R}^n$. Plus généralement, on peut définir l'intégrale d'une p-forme de $\mathbb{R}^n$ sur des sous-ensembles de $\mathbb{R}^n$ appelés sous-variétés de dimension p de $\mathbb{R}^n$. Sans entrer dans les détails, une sous-variété de dimension 1 est une courbe et une sous-variété de dimension 2 est une surface. La notion générale de sous-variété ne sera pas abordée dans ce cours, nous allons uniquement donner quelques-unes de leurs propriétés, utiles pour la suite du chapitre, en les illustrant dans le cas $p \leqslant 3$ et $n = 3$.

Exemples

1. Les sous-variétés de dimension 0 de $\mathbb{R}^3$ correspondent aux points de $\mathbb{R}^3$.

2. La spirale $V_1 = \{(\cos(t), \sin(t), t) \in \mathbb{R}^3 \ ; \ t \in [0, 2\pi]\}$ est une sous-variété de dimension 1 de $\mathbb{R}^3$.

3. Le « morceau » de paraboloïde $V_2 = \{(x_1, x_2, x_1^2 + x_2^2) \ ; \ x_1^2 + x_2^2 \leqslant 1\}$ est une sous-variété de dimension 2 de $\mathbb{R}^3$.

4. La boule fermée $V_3 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \ ; \ x_1^2 + x_2^2 + x^3 \leqslant 1\}$ est une sous-variété de dimension 3 de $\mathbb{R}^3$.

---

(64) En effet, la 2-forme $\mathrm{d}x_i \wedge \mathrm{d}x_j$ est l'application $(\mathbf{u}, \mathbf{v}) \in \mathbb{R}^n \times \mathbb{R}^n \mapsto \det \begin{pmatrix} u_1 & v_1 \\ u_2 & v_2 \end{pmatrix}$ et nous avons vu p. 986 que le déterminant $\det \begin{pmatrix} u_1 & v_1 \\ u_2 & v_2 \end{pmatrix}$ s'interprète comme l'aire algébrique du parallélogramme construit à partir des vecteurs $\mathbf{u}$ et $\mathbf{v}$.

Certaines sous-variétés possèdent des points particuliers qui forment ce que l'on appelle le bord de la sous-variété. Si une sous-variété V est de dimension $p \in \mathbb{N}^*$ et possède un bord, noté ∂V, alors celui-ci est une sous-variété de dimension $p-1$. Reprenons les exemples précédents.

Exemples

1. La spirale V_1 est une sous-variété à bord. Son bord correspond aux deux extrémités de la courbe, c'est-à-dire que l'on a $\partial V_1 = \{(1,0,0),(1,0,2\pi)\}$, voir la figure 9 à droite.

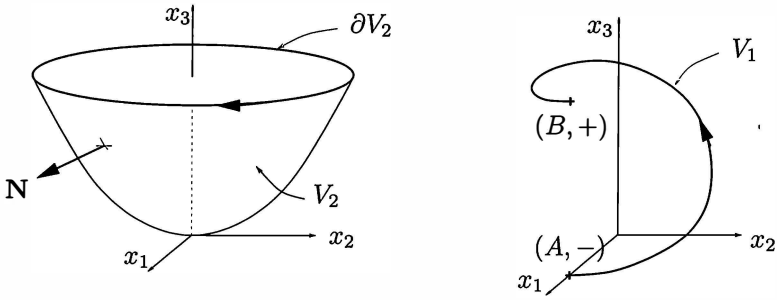


Fig. 9 Orientation d'une variété de dimension 2 (à gauche) ou de dimension 1 (à droite) et l'orientation induite de son bord.

2. Le « morceau » de paraboloïde V_2 est une sous-variété dont le bord ∂V_2 est le cercle dans le plan $x_3 = 1$ de centre $(0,0,1)$ et de rayon 1, voir la figure 9 à gauche. L'ensemble ∂V_2 est une sous-variété de dimension 1 qui n'a pas de bord.

3. Le bord ∂V_3 de la boule fermée V_3 est la sphère de $\mathbb{R}^3$ de centre $(0,0,0)$ et de rayon 1. L'ensemble ∂V_3 est une sous-variété de dimension 2 qui n'a pas de bord.

On a vu que l'orientation du domaine jouait un rôle dans les calculs d'intégrales, que ce soit pour intégrer une 1-forme sur une courbe ou pour intégrer une 2-forme sur une surface. Pour intégrer une p-forme sur une sous-variété de dimension p, il faut que celle-ci soit orientable. On peut noter que si une sous-variété possédant un bord est orientable alors il existe une orientation « induite » pour le bord de cette variété.

Exemples

1. On a vu que toutes les courbes régulières sont orientables (l'orientation étant donnée par le vecteur tangent). Le bord d'une sous-variété de dimension 1 étant réduit aux deux extrémités de la courbe, on oriente ce bord en affectant les signes $+$ et $-$ aux deux extrémités (voir la figure 9 à droite dans le cas de la variété V_1).

2. Une surface orientable peut être orientée en imposant le sens du vecteur normal en tout point. L'orientation induite sur le bord d'une telle surface s'obtient en appliquant la règle du bonhomme d'Ampère : « laisser la surface sur la gauche lorsqu'on parcourt le bord » (voir la figure 9 à gauche dans le cas de la variété V_2).

13.4.4 Différentielle extérieure

Rappelons[65] que U étant un ouvert de $\mathbb{R}^n$ et k un entier naturel, on note $\Lambda^k(U)$ l'ensemble des k-formes sur U. En particulier, $\Lambda^0(U)$ correspond à l'ensemble des applications de U dans $\mathbb{R}$. Afin de regrouper toutes les formes différentielles dans un seul ensemble, on définit

$$\Lambda(U) = \bigcup_{k\in\mathbb{N}} \Lambda^k(U).$$

PROPOSITION 13.12 *Il existe un unique opérateur* d *nommé* différentielle *qui à toute* $(k-1)$*-forme* ω*, où* $k \in \mathbb{N}^*$*, de classe* $\mathcal{C}^1$ *fait correspondre une* k*-forme* $\mathrm{d}\omega$ *en satisfaisant aux conditions suivantes.*

1. *L'opérateur* d *est* $\mathbb{R}$*-linéaire, c'est-à-dire*

$$\forall(\lambda,\mu)\in\mathbb{R}^2 \quad \forall(\omega,\widetilde{\omega})\in\Lambda(U) \quad \mathrm{d}(\lambda\omega+\mu\widetilde{\omega}) = \lambda\,\mathrm{d}\omega + \mu\,\mathrm{d}\widetilde{\omega}.$$

2. *L'opérateur* d *associe à toute application* f *de* U *dans* $\mathbb{R}$ *de classe* $\mathcal{C}^1$ *sa différentielle, c'est-à-dire*

$$\forall f\in\mathcal{C}^1(U,\mathbb{R}) \quad \mathrm{d}f = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}\mathrm{d}x_j.$$

3. *Pour tous* $f\in\mathcal{C}^1(U,\mathbb{R})$ *et* $\omega\in\Lambda(U)$ *on a* $\mathrm{d}(f\omega) = \mathrm{d}f\wedge\omega + f\mathrm{d}\omega$.
4. *Pour tout* $\omega\in\Lambda(U)$ *de classe* $\mathcal{C}^2$ *on a* $\mathrm{d}(\mathrm{d}\omega) = 0$.

Comme l'indique la propriété 2 de la proposition 13.12, cet opérateur d prolonge l'« opérateur de différentiation » vu pour les fonctions de plusieurs variables (voir le chap. 8). Étudions plus précisément comment il « transforme » une 1-forme en une 2-forme. Considérons ω une 1-forme définie sur un ouvert U de $\mathbb{R}^n$, de classe $\mathcal{C}^1$. D'après la proposition 13.6, il existe n applications $\alpha_1,\ldots,\alpha_n$ de U dans $\mathbb{R}$ telles que

$$\omega = \sum_{j=1}^{n} \alpha_j\,\mathrm{d}x_j. \tag{20}$$

En utilisant les règles de calcul données pour l'opérateur différentiel d dans la proposition 13.12, on obtient

$$\mathrm{d}\omega = \mathrm{d}\left(\sum_{j=1}^{n}\alpha_j\,\mathrm{d}x_j\right) = \sum_{j=1}^{n}\mathrm{d}(\alpha_j\,\mathrm{d}x_j) = \sum_{j=1}^{n}\big(\mathrm{d}\alpha_j\wedge\mathrm{d}x_j + \alpha_j\,\mathrm{d}(\mathrm{d}x_j)\big)$$
$$= \sum_{j=1}^{n}\left(\sum_{i=1}^{n}\frac{\partial\alpha_j}{\partial x_i}\,\mathrm{d}x_i\wedge\mathrm{d}x_j\right).$$

[65] Voir la définition 13.15 p. 988.

Puisque pour tout $(i,j) \in \{1,\dots,n\}^2$ on a $\mathrm{d}x_i \wedge \mathrm{d}x_j = -\mathrm{d}x_j \wedge \mathrm{d}x_i$, on en déduit que

$$\mathrm{d}\omega = \sum_{i=1}^{n} \sum_{j=1}^{i-1} \left(\frac{\partial \alpha_j}{\partial x_i} - \frac{\partial \alpha_i}{\partial x_j} \right) \mathrm{d}x_i \wedge \mathrm{d}x_j. \tag{21}$$

Ainsi, la différentielle d'une 1-forme donnée par la relation (20) est la 2-forme donnée par la relation (21).

Par exemple, la différentielle de la 1-forme de $\mathbb{R}^2$ définie par $\omega = \alpha_1 \mathrm{d}x_1 + \alpha_2 \mathrm{d}x_2$ vaut $\mathrm{d}\omega = \left(\dfrac{\partial \alpha_2}{\partial x_1} - \dfrac{\partial \alpha_1}{\partial x_2} \right) \mathrm{d}x_1 \wedge \mathrm{d}x_2$.

L'opérateur d « transforme » de la même manière les 2-formes en 3-formes. Considérons une 2-forme ω sur $\mathbb{R}^3$. On sait d'après la proposition 13.11 qu'il existe 3 applications α_1, α_2 et α_3 de $\mathbb{R}^2$ dans $\mathbb{R}$ telles que

$$\omega = \alpha_1 \mathrm{d}x_2 \wedge \mathrm{d}x_3 + \alpha_2 \mathrm{d}x_3 \wedge \mathrm{d}x_1 + \alpha_3 \mathrm{d}x_1 \wedge \mathrm{d}x_2.$$

Compte tenu du fait que $\mathrm{d}x_i \wedge \mathrm{d}x_j \wedge \mathrm{d}x_k = 0$ dès que deux indices parmi $\{i,j,k\}$ sont égaux, d'après la proposition 13.12, la 3-forme $\mathrm{d}\omega$ est donnée par

$$\mathrm{d}\omega = \left(\frac{\partial \alpha_1}{\partial x_1} + \frac{\partial \alpha_2}{\partial x_2} + \frac{\partial \alpha_3}{\partial x_3} \right) \mathrm{d}x_1 \wedge \mathrm{d}x_2 \wedge \mathrm{d}x_3.$$

En introduisant les notations

$$\mathrm{d}\mathbf{s} = \begin{pmatrix} \mathrm{d}x_2 \wedge \mathrm{d}x_3 \\ \mathrm{d}x_3 \wedge \mathrm{d}x_1 \\ \mathrm{d}x_1 \wedge \mathrm{d}x_2 \end{pmatrix} \quad \text{et} \quad \mathrm{d}v = \mathrm{d}x_1 \wedge \mathrm{d}x_2 \wedge \mathrm{d}x_3$$

on obtient formellement $\omega = \alpha \cdot \mathrm{d}\mathbf{s}$ et $\mathrm{d}\omega = \mathrm{div}(\alpha)\, \mathrm{d}v$.

Exercice 13.12 Soit ω une 1-forme de $\mathbb{R}^3$ dont la décomposition dans la famille $\{\mathrm{d}x_1, \mathrm{d}x_2, \mathrm{d}x_3\}$ est $\omega = \alpha_1 \mathrm{d}x_1 + \alpha_2 \mathrm{d}x_2 + \alpha_3 \mathrm{d}x_3$ où pour tout $i \in \{1,2,3\}$, α_i est une application de $\mathbb{R}^3$ dans $\mathbb{R}$ de classe $\mathcal{C}^1$.
1 - Déterminer l'expression de la différentielle $\mathrm{d}\omega$ en la décomposant selon la famille $\{\mathrm{d}x_2 \wedge \mathrm{d}x_3, \mathrm{d}x_3 \wedge \mathrm{d}x_1, \mathrm{d}x_1 \wedge \mathrm{d}x_2\}$.
2 - En utilisant les notations

$$\mathrm{d}\mathbf{x} = \begin{pmatrix} \mathrm{d}x_1 \\ \mathrm{d}x_2 \\ \mathrm{d}x_3 \end{pmatrix} \quad \text{et} \quad \mathrm{d}\mathbf{s} = \begin{pmatrix} \mathrm{d}x_2 \wedge \mathrm{d}x_3 \\ \mathrm{d}x_3 \wedge \mathrm{d}x_1 \\ \mathrm{d}x_1 \wedge \mathrm{d}x_2 \end{pmatrix},$$

vérifier que l'on peut écrire « formellement » $\omega = \alpha \cdot \mathrm{d}\mathbf{x}$ et $\mathrm{d}\omega = \mathrm{rot}(\alpha) \cdot \mathrm{d}\mathbf{s}$.

13.4.5 Théorème de Stokes

Imaginons un conservateur de musée voulant contrôler à chaque instant le nombre de personnes qui est à l'intérieur de son musée. Il décide d'installer des portillons automatiques à l'entrée et à la sortie du musée. Si au début de la journée le musée est vide alors le conservateur peut à tout instant de la journée avoir, grâce à cette donnée « au bord », une information globale à l'intérieur du musée[66].

De façon similaire, le théorème de Stokes permet de connaître l'intégrale (assimilable à une somme) d'une forme différentielle $\mathrm{d}\omega$ sur une variété Ω à partir de la donnée de la forme différentielle ω sur le bord $\partial\Omega$ de Ω.

THÉORÈME 13.2 (Théorème de Stokes)
Soit $(n,p) \in \mathbb{N}^* \times \mathbb{N}^*$ tel que $p \leqslant n$. Si ω est une forme différentielle de degré $p-1$ et de classe $\mathcal{C}^1$ sur un ouvert U de $\mathbb{R}^n$ et Ω une sous-variété bornée orientable de $\mathbb{R}^n$ de dimension p contenue dans U alors

$$\int_\Omega \mathrm{d}\omega = \int_{\partial\Omega} \omega. \tag{22}$$

La démonstration de cette égalité dépasse le cadre de ce cours. Son énoncé est néanmoins intéressant car le théorème de Stokes, dont la formulation est très générale, englobe un certain nombre de formules intégrales très utilisées en physique : la formule de Green-Riemann, celle de Stokes et la formule d'Ostrogradski. Ces différentes formules correspondent au théorème 13.2 dans le cas de valeurs particulières des nombres p et n. Ce sont ces cas que nous allons détailler dans ce qui suit.

Remarque On a vu à plusieurs reprises dans ce chapitre que l'on sait intégrer des objets de « dimension » k sur des ensembles de même « dimension ». On peut s'assurer que dans la relation (22), ce principe est respecté : $\mathrm{d}\omega$ est une p-forme que l'on intègre sur Ω qui est de dimension p ; de même ω est une $(p-1)$-forme que l'on intègre sur $\partial\Omega$ qui est de dimension $p-1$. ◇

Le théorème de Stokes pour les fonctions de la variable réelle

Écrivons le théorème de Stokes lorsque $n = 1$. L'entier $p \in \mathbb{N}^*$ doit vérifier $p \leqslant n$ et on a par conséquent $p = 1$. Le théorème de Stokes affirme que si ω est une forme différentielle de degré 0 et de classe $\mathcal{C}^1$ sur $\mathbb{R}$ et si Ω une sous-variété orientable de $\mathbb{R}$ de dimension 1 alors $\int_\Omega \mathrm{d}\omega = \int_{\partial\Omega} \omega$. Explicitons ce résultat.

[66] On peut remarquer que ce principe fonctionne grâce à une hypothèse implicite de conservation du nombre de personne : les gens ne peuvent ni disparaître ni apparaître dans le musée ! Un exemple où le nombre de personne à l'intérieur d'un édifice ne dépend pas uniquement des entrées et sorties de cet édifice est celui d'une maternité où la relation entre le débit aux entrée et sortie ne peut pas être si simplement lié au nombre de personnes dans la maternité. De la même façon, l'évolution de la population d'un pays ne peut pas se déduire simplement de la comptabilité des passages aux frontières...

STOKES, George-Gabriel (1819, Skreen, Irlande - 1903, Cambridge, Angleterre).



Stokes est connu pour ses recherches en mécanique des fluides, en particulier sur la modélisation de la viscosité. Au cours de ses travaux, il a été amené à analyser des variétés (courbes de l'espace, surfaces, volumes) décrivant les systèmes physiques étudiés. L'énoncé de la formule de Stokes est dû à Lord Kelvin. Stokes reprend cet énoncé dans une question d'examen pour le prix Smith à l'Université de Cambridge en 1854. Cependant la première démonstration publiée de cette formule semble avoir été celle de Hermann Hankel en 1861.

Les formes différentielles de degré 0 sur $\mathbb{R}$ sont les applications de $\mathbb{R}$ dans $\mathbb{R}$ et les sous-variétés de $\mathbb{R}$ de dimension 1 sont des segments (plus exactement des réunions de segments) du type $]a, b[$. La différentielle d'une application $f \in \mathcal{C}^1(\mathbb{R}, \mathbb{R})$ est la 1-forme $\mathrm{d}f = f' \,\mathrm{d}x$ et le bord du segment $]a, b[$ est le couple orienté[67] de points $\{(a, -), (b, +)\}$. Le théorème de Stokes s'énonce dans ce cas particulier de la manière suivante : si f est une application de $\mathbb{R}$ dans $\mathbb{R}$ de classe $\mathcal{C}^1$ et $]a, b[$ un segment de $\mathbb{R}$ alors

$$\int_a^b f'(x) \,\mathrm{d}x = f(b) - f(a).$$

Remarque Écrite ainsi, la formule de Stokes suggère que l'intégration est la « procédure inverse » de la dérivation. L'intégrale d'une application f de $\mathbb{R}$ dans $\mathbb{R}$ entre deux réels a et b correspond à l'aire algébrique de la surface comprise entre la représentation graphique de l'application f et l'axe des abscisses. La dérivée d'une application de $\mathbb{R}$ dans $\mathbb{R}$ correspond à la pente de la tangente à la courbe représentant f. Le théorème de Stokes qui relie les deux notions d'intégration et de dérivation est parfois appelé *théorème fondamental de l'analyse.* ◇

Formule de Green-Riemann

Explicitons le théorème de Stokes lorsque $n = 2$; l'entier p peut prendre les deux valeurs $p = 1$ ou $p = 2$.

Lorsque $p = 1$, l'ensemble Ω est une courbe, que l'on notera C, du plan $\mathbb{R}^2$ dont le bord est réduit à deux points A et B (que l'on peut orienter grâce à l'orientation de la courbe C). Les 0-formes de $\mathbb{R}^2$ étant assimilées aux applications de $\mathbb{R}^2$ dans $\mathbb{R}$, la formule de Stokes s'écrit

$$\int_C \mathrm{d}f = f(B) - f(A).$$

On retrouve la proposition 13.8 p. 970.

[67] Orienter un point signifie ici lui attribuer un signe. La notation $\{(a, -), (b, +)\}$ signifie que l'on donne le signe positif au point b et négatif au point a, voir p. 994.

GREEN, George (1793, Sneinton, Angleterre - 1841, Sneinton).



Physicien anglais autodidacte (il n'a passé qu'un an environ à l'école, entre 8 et 9 ans et fut apprenti boulanger), Green intégra l'Université de Cambridge comme étudiant en 1833 à l'âge de 40 ans. Il avait écrit en 1828 un *Essai sur l'application de l'analyse mathématique aux théories de l'électricité et du magnétisme.* Il enseigna à Cambridge jusqu'à ce qu'il tombe malade en 1840. Ses travaux sur l'électricité, le magnétisme et l'électrostatique le conduirent à d'importants résultats sur l'intégration des formes différentielles sur un contour du plan mais ceux-ci furent peu reconnus par la communauté scientifique de son vivant.

Intéressons-nous au cas $p = 2$. Une sous-variété Ω de dimension 2 dans le plan $\mathbb{R}^2$ est un sous-ensemble de $\mathbb{R}^2$, que l'on notera D, contenant un ouvert non vide et possédant un bord qui est une courbe C du plan, régulière et sans point double. Lorsque $p = 2$, la forme différentielle ω est une 1-forme de $\mathbb{R}^2$, de classe $\mathcal{C}^1$. On sait[68] qu'elle s'écrit $\omega = \alpha_1 \mathrm{d}x_1 + \alpha_2 \mathrm{d}x_2$ où α_1 et α_2 sont deux applications de classe $\mathcal{C}^1$ de $\mathbb{R}^2$ dans $\mathbb{R}$. On a vu à la p. 997 que la 2-forme $\mathrm{d}\omega$ est

$$\mathrm{d}\omega = \left(\frac{\partial \alpha_2}{\partial x_1} - \frac{\partial \alpha_1}{\partial x_2}\right) \mathrm{d}x_1 \wedge \mathrm{d}x_2.$$

Dans ce cas ($n = 2$ et $p = 2$) le théorème de Stokes prend le nom de formule de Green-Riemann.

PROPOSITION 13.13 (Formule de Green-Riemann)
Soit D un sous-ensemble borné de $\mathbb{R}^2$ possédant un bord C qui est une courbe régulière et sans point double. Si α_1 et α_2 sont deux fonctions de classe $\mathcal{C}^1$ sur $\overline{D}$ alors

$$\int_D \left(\frac{\partial \alpha_2}{\partial x_1} - \frac{\partial \alpha_1}{\partial x_2}\right) \mathrm{d}x_1 \mathrm{d}x_2 = \int_C (\alpha_1 \mathrm{d}x_1 + \alpha_2 \mathrm{d}x_2).$$

Remarque Dans cette égalité, $\mathrm{d}x_1\mathrm{d}x_2$ est vue comme une notation. On devrait écrire $\mathrm{d}x_1 \wedge \mathrm{d}x_2$. On remarque que la 1-forme $\mathrm{d}x_1 \wedge \mathrm{d}x_2$ engendre l'ensemble $\Lambda^2(\mathbb{R}^2)$ et prend en compte l'orientation du sous-ensemble D de $\mathbb{R}^2$. ◇

Une application importante de la formule de Green-Riemann est donnée par le corollaire suivant.

[68] Voir la proposition 13.6 p. 956.

Corollaire 13.3 *Si D est un sous-ensemble borné de $\mathbb{R}^2$ possédant un bord qui est une courbe C régulière et sans point double alors l'aire de D est donnée par*

$$\text{Aire}(D) = \int_C x_1 \, \mathrm{d}x_2 = \int_C -x_2 \, \mathrm{d}x_1 = \frac{1}{2}\int_C (x_1 \mathrm{d}x_2 - x_2 \mathrm{d}x_1).$$

Démonstration Rappelons que l'aire de D est donnée par $\int_D \mathrm{d}x_1\mathrm{d}x_2$. Le résultat du corollaire 13.3 se démontre en utilisant la proposition 13.13 avec les applications $\alpha_1 : (x_1, x_2) \in \mathbb{R}^2 \mapsto 0$ et $\alpha_2 : (x_1, x_2) \in \mathbb{R}^2 \mapsto x_1$. Pour tout $(x_1, x_2) \in \mathbb{R}^2$ on a

$$\frac{\partial \alpha_2}{\partial x_1}(x_1, x_2) = 1 \quad \text{et} \quad \frac{\partial \alpha_1}{\partial x_2}(x_1, x_2) = 0,$$

d'où la relation

$$\int_D \mathrm{d}x_1\mathrm{d}x_2 = \int_C \alpha_2 \, \mathrm{d}x_2.$$

En notant abusivement $\int_C \alpha_2 \, \mathrm{d}x_2 = \int_C x_1 \, \mathrm{d}x_2$, on obtient la première égalité du corollaire. On obtient la seconde égalité de la même manière en choisissant cette fois-ci $\alpha_1 : (x_1, x_2) \in \mathbb{R}^2 \mapsto x_2$ et $\alpha_2 : (x_1, x_2) \in \mathbb{R}^2 \mapsto 0$. La troisième égalité s'obtient quant à elle en choisissant dans le théorème 13.13 les applications $\alpha_1 : (x_1, x_2) \in \mathbb{R}^2 \mapsto -\frac{1}{2}x_2$ et $\alpha_2 : (x_1, x_2) \in \mathbb{R}^2 \mapsto \frac{1}{2}x_1$. □

Exemple Considérons le Folium de Descartes qui est la courbe C du plan $\mathbb{R}^2$ d'équation implicite

$$x_1^3 + x_2^3 = 3x_1x_2.$$

Cette courbe fait une boucle (voir la figure 10) dont il semble *a priori* difficile de calculer l'aire. Le corollaire 13.3 permet d'effectuer ce calcul.

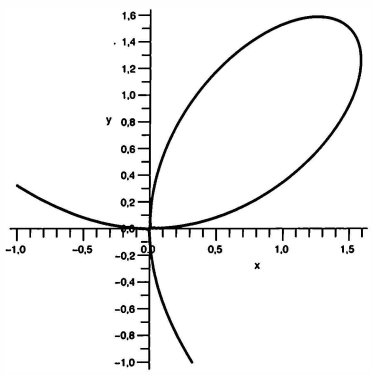


Fig. 10 Le Folium de Descartes

Dans un premier temps, on détermine un paramétrage de la courbe C en cherchant[(69)] l'intersection de C avec toutes les droites D_t d'équation $x_2 = tx_1$, $t \in \mathbb{R}$, passant par $(0,0)$. On a

$$(x_1,x_2) \in C \cap D_t \iff \Big(x_1^3 + x_2^3 = 3x_1x_2 \quad \text{et} \quad x_2 = tx_1 \Big).$$

En utilisant l'expression $x_2 = tx_1$ dans l'équation de C, on a

$$(x_1,x_2) \in C \cap D_t \iff \Big(x_1^3 + t^3x_1^3 = 3tx_1^2 \quad \text{et} \quad x_2 = tx_1 \Big).$$

Il s'agit donc de résoudre, pour $t \in \mathbb{R}$ fixé, l'équation $x_1^3 + t^3x_1^3 = 3tx_1^2$. On remarque que si $t = -1$, cette équation admet $x_1 = 0$ pour unique solution alors que si $t \neq -1$, c'est une équation polynomiale de degré 3 qui admet 0 comme racine double, l'autre racine étant donnée par $x_1 = 3t/(1+t^3)$. On a donc $C \cap D_{-1} = (0,0)$ et pour tout $t \neq -1$

$$(x_1,x_2) \in C \cap D_t \setminus (0,0) \iff \Big(x_1 = \frac{3t}{1+t^3} \quad \text{et} \quad x_2 = \frac{3t^2}{1+t^3} \Big).$$

Enfin, l'intersection de C avec la droite D_∞ d'équation $x_1 = 0$ est réduite à l'origine $(0,0)$. On en déduit[(69)] qu'un paramétrage de la courbe C est

$$\varphi : t \in \mathbb{R} \setminus \{-1\} \longmapsto \left(\frac{3t}{1+t^3}, \frac{3t^2}{1+t^3} \right) \in \mathbb{R}^2.$$

Le bord de la boucle est obtenu pour t compris entre 0 et $+\infty$. On peut alors utiliser le corollaire 13.3 pour calculer l'aire de cette boucle ; on a

$$\begin{aligned} \text{Aire(Boucle)} &= \frac{1}{2}\int_C (x_1 \mathrm{d}x_2 - x_2 \mathrm{d}x_1) = \frac{1}{2}\int_0^{+\infty} (x_1(t)x_2'(t) - x_2(t)x_1'(t))\ \mathrm{d}t \\ &= \frac{1}{2}\int_0^{+\infty} \frac{9t^2}{(1+t^3)^2}\ \mathrm{d}t = \frac{3}{2}\left[\frac{-1}{1+t^3}\right]_0^{+\infty} = \frac{3}{2}. \end{aligned}$$

L'aire de la boucle du Folium de Descartes vaut donc 3/2.

Exercice 13.13 Considérons l'ellipse d'équation $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$ où a et b sont deux réels strictement positifs.
1 - Donner un paramétrage de cette ellipse.
2 - En utilisant le corollaire 13.3, retrouver l'aire du domaine délimité par cette ellipse.

Un autre corollaire très utile de la formule de Green-Riemann est la formule de la divergence qui s'obtient de la manière suivante. Considérons l'application suivante définie sur un voisinage U d'un sous-ensemble borné D de $\mathbb{R}^2$:

$$\mathbf{u} : (x_1,x_2) \in U \longmapsto (\alpha_2(x_1,x_2), -\alpha_1(x_1,x_2))$$

(69) Le plan étant la réunion de toutes les droites passant par $(0,0)$: $\mathbb{R}^2 = \bigcup_{t\in\mathbb{R}\cup\{\infty\}} D_t$, la courbe C vérifie $C = \bigcup_{t\in\mathbb{R}\cup\{\infty\}} (C \cap D_t)$.

où α_1 et α_2 sont deux applications de classe $\mathcal{C}^1$ sur U. La divergence de $\mathbf{u}$ est l'application

$$\operatorname{div}(\mathbf{u}) = \frac{\partial \alpha_2}{\partial x_1} - \frac{\partial \alpha_1}{\partial x_2}.$$

Si le bord de l'ensemble D est une courbe régulière C sans point double alors la formule de Green-Riemann s'écrit :

$$\int_D \operatorname{div}(\mathbf{u})\, \mathrm{d}x_1\mathrm{d}x_2 = \int_C (\alpha_1 \mathrm{d}x_1 + \alpha_2 \mathrm{d}x_2).$$

Pour une telle courbe C, on sait[70] qu'il existe toujours un paramétrage $\psi : \tau \in [a,b] \mapsto (\psi_1(\tau), \psi_2(\tau)) \in \mathbb{R}^2$ pour lequel le vecteur tangent à la courbe, qui est donné par $(\psi_1'(\tau), \psi_2'(\tau))$, est de norme 1. On dit que la courbe est paramétrée par l'abscisse curviligne. Par définition de la circulation d'une 1-forme[71], on a

$$\int_C (\alpha_1 \mathrm{d}x_1 + \alpha_2 \mathrm{d}x_2) = \int_a^b \Big(\alpha_1(\psi_1(\tau), \psi_2(\tau))\psi_1'(\tau) + \alpha_2(\psi_1(\tau), \psi_2(\tau))\psi_2'(\tau)\Big)\, \mathrm{d}\tau.$$

Au point $(\psi_1(\tau), \psi_2(\tau))$, un vecteur normal unitaire à C est $\mathbf{n} = (\psi_2'(\tau), -\psi_1'(\tau))$ et la valeur prise par l'application $\mathbf{u}$ est $(\alpha_2(\psi_1(\tau), \psi_2(\tau)), -\alpha_1(\psi_1(\tau), \psi_2(\tau)))$. La relation précédente se réécrit donc de la manière suivante

$$\int_C (\alpha_1 \mathrm{d}x_1 + \alpha_2 \mathrm{d}x_2) = \int_C \mathbf{u} \cdot \mathbf{n}\, \mathrm{d}\tau.$$

Cette formule est connue sous le nom de formule de la divergence.

Corollaire 13.4 (Formule de la divergence)
Soit D un sous-ensemble borné de $\mathbb{R}^2$ dont le bord C est une courbe régulière, sans point double et dont la normale unitaire « sortante » est notée $\mathbf{n}$. Si $\mathbf{u}$ est une application de $\overline{D}$ dans $\mathbb{R}^2$ de classe $\mathcal{C}^1$ alors

$$\int_D \operatorname{div}(\mathbf{u})\, \mathrm{d}x_1\mathrm{d}x_2 = \int_C \mathbf{u} \cdot \mathbf{n}\, \mathrm{d}\tau,$$

où τ désigne l'abscisse curviligne.

Exemple L'évolution de la masse d'un fluide contenue dans un domaine $\mathcal{V}$ du plan $\mathbb{R}^2$ exprime le fait que la variation de la masse totale de fluide contenu dans $\mathcal{V}$ à l'instant t est égale au flux de masse entrant à travers le bord $\partial\mathcal{V}$ de $\mathcal{V}$. En notant $\rho(t,\mathbf{x})$ (respectivement $\mathbf{u}(t,\mathbf{x})$) la densité (respectivement la vitesse) du fluide au point $\mathbf{x}$ de l'espace (c'est-à-dire $\mathbf{x} \in \mathcal{V}$) et à l'instant $t \in \mathbb{R}$, le principe de conservation de la masse s'écrit :

$$\frac{d}{dt}\left(\int_{\mathcal{V}} \rho(t,\mathbf{x})\, \mathrm{d}\mathbf{x}\right) = -\int_{\partial\mathcal{V}} \rho(t,\mathbf{x})\mathbf{u}(t,\mathbf{x}) \cdot \mathbf{n}\, \mathrm{d}\tau.$$

[70] Voir la proposition 13.5 p. 953.
[71] Voir la définition 13.1 p. 938 et la définition 13.9 p. 969.

En utilisant la formule de la divergence, on obtient

$$\int_{\mathcal{V}} \left(\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho \mathbf{u}) \right) \, \mathrm{d}\mathbf{x} = 0.$$

Cette relation étant vraie à tout instant t et pour tout voisinage $\mathcal{V}$ d'un point x quelconque de l'ouvert sur lequel on étudie le fluide, nous obtenons une expression locale de la conservation de la masse qui est :

$$\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho \mathbf{u}) = 0.$$

Cette dernière équation, qui correspond à la forme « locale » de la conservation de la masse d'un fluide, est très utilisée en mécanique des fluides.

Formule de Stokes

Étudions le théorème de Stokes lorsque $n = 3$; les valeurs possibles pour p sont $p = 1$, $p = 2$ et $p = 3$.

Lorsque $p = 1$, l'ensemble Ω est une courbe dans l'espace $\mathbb{R}^3$: le bord de Ω est réduit à deux points A et B (que l'on peut orienter grâce à l'orientation induite par la courbe C, voir p. 994). Le cas est similaire au cas $n = 2$, $p = 1$ et la formule de Stokes s'écrit

$$\int_C \mathrm{d}f = f(B) - f(A).$$

On retrouve la proposition 13.8.

Considérons le cas $p = 2$; l'ensemble Ω est une surface de $\mathbb{R}^3$, que l'on notera S, dont le bord est une courbe[72], que l'on notera C, et ω est une 1-forme de classe $\mathcal{C}^1$ sur $\mathbb{R}^3$. Cette 1-forme s'écrit[73] $\omega = f_1 \mathrm{d}x_1 + f_2 \mathrm{d}x_2 + f_3 \mathrm{d}x_3$ où pour tout $i \in \{1, 2, 3\}$ l'application f_i est une application de $\mathbb{R}^3$ dans $\mathbb{R}$ de classe $\mathcal{C}^1$. La 2-forme $\mathrm{d}\omega$ vaut[74]

$$\mathrm{d}\omega = \left(\frac{\partial f_3}{\partial x_2} - \frac{\partial f_2}{\partial x_3} \right) \mathrm{d}x_2 \wedge \mathrm{d}x_3 + \left(\frac{\partial f_1}{\partial x_3} - \frac{\partial f_3}{\partial x_1} \right) \mathrm{d}x_3 \wedge \mathrm{d}x_1 + \left(\frac{\partial f_2}{\partial x_1} - \frac{\partial f_1}{\partial x_2} \right) \mathrm{d}x_1 \wedge \mathrm{d}x_2.$$

En utilisant les notations

$$\mathrm{d}\mathbf{x} = \begin{pmatrix} \mathrm{d}x_1 \\ \mathrm{d}x_2 \\ \mathrm{d}x_3 \end{pmatrix} \quad \text{et} \quad \mathrm{d}\mathbf{s} = \begin{pmatrix} \mathrm{d}x_2 \wedge \mathrm{d}x_3 \\ \mathrm{d}x_3 \wedge \mathrm{d}x_1 \\ \mathrm{d}x_1 \wedge \mathrm{d}x_2 \end{pmatrix},$$

[72] Le bord d'une surface peut éventuellement être vide, c'est le cas par exemple du bord d'une sphère.

[73] Voir la proposition 13.6 p. 956.

[74] Voir l'exercice 13.12 p. 997.

on peut écrire formellement $\omega = \mathbf{f} \cdot \mathrm{d}\mathbf{x}$ et $\mathrm{d}\omega = \mathrm{rot}(\mathbf{f}) \cdot \mathrm{d}\mathbf{s}$. Le théorème de Stokes donne alors lieu à la proposition suivante connue sous le nom de formule de Stokes.

PROPOSITION 13.14 (Formule de Stokes)
Soit S une surface bornée de $\mathbb{R}^3$ dont le bord C est une courbe régulière sans point double. Si f est une application de $\overline{S}$ dans $\mathbb{R}^3$ de classe $\mathcal{C}^1$ alors

$$\int_S \mathrm{rot}(f) \cdot \mathrm{d}\mathbf{s} = \int_C f \cdot \mathrm{d}\mathbf{x}.$$

Remarque Si S est une surface sans bord (par exemple si S est une sphère) alors la formule de Stokes affirme que le flux d'un rotationnel à travers S est toujours nul. ◇

Exemples

1. Calculons, en utilisant la formule de Stokes, l'intégrale

$$I = \int_C (x_2 + x_3)\ \mathrm{d}x_1 + (x_3 + x_1)\ \mathrm{d}x_2 + (x_1 + x_2)\ \mathrm{d}x_3,$$

où le cercle C est le bord de la demi-sphère S d'équation $x_1^2 + x_2^2 + x_3^2 = 1$ limitée par la condition $x_1 + x_2 + x_3 \geqslant 0$. On a $I = \int_S \mathrm{rot}(f) \cdot \mathrm{d}\mathbf{s}$ où l'application f est définie par $f : (x_1, x_2, x_3) \in \mathbb{R}^3 \mapsto (x_2 + x_3,\ x_3 + x_1,\ x_1 + x_2)$. Un calcul montre que $\mathrm{rot}(f) = \mathbf{0}$ de sorte que $I = 0$.

2. Calculons à l'aide de la formule de Stokes la circulation J de l'application $f : (x_1, x_2, x_3) \in \mathbb{R}^3 \mapsto (2x_1 + x_2x_3,\ x_1x_3,\ x_1x_2)$ le long du cercle de centre $(0, 0)$ de rayon 1 inclus dans le plan d'équation $x_3 = 0$. Par définition de la circulation de f le long de C, on a $J = \int_C f \cdot \mathrm{d}\mathbf{x}$. En utilisant la formule de Stokes et le fait que $\mathrm{rot}(f) = \mathbf{0}$, on en déduit là aussi que $J = 0$.

Formule d'Ostrogradski

Intéressons-nous au théorème de Stokes dans le cas où $n = 3$ et $p = 3$. L'ensemble Ω est un sous-ensemble de $\mathbb{R}^3$ que l'on notera B et le bord de Ω est une surface que l'on notera S. Une 2-forme sur $\mathbb{R}^3$ est engendrée[(75)] par la famille $\{\mathrm{d}x_2 \wedge \mathrm{d}x_3, \mathrm{d}x_3 \wedge \mathrm{d}x_1, \mathrm{d}x_1 \wedge \mathrm{d}x_2\}$ et s'écrit

$$\omega = f_1 \mathrm{d}x_2 \wedge \mathrm{d}x_3 + f_2 \mathrm{d}x_3 \wedge \mathrm{d}x_1 + f_3 \mathrm{d}x_1 \wedge \mathrm{d}x_2,$$

où, pour tout $i \in \{1, 2, 3\}$, f_i est une application de $\mathbb{R}^3$ dans $\mathbb{R}$ de classe $\mathcal{C}^1$. La 3-forme $\mathrm{d}\omega$ s'écrit

$$\mathrm{d}\omega = \mathrm{d}f_1 \wedge \mathrm{d}x_2 \wedge \mathrm{d}x_3 + \mathrm{d}f_2 \wedge \mathrm{d}x_3 \wedge \mathrm{d}x_1 + \mathrm{d}f_3 \wedge \mathrm{d}x_1 \wedge \mathrm{d}x_2,$$

(75) Voir la proposition 13.11 p. 989.

OSTROGRADSKI, Mikhaïl Vassilievitch (1801, Pachenna - 1861, Poltava).



Né en Ukraine, Ostrogradski, après avoir commencé ses études de mathématiques à l'Université de Kharkov, vint les poursuivre à Paris. Entre 1824 et 1832, il présente à l'Académie des Sciences de Paris de nombreux mémoires contenant des découvertes nouvelles dans le domaine de la physique mathématique et du calcul intégral, ainsi qu'un mémoire sur l'hydrodynamique. Scientifique éminent dans le domaine de la mécanique céleste, il s'intéresse également au problème des petites déformations des corps élastiques, aux mouvements des corps dans un milieu visqueux et aux conditions d'équilibre des fluides incompressibles.

où $\mathrm{d}f_i = \frac{\partial f_i}{\partial x_1}\mathrm{d}x_1 + \frac{\partial f_i}{\partial x_2}\mathrm{d}x_2 + \frac{\partial f_i}{\partial x_3}\mathrm{d}x_3$ pour $i \in \{1,2,3\}$. On a donc

$$\mathrm{d}\omega = \left(\frac{\partial f_1}{\partial x_1} + \frac{\partial f_2}{\partial x_2} + \frac{\partial f_3}{\partial x_3}\right) \mathrm{d}x_1 \wedge \mathrm{d}x_2 \wedge \mathrm{d}x_3.$$

Si on utilise les notations

$$f = \begin{pmatrix} f_1 \\ f_2 \\ f_3 \end{pmatrix}, \quad \mathrm{d}\mathbf{s} = \begin{pmatrix} \mathrm{d}x_2 \wedge \mathrm{d}x_3 \\ \mathrm{d}x_3 \wedge \mathrm{d}x_1 \\ \mathrm{d}x_1 \wedge \mathrm{d}x_2 \end{pmatrix} \quad \text{et} \quad \mathrm{d}v = \mathrm{d}x_1 \wedge \mathrm{d}x_2 \wedge \mathrm{d}x_3$$

alors on a $\omega = f \cdot \mathrm{d}\mathbf{s}$ et $\mathrm{d}\omega = \mathrm{div}(f)\ \mathrm{d}v$. Le théorème de Stokes fournit la formule d'Ostrogradski.

PROPOSITION 13.15 (Formule d'Ostrogradski)
Soit B un sous-ensemble borné de $\mathbb{R}^3$ dont le bord S est une surface régulière sans point double. Si f est une application de $\overline{B}$ dans $\mathbb{R}^3$ de classe C^1 alors

$$\int_B \mathrm{div}(f)\ \mathrm{d}v = \int_S f \cdot \mathrm{d}\mathbf{s}.$$

La formule d'Ostrogradski permet entre autres d'exprimer le volume d'un sous-ensemble B borné de $\mathbb{R}^3$ à l'aide de l'intégrale de surface d'une 2-forme sur le bord de B. Considérons l'application $f : (x_1,x_2,x_3) \in \mathbb{R}^3 \mapsto (x_1,0,0) \in \mathbb{R}^3$. Puisque $\mathrm{div}(f) = 1$, on a $\int_B \mathrm{div}(f)\ \mathrm{d}v = \mathrm{Vol}(B)$. Supposons que le bord S de l'ensemble B soit une surface régulière sans point double. On a vu à la p. 993 que $\int_S f \cdot \mathrm{d}\mathbf{s} = \int_S x_1 \mathrm{d}x_2 \wedge \mathrm{d}x_3$. La formule d'Ostrogradski s'écrit

$$\mathrm{Vol}(B) = \int_S x_1\ \mathrm{d}x_2 \wedge \mathrm{d}x_3. \tag{23}$$

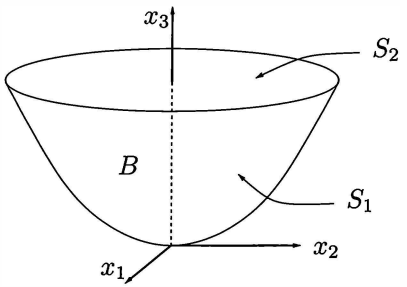
De même, en considérant l'application $f : (x_1, x_2, x_3) \in \mathbb{R}^3 \mapsto (0, x_2, 0) \in \mathbb{R}^3$ puis l'application $f : (x_1, x_2, x_3) \in \mathbb{R}^3 \mapsto (0, 0, x_3) \in \mathbb{R}^3$, on montré que l'on a aussi

$$\mathrm{Vol}(B) = \int_S x_2 \, \mathrm{d}x_3 \wedge \mathrm{d}x_1 = \int_S x_3 \, \mathrm{d}x_1 \wedge \mathrm{d}x_2.$$

Exemple Calculons le volume de l'ensemble

$$B = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \ ; \ x_1^2 + x_2^2 \leqslant x_3 \leqslant 1\}$$

représenté sur la figure ci-dessous.



Le bord de B est composé d'un « morceau » S_1 du paraboloïde d'équation $x_3 = x_1^2 + x_2^2$ et du disque S_2 de centre $(0, 0, 1)$ et de rayon 1 dans le plan d'équation $x_3 = 1$.

La relation (23) permet d'exprimer le volume de B de la manière suivante

$$\mathrm{Vol}(B) = \int_{S_1} x_1 \, \mathrm{d}x_2 \wedge \mathrm{d}x_3 + \int_{S_2} x_1 \, \mathrm{d}x_2 \wedge \mathrm{d}x_3.$$

On a établi à la p. 993 que $\int_{S_1} x_1 \mathrm{d}x_2 \wedge \mathrm{d}x_3 = \pi/2$ et dans l'exercice 13.11 que $\int_{S_2} x_1 \mathrm{d}x_2 \wedge \mathrm{d}x_3 = 0$. On en conclut que $\mathrm{Vol}(B) = \pi/2$.

Remarque La formule d'Ostrogradski peut être considérée comme la généralisation à $\mathbb{R}^3$ de la formule de la divergence dans $\mathbb{R}^2$ (voir le corollaire 13.4). En effet, on peut écrire la formule d'Ostrogradski sous la forme

$$\int_B \mathrm{div}(f) \, \mathrm{d}v = \int_S f \cdot \mathbf{n} \, \mathrm{d}\sigma,$$

le vecteur $\mathbf{n}$ étant le vecteur normal unitaire sortant à la surface S et $\mathrm{d}\sigma$ une mesure « adaptée »[(76)] sur cette surface. ◇

La formule d'Ostrogradski est très utilisée en analyse numérique pour écrire des équations aux dérivées partielles sous des formes dites « faibles ». Ces formulations sont basées sur les relations suivantes.

[(76)] La notion de mesure sort du cadre de ce cours.

Corollaire 13.5 *Soit B un sous-ensemble borné de $\mathbb{R}^3$ dont le bord S est une surface régulière sans point double. Si f est une application de $\overline{B}$ dans $\mathbb{R}^3$ de classe $\mathcal{C}^2$ et g une application de $\overline{B}$ dans $\mathbb{R}^3$ de classe $\mathcal{C}^1$ alors*

$$\int_B \Delta(f)\ g\ \mathrm{d}v = -\int_B \nabla f \cdot \nabla g\ \mathrm{d}v + \int_S g\ \frac{\partial f}{\partial \mathbf{n}}\ \mathrm{d}\sigma,$$

où $\dfrac{\partial f}{\partial \mathbf{n}}$ correspond à la dérivée de l'application f dans la direction[77] *du vecteur normal $\mathbf{n}$.*

Démonstration D'après les propriétés[78] des opérateurs différentiels div, ∇ et Δ on a

$$\mathrm{div}(g\ \nabla f) = g\ \mathrm{div}(\nabla f) + \nabla g \cdot \nabla f = \Delta(f)\ g + \nabla f \cdot \nabla g.$$

Étant donné que d'après les hypothèses du corollaire 13.5, tous les termes apparaissant dans la relation ci-dessus sont continus sur $\overline{B}$, on obtient en intégrant sur B

$$\int_B \mathrm{div}(g\ \nabla f)\ \mathrm{d}v = \int_B \Delta(f)\ g\ \mathrm{d}v + \int_B \nabla f \cdot \nabla g\ \mathrm{d}v.$$

D'après la formule d'Ostrogradski, on a

$$\int_S g\ \nabla f \cdot \mathbf{n}\ \mathrm{d}\sigma = \int_B \Delta(f)\ g\ \mathrm{d}v + \int_B \nabla f \cdot \nabla g\ \mathrm{d}v.$$

Par définition de la dérivée directionnelle d'une application, voir la définition 10.5 p. 627, le produit scalaire $\nabla f \cdot \mathbf{n}$ correspond à la dérivée de l'application f dans la direction du vecteur normal $\mathbf{n}$. Autrement dit, on a l'égalité $\nabla f \cdot \mathbf{n} = \dfrac{\partial f}{\partial \mathbf{n}}$ ce qui achève la preuve. □

Exemple Le corollaire 13.5 permet d'obtenir de nombreux résultats en analyse. Par exemple, il permet d'établir que la seule application $f \in \mathcal{C}^2(\overline{B}, \mathbb{R})$, où B est un ouvert de $\mathbb{R}^3$ dont le bord S est supposé être une surface régulière sans point double, qui soit la solution du système suivant

$$\Delta f = 0 \quad \text{sur } B \qquad \text{et} \qquad f = 0 \quad \text{sur } S \tag{24}$$

est l'application nulle $f = 0$.

D'une part il est immédiat que l'application nulle est une solution du système (24). D'autre part, si $f \in \mathcal{C}^2(\overline{B}, \mathbb{R})$ est une solution du système (24) alors, en utilisant le corollaire 13.5 avec $g = f$ on en déduit que

$$0 = \int_B \Delta(f)\ f\ \mathrm{d}v = -\int_B |\nabla f|^2\ \mathrm{d}v + \int_S f\ \frac{\partial f}{\partial \mathbf{n}}\ \mathrm{d}\sigma = -\int_B |\nabla f|^2\ \mathrm{d}v.$$

[77] Voir la définition 10.5 p. 627.

[78] Voir la proposition 10.16 p. 661 et la proposition 10.17 p. 661.

Par conséquent[79], $\nabla f = 0$ sur B ce qui implique[80] que l'application f est constante sur chaque composante connexe de B. Étant nulle sur la frontière S et continue sur $\overline{B}$ on en conclut qu'elle est nulle sur B.

13.5 Exercices de synthèse

Exercice 13.14 *Navigation sur la terre*

Dans tout cet exercice, la terre est notée $\mathcal{S}$ et est assimilée à une sphère de $\mathbb{R}^3$ de centre $\mathbf{0}$ et de rayon R. Les points de la terre peuvent être représentés par leurs coordonnées sphériques θ et φ, θ représentant la longitude et φ la latitude (en radians) ou par leurs coordonnées cartésiennes x, y et z, la correspondance se faisant par

$$x = R\cos(\theta)\cos(\varphi), \quad y = R\sin(\theta)\cos(\varphi), \quad z = R\sin(\varphi).$$

La norme $\|\cdot\|$ désigne la norme euclidienne de $\mathbb{R}^3$.

Pour se déplacer sur la surface de la terre, un des moyens les plus anciens consiste à utiliser une boussole et à garder un cap constant (c'est-à-dire un angle constant avec les parallèles). Autrement dit, à l'aide d'une boussole il est facile pour un avion ou un bateau de suivre des courbes tracées sur la terre appelées des loxodromies. Une loxodromie d'angle α et passant par le point de la terre de coordonnées (θ_1, φ_1) est donnée[81] par la relation suivante entre la latitude et la longitude :

$$\ln\left(\tan\left(\frac{\varphi}{2}+\frac{\pi}{4}\right)\right) - \ln\left(\tan\left(\frac{\varphi_1}{2}+\frac{\pi}{4}\right)\right) = (\theta-\theta_1)\tan(\alpha). \tag{25}$$

On suppose qu'un avion part du point de la terre de coordonnées (θ_1, φ_1) et qu'il suit suit une loxodromie d'angle $\alpha \in]-\pi/2, \pi/2[$ avec $\alpha \neq 0$.

1 - Donner un arc paramétré F représentant la trajectoire de l'avion à l'aide de l'angle φ. On pourra, pour simplifier, noter $\theta(\varphi)$ la valeur de θ obtenue en fonction de φ à partir de la relation (25).

2 - Montrer que la longueur de la trajectoire que parcourt l'avion entre son point de départ (θ_1, φ_1) et le point d'arrivée (θ_2, φ_2) s'écrit

$$L = C\left|\frac{\varphi_2-\varphi_1}{\sin(\alpha)}\right|$$

où l'on déterminera la constante C (qui ne dépend que de R).

---

[79] Voir la proposition 12.8 p. 874.

[80] Voir la proposition 10.19 p. 682.

[81] On a montré lors de l'exercice 11.15 que les coordonnées (θ, φ) des points sur une courbe ayant un angle constant α avec les parallèles vérifient la relation (25).

Le rayon de la terre est approximativement $R = 6370$ km. Le point P représente la ville de Lyon et le point Q représente la ville de Wuhan en Chine. Les coordonnées de ces deux villes sont données (en degrés) par

$$\theta_1 = 4, \quad \varphi_1 = 46, \quad \theta_2 = 110 \quad \text{et} \quad \varphi_2 = 30.$$

3 - Si un avion décide d'aller de Lyon à Wuhan en suivant une loxodromie, quel angle (constant) serait-il le plus judicieux de suivre sur sa boussole[82] ?

4 - Quelle distance parcourra cet avion entre Lyon et Wuhan ?

Exercice 13.15 ***Flot conservant le volume***

Soit f une application de $\mathbb{R}^n$ dans $\mathbb{R}^n$ de classe $\mathcal{C}^1$. Résoudre l'équation différentielle

$$\mathbf{x}'(t) = f(\mathbf{x}(t)) \quad \forall t \in \mathbb{R} \tag{26}$$

consiste à déterminer n applications $x_1, \ldots, x_n$ de $\mathbb{R}$ dans $\mathbb{R}$ satisfaisant pour tout $i \in \{1, \ldots, n\}$ et pour tout $t \in \mathbb{R}$ la relation $x_i'(t) = f_i(x_1(t), \ldots, x_n(t))$. On suppose que

$$\forall \mathbf{x}_0 \in \mathbb{R}^n \quad \exists! \ \mathbf{x} \in \left(\mathcal{C}^2(\mathbb{R}, \mathbb{R})\right)^n \quad \Big(\ \forall t \in \mathbb{R} \quad \mathbf{x}'(t) = f(\mathbf{x}(t)) \ \Big) \quad \text{et} \quad \mathbf{x}(0) = \mathbf{x}_0,$$

autrement dit, on suppose que pour tout $\mathbf{x}_0$ de $\mathbb{R}^n$, il existe une unique solution du système (26) vérifiant la condition initiale $\mathbf{x}(0) = \mathbf{x}_0$. Pour tout $t \in \mathbb{R}$, on note g_t l'application de $\mathbb{R}^n$ dans $\mathbb{R}^n$ définie par

$$g_t : \mathbf{x}_0 \in \mathbb{R}^n \longmapsto \mathbf{x}(t) \in \mathbb{R}^n \quad \text{où } \mathbf{x} \text{ est la solution de (26) vérifiant } \mathbf{x}(0) = \mathbf{x}_0.$$

On admettra que pour tout $t \in \mathbb{R}$ l'application g_t est de classe $\mathcal{C}^1$ sur $\mathbb{R}^n$. L'ensemble des applications g_t où $t \in \mathbb{R}$ s'appelle le *flot* associé à l'application f.

PARTIE 1 - ÉTUDE D'UN EXEMPLE

1 - Déterminer la solution de l'équation différentielle

$$y''(t) + y(t) = 0 \quad \forall t \in \mathbb{R} \tag{27}$$

telle que $y(0) = x_{01}$ et $y'(0) = x_{02}$ où x_{01} et x_{02} sont deux réels donnés.

2 - Écrire l'équation (27) sous la forme du système différentiel (26).

3 - En déduire le flot associé à l'équation (27).

4 - Montrer que le flot associé à l'équation (27), muni de la loi de composition des applications, correspond au groupe des rotations[83] du plan $\mathbb{R}^2$.

---

[82] Sur une boussole, 0 degré correspond au Nord alors que 90 degrés correspond à l'ouest.

[83] Voir p. 497 et la définition 8.21 p. 498.

Partie 2 - Étude du flot

Revenons au cas général et considérons le système différentiel (26). On admet que le flot $\{g_t\ ;\ t \in \mathbb{R}\}$ est un groupe pour la loi de composition des applications.

5 - Montrer que pour tout $\mathbf{x}_0 \in \mathbb{R}^n$ on a

$$g_t(\mathbf{x}_0) = \mathbf{x}_0 + tf(\mathbf{x}_0) + \mathcal{O}_0(t^2).$$

Soit D un sous-ensemble borné de $\mathbb{R}^n$. Pour tout $t \in \mathbb{R}$, on note $D(t)$ l'image de D par l'application g_t et $v(t)$ le volume de l'ensemble $D(t)$.

6 - Montrer que $v(t) = \int_{D(0)} |\mathcal{J}_{g_t}(\mathbf{x})|\ \mathrm{d}\mathbf{x}$ où $\mathcal{J}_{g_t}$ désigne le jacobien de l'application g_t.

7 - Montrer que pour toute matrice $\mathrm{A} \in \mathcal{M}_n(\mathbb{R})$ on a

$$\det(\mathrm{Id} + t\mathrm{A}) = 1 + t\,\mathrm{Tr}(\mathrm{A}) + \mathcal{O}_0(t^2).$$

8 - En déduire que pour tout $\mathbf{x}_0 \in \mathbb{R}^n$ on a

$$\mathcal{J}_{g_t}(\mathbf{x}_0) = 1 + t\ \mathrm{div} f(\mathbf{x}_0) + \mathcal{O}_0(t^2).$$

9 - Montrer que si l'application f est à divergence nulle alors le flot conserve les volumes, c'est-à-dire que $v(t) = v(0)$ pour tout $t \in \mathbb{R}$.

Partie 3 - Théorème du retour de Poincaré

Soient U un sous-ensemble de $\mathbb{R}^n$ et g une application de U à valeurs dans U. On suppose que g est bijective, continue et qu'elle conserve les volumes.
Pour tout ouvert V non vide inclus dans U, on note $g(V)$ l'image de V par l'application g et on définit par récurrence $g^n(V)$ pour $n \in \mathbb{N} \setminus \{0, 1\}$, comme étant l'image de $g^{n-1}(V)$ par l'application g.

10 - Montrer qu'il existe $(p, q) \in \mathbb{N}^2$ tel que $p \neq q$ et $g^p(V) \cap g^q(V) \neq \varnothing$.

11 - En déduire qu'il existe un élément $\mathbf{x} \in V$ qui « retourne » dans l'ouvert V, c'est-à-dire un élément $\mathbf{x} \in V$ pour lequel il existe $n \in \mathbb{N}^*$ avec $g^n(\mathbf{x}) \in V$. Ce résultat est connu sous le nom de théorème de retour de Poincaré.

12 - Notons C le cercle du plan de centre $(0, 0)$ et de rayon 1 et g une rotation d'angle $\alpha \in \mathbb{R}$. Montrer qu'à partir de tout point $\mathbf{x} \in C$, en appliquant un nombre fini de fois la rotation g, on peut se rapprocher autant que l'on souhaite du point initial $\mathbf{x}$.

13 - Selon le principe de moindre action de Hamilton[84], le mouvement d'une particule située au temps $t \in \mathbb{R}$ à la position $p(t) \in \mathbb{R}^3$ et animée d'une vitesse $q(t) \in \mathbb{R}^3$ est régi par les équations suivantes[85], appelées équations de Hamilton :

$$p'(t) = -\nabla_q H(p(t), q(t)) \quad \text{et} \quad q'(t) = \nabla_p H(p(t), q(t)),$$

où H est une application[86] de $\mathbb{R}^6$ dans $\mathbb{R}$ de classe C^2. Montrer le résultat suivant : si l'on ouvre la cloison séparant une chambre remplie de gaz (composé d'un nombre fini de particules) et une chambre sous vide (voir la figure 11) alors au bout d'un moment toutes les molécules de gaz repasseront en même temps dans la première chambre.

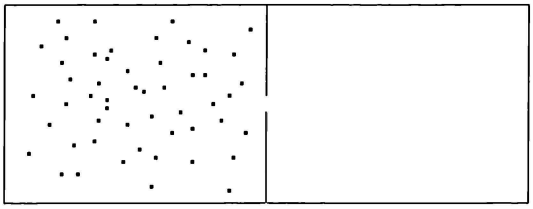


Fig. 11 Position des particules à l'instant initial. Il existe un temps (sans doute très grand) pour lequel toutes les particules retournerons dans la chambre de gauche.

---

[84] HAMILTON, William Rowan (1805, Dublin - 1865, Dublin). Mathématicien, physicien et astronome irlandais.

[85] Dans ces équations, si $p_x(t)$, $p_y(t)$ et $p_z(t)$ désignent les trois coordonnées de la position $p(t)$ alors $p'(t)$ est le vecteur $(p'_x(t), p'_y(t), p'_z(t))$ et la dérivée $\nabla_p H$ correspond au vecteur $(\frac{\partial H}{\partial p_x}, \frac{\partial H}{\partial p_y}, \frac{\partial H}{\partial p_z})$. De même, si $q_x(t)$, $q_y(t)$ et $q_z(t)$ désignent les trois coordonnées de la vitesse $q(t)$ alors $q'(t)$ est le vecteur $(q'_x(t), q'_y(t), q'_z(t))$ et la dérivée $\nabla_q H$ correspond au vecteur $(\frac{\partial H}{\partial q_x}, \frac{\partial H}{\partial q_y}, \frac{\partial H}{\partial q_z})$.

[86] Physiquement, l'application H est déterminée à partir des énergies cinétiques et potentielles du système considéré.

13.6 Solutions des exercices

Solution de l'exercice 13.1

1 - On paramètre le carré $CASI$ en utilisant un arc continu et de classe $\mathcal{C}^1$ par morceaux :

$$\varphi : t \in [0,4] \longmapsto \begin{cases} (1-2t,1) & \text{si } t \in [0,1], \quad \text{(segment } [CA]) \\ (-1,3-2t) & \text{si } t \in [1,2], \quad \text{(segment } [AS]) \\ (-5+2t,-1) & \text{si } t \in [2,3], \quad \text{(segment } [SI]) \\ (1,-7+2t) & \text{si } t \in [3,4]. \quad \text{(segment } [IC]) \end{cases}$$

En utilisant la relation de Chasles, on en déduit que l'intégrale curviligne de f le long du carré $CASI$ est la somme des quatre intégrales curvilignes de f le long des quatre côtés du carré. Le long du côté $[CA]$, on a $\varphi(t) = (1-2t, 1)$ donc $\varphi'(t) = (-2, 0)$ et $f(\varphi(t)) = (2-2t, 1-2t)$ de sorte que l'intégrale curviligne de f le long du segment $[CA]$ vaut

$$\int_{[CA]} f \cdot \, \mathrm{d}x = \int_0^1 f(\varphi(t)) \cdot \varphi'(t) \, \mathrm{d}t = \int_0^1 4t - 4 \, \mathrm{d}t = -2.$$

En effectuant un calcul similaire sur chacun des quatre côtés du carré, on obtient

$$\int_{CASI} f \cdot \mathrm{d}x = \underbrace{\int_{[CA]} f \cdot \mathrm{d}x}_{=-2} + \underbrace{\int_{[AS]} f \cdot \mathrm{d}x}_{=0} + \underbrace{\int_{[SI]} f \cdot \mathrm{d}x}_{=-2} + \underbrace{\int_{[IC]} f \cdot \mathrm{d}x}_{=0} = -4$$

2 - Un paramétrage du cercle Γ de centre 0 et de rayon $R \in \mathbb{R}_+^*$ est $\varphi : t \in [0, 2\pi] \mapsto (R\cos(t), R\sin(t))$. On vérifie que ce paramétrage respecte l'orientation indiquée. Le vecteur tangent à Γ en $\varphi(t)$ est donné par $\varphi'(t) = (-R\sin(t), R\cos(t))$. On en déduit que l'intégrale curviligne de f le long du cercle Γ vaut :

$$\int_\Gamma f \cdot \mathrm{d}\mathbf{x} = \int_0^{2\pi} R\big(\sin^2(t) + \cos^2(t)\big) \, \mathrm{d}t = 2\pi R.$$

Solution de l'exercice 13.2

Un paramétrage de la parabole est $\varphi : t \in [0,a] \mapsto (t, t^2/2)$. Pour $t \in [0,a]$ on a $\|\varphi'(t)\| = \sqrt{1+t^2}$ et la longueur de la parabole est

$$\mathrm{Long}(a) = \int_0^a \sqrt{1+t^2} \, \mathrm{d}t.$$

Pour obtenir la valeur de cette intégrale, on peut effectuer le changement de variable $t = \mathrm{sh}(u)$. On a

$$\begin{aligned} \int_0^a \sqrt{1+t^2} \, \mathrm{d}t &= \int_{\mathrm{sh}(0)}^{\mathrm{sh}(\mathrm{argsh}(a))} \sqrt{1+t^2} \, \mathrm{d}t = \int_0^{\mathrm{argsh}(a)} \sqrt{1+\mathrm{sh}(u)^2}\,\mathrm{ch}(u) \, \mathrm{d}u \\ &= \int_0^{\mathrm{argsh}(a)} \mathrm{ch}(u)^2 \, \mathrm{d}u. \end{aligned}$$

On « linéarise » $\operatorname{ch}(u)^2$ en écrivant que $\operatorname{ch}(u)^2 = \frac{1}{2}(\operatorname{ch}(2u)+1)$ et on obtient

$$\begin{aligned}\int_0^a \sqrt{1+t^2}\,\mathrm{d}t &= \tfrac{1}{2}\int_0^{\operatorname{argsh}(a)} \operatorname{ch}(2u)\,\mathrm{d}u + \tfrac{1}{2}\operatorname{argsh}(a)\\ &= \tfrac{1}{4}\Big[\operatorname{sh}(2u)\Big]_0^{\operatorname{argsh}(a)} + \tfrac{1}{2}\operatorname{argsh}(a) = \tfrac{1}{4}\operatorname{sh}(2\operatorname{argsh}(a)) + \tfrac{1}{2}\operatorname{argsh}(a).\end{aligned}$$

Comme pour tout $u \in \mathbb{R}$ on a $\operatorname{sh}(2u) = 2\operatorname{sh}(u)\operatorname{ch}(u) = 2\operatorname{sh}(u)\sqrt{1+\operatorname{sh}(u)^2}$, on en déduit que

$$\operatorname{Long}(a) = \tfrac{1}{2}\big(a\sqrt{1+a^2} + \operatorname{argsh}(a)\big).$$

Une autre manière de procéder pour calculer $\int_0^a \sqrt{1+t^2}\,\mathrm{d}t$ consiste à intégrer par parties.

Solution de l'exercice 13.3

Un paramétrage du fil dans le plan est $\varphi : t \in [0,L] \mapsto (t,0) \in \mathbb{R}^2$. Introduisons les notations $\rho : (x,y) \in \mathbb{R}^2 \mapsto 1+x^2$ et $f_x : (x,y) \in \mathbb{R}^2 \mapsto x\rho(x,y)$, $f_y : (x,y) \in \mathbb{R}^2 \mapsto y\rho(x,y)$. Par définition, la masse du fil vaut

$$\mathrm{M} = \int_C \rho\,\mathrm{d}\tau = \int_0^L (1+t^2)\,\mathrm{d}t = L + \frac{L^3}{3},$$

et les coordonnées de son centre de gravité sont données par

$$\begin{aligned}x_G &= \frac{1}{M}\int_C f_x\,\mathrm{d}\tau = \frac{1}{M}\int_0^L t(1+t^2)\,\mathrm{d}t = \frac{6L+3L^3}{12+4L^2},\\ y_G &= \frac{1}{M}\int_C f_y\,\mathrm{d}\tau = \frac{1}{M}\int_0^L 0\times(1+t^2)\,\mathrm{d}t = 0.\end{aligned}$$

Solution de l'exercice 13.4

1 - La fonction f est bien définie sur l'ouvert U et pour tout $(x_1,x_2,x_3) \in U$ on a $f(x_1,x_2,x_3) > 0$ ce qui permet de définir $h(x_1,x_2,x_3) = \ln(f(x_1,x_2,x_3))$. Comme pour $(x_1,x_2,x_3) \in U$ on a $x_1 > 0$, $x_2 > 0$ et $4x_1 - x_2x_3^2 > 0$, on peut écrire

$$h(x_1,x_2,x_3) = \ln(x_1) + \tfrac{1}{2}\ln(x_2) - \tfrac{1}{2}\ln(4x_1 - x_2x_3^2).$$

2 - L'application h est de classe $\mathcal{C}^\infty$, ses dérivées partielles en $(x_1,x_2,x_3) \in U$ sont données par

$$\begin{aligned}\frac{\partial h}{\partial x_1}(x_1,x_2,x_3) &= \frac{1}{x_1} - \frac{2}{4x_1 - x_2x_3^2},\\ \frac{\partial h}{\partial x_2}(x_1,x_2,x_3) &= \frac{1}{2x_2} + \frac{x_3^2}{2(4x_1 - x_2x_3^2)},\\ \frac{\partial h}{\partial x_3}(x_1,x_2,x_3) &= \frac{x_2x_3}{4x_1 - x_2x_3^2}.\end{aligned}$$

Ainsi, pour $\mathbf{x} = (x_1, x_2, x_3) \in U$ on a

$$\mathrm{d}h_{\mathbf{x}} = \left(\frac{1}{x_1} - \frac{2}{4x_1 - x_2x_3^2}\right)\,\mathrm{d}x_1 + \left(\frac{1}{2x_2} + \frac{x_3^2}{2(4x_1 - x_2x_3^2)}\right)\,\mathrm{d}x_2 + \frac{x_2x_3}{4x_1 - x_2x_3^2}\,\mathrm{d}x_3.$$

3 - D'après les règles de dérivations des fonctions composées, pour $i \in \{1, 2, 3\}$ on a $\dfrac{\partial h}{\partial x_i} = \dfrac{\partial(\ln(f))}{\partial x_i} = \dfrac{1}{f}\dfrac{\partial f}{\partial x_i}$. Par conséquent pour tout $\mathbf{x} \in U$ on a

$$\begin{aligned}\frac{\mathrm{d}f}{f}(\mathbf{x}) &= \frac{1}{f(\mathbf{x})}\mathrm{d}f(\mathbf{x}) = \frac{1}{f(\mathbf{x})}\left(\frac{\partial f}{\partial x_1}(\mathbf{x})\,\mathrm{d}x_1 + \frac{\partial f}{\partial x_2}(\mathbf{x})\,\mathrm{d}x_2 + \frac{\partial f}{\partial x_3}(\mathbf{x})\,\mathrm{d}x_3\right)\\ &= \frac{\partial h}{\partial x_1}(\mathbf{x})\,\mathrm{d}x_1 + \frac{\partial h}{\partial x_2}(\mathbf{x})\,\mathrm{d}x_2 + \frac{\partial h}{\partial x_3}(\mathbf{x})\,\mathrm{d}x_3 = \mathrm{d}h(\mathbf{x}).\end{aligned}$$

Cela implique que $\dfrac{\mathrm{d}f}{f} = \mathrm{d}h$.

Solution de l'exercice 13.5

Soit y une solution de l'équation différentielle $3x^2y + (x^3 - \sin(y))y' = 0$. Introduisons la 1-forme de $\mathbb{R}^2$, $\omega(x, y) = 3x^2y\mathrm{d}x + (x^3 - \sin(y))\mathrm{d}y$. Cette 1-forme est fermée et $\mathbb{R}^2$ est un ouvert étoilé. Par conséquent la 1-forme ω est exacte : il existe une application $f \in \mathcal{C}^1(\mathbb{R}^2, \mathbb{R})$ telle que $\omega = \mathrm{d}f$. Pour tout $(x, y) \in \mathbb{R}^2$ on a

$$\begin{cases}\dfrac{\partial f}{\partial x}(x, y) = 3x^2y,\\ \dfrac{\partial f}{\partial y}(x, y) = x^3 - \sin(y).\end{cases}$$

Pour que f vérifie la première de ces deux relations, il faut et il suffit qu'il existe une application $g \in \mathcal{C}^1(\mathbb{R}, \mathbb{R})$ telle que pour tout $(x, y) \in \mathbb{R}^2$ on ait $f(x, y) = x^3y + g(y)$. En reportant cette relation dans la seconde équation du système, on en déduit que pour tout $y \in \mathbb{R}$ on a $g'(y) = -\sin(y)$, c'est-à-dire $g(y) = \cos(y) + k$ où k est une constante réelle. Les applications f telles que $\mathrm{d}f = \omega$ sont donc les applications $f : (x, y) \in \mathbb{R}^2 \mapsto x^3y + \cos(y) + k$ où k est une constante réelle. Finalement, si y est une solution de l'équation différentielle $3x^2y + (x^3 - \sin(y))y' = 0$ alors on a $\mathrm{d}f = 0$, c'est-à-dire f est constante et donc pour tout $x \in \mathbb{R}$

$$x^3y(x) + \cos(y(x)) = C \quad \text{où } C \text{ est une constante réelle.}$$

On en déduit que pour tout $x \in \mathbb{R}$ on a $C - 1 \leqslant x^3y(x) \leqslant C + 1$, autrement dit pour tout $x \in \mathbb{R}_+^*$ on a

$$\frac{C-1}{x^3} \leqslant y(x) \leqslant \frac{C+1}{x^3}.$$

D'après le théorème d'encadrement, on en déduit que $\lim\limits_{x \to +\infty} y(x) = 0$. De même, pour tout $x \in \mathbb{R}_-^*$ on a

$$\frac{C+1}{x^3} \leqslant y(x) \leqslant \frac{C-1}{x^3},$$

et par conséquent $\lim_{x\to-\infty} y(x) = 0$.

Solution de l'exercice 13.6

Pour $(t,\lambda) \in [0,h] \times [0,1]$, on a

$$\mathbf{N}(t,\lambda) = \big(\sin(t), -\cos(t), \lambda\big)\,.$$

La norme de ce vecteur vaut $\sqrt{1+\lambda^2}$; elle n'est jamais nulle et par conséquent le vecteur $\mathbf{N}(t,\lambda)$ n'est pas nul[87].

Notons I l'intégrale de $\mathbf{N}/\|\mathbf{N}\|$ à travers la surface S. On a

$$\begin{aligned} I &= \int_{[0,h]\times[0,1]} \frac{\mathbf{N}(t,\lambda)}{\|\mathbf{N}(t,\lambda)\|} \cdot \mathbf{N}(t,\lambda)\ \mathrm{d}t\ \mathrm{d}\lambda = \int_{[0,h]\times[0,1]} \|\mathbf{N}(t,\lambda)\|\ \mathrm{d}t\ \mathrm{d}\lambda \\ &= \int_0^1 \int_0^h \sqrt{1+\lambda^2}\ \mathrm{d}t\ \mathrm{d}\lambda = h \int_0^1 \sqrt{1+\lambda^2}\ \mathrm{d}\lambda. \end{aligned}$$

L'intégrale $\int_0^1 \sqrt{1+\lambda^2}\ \mathrm{d}\lambda$ vaut[88] $\frac{1}{2}\big(\sqrt{2} + \operatorname{argsh}(1)\big) \approx 1,148$ de sorte que

$$I = \frac{h}{2}\Big(\sqrt{2} + \operatorname{argsh}(1)\Big) \approx 1,148\ h.$$

Solution de l'exercice 13.7

Le cône considéré est la surface de révolution engendrée par la rotation autour de l'axe $(O;\mathbf{e}_3)$ du segment $[OA]$ où A est le point de $\mathbb{R}^3$ de coordonnées $(h\tan(\alpha), 0, h)$. Le segment $[OA]$ est paramétré par

$$f : r \in [0, h\tan(\alpha)] \longmapsto \left(r, 0, \frac{1}{\tan(\alpha)} r\right).$$

On en déduit qu'un paramétrage du cône est

$$\varphi : (r,\theta) \in [0, h\tan(\alpha)] \times [0, 2\pi] \mapsto (r\cos(\theta), r\sin(\theta), f(r)).$$

Ainsi, une normale au cône est donnée au point $\varphi(r,\theta)$ par

$$\begin{aligned} \mathbf{N}(r,\theta) &= \frac{\partial\varphi}{\partial r}(r,\theta) \wedge \frac{\partial\varphi}{\partial\theta}(r,\theta) \\ &= \big(\cos(\theta), \sin(\theta), f'(r)\big) \wedge \big(-r\sin(\theta), r\cos(\theta), 0\big) \\ &= \big(-f'(r) r\cos(\theta), -f'(r) r\sin(\theta), r\big) \end{aligned}$$

[87] Ceci indique que la surface S est une surface paramétrée régulière.

[88] Le calcul de cette intégrale a été effectué à la p. 949.

et sa norme vaut $\|\mathbf{N}(r,\theta)\| = r\sqrt{1+(f'(r))^2} = r\sqrt{1+\tan^{-2}(\alpha)} = \dfrac{r}{\sin(\alpha)}$.
On en déduit que

$$\begin{aligned}\mathcal{A}(\alpha,h) &= \int_{[0,h\tan(\alpha)]\times[0,2\pi]} \|\mathbf{N}(r,\theta)\|\,\mathrm{d}r\mathrm{d}\theta = \int_0^{2\pi}\int_0^{h\tan(\alpha)} \frac{r}{\sin(\alpha)}\,\mathrm{d}r\mathrm{d}\theta \\ &= \frac{1}{\sin(\alpha)}\left(\int_0^{h\tan(\alpha)} r\,\mathrm{d}r\right)\left(\int_0^{2\pi}\mathrm{d}\theta\right) = \frac{\pi h^2\sin(\alpha)}{\cos^2(\alpha)}.\end{aligned}$$

Solution de l'exercice 13.8

1 - On vérifie facilement que pour tout $(\mathbf{u},\mathbf{v}) \in (\mathbb{R}^3)^2$ on a $f(\mathbf{u},\mathbf{v}) = -f(\mathbf{v},\mathbf{u})$. Cela signifie que si l'application f est bilinéaire, elle est alternée. De plus, cette observation permet, pour montrer que f est bilinéaire, de se contenter de montrer qu'elle est linéaire par rapport à sa première variable. Pour tout $(\mathbf{u},\mathbf{v},\mathbf{w}) \in (\mathbb{R}^3)$ et pour tout $(\lambda,\mu) \in \mathbb{R}^2$, on a

$$\begin{aligned}f(\lambda\mathbf{u}+\mu\mathbf{w},\mathbf{v}) &= \begin{pmatrix}(\lambda u_2+\mu w_2)v_3-(\lambda u_3+\mu w_3)v_2\\(\lambda u_3+\mu w_3)v_1-(\lambda u_1+\mu w_1)v_3\\(\lambda u_1+\mu w_1)v_2-(\lambda u_2+\mu w_2)v_1\end{pmatrix} \\ &= \lambda\begin{pmatrix}u_2v_3-u_3v_2\\u_3v_1-u_1v_3\\u_1v_2-u_2v_1\end{pmatrix} + \mu\begin{pmatrix}w_2v_3-w_3v_2\\w_3v_1-w_1v_3\\w_1v_2-w_2v_1\end{pmatrix} = \lambda f(\mathbf{u},\mathbf{v}) + \mu f(\mathbf{w},\mathbf{v}).\end{aligned}$$

L'application f étant linéaire par rapport à sa première variable, on en déduit des remarques préliminaires que f est une application bilinéaire alternée.

2 - Comme f est une application bilinéaire alternée, pour tous $\mathbf{u},\mathbf{v},\mathbf{w} \in \mathbb{R}^3$, on a

$$\begin{aligned}f(\mathbf{u},\mathbf{v}) &= -f(\mathbf{v},\mathbf{u})\\ f(\mathbf{u},\mathbf{u}) &= 0\\ f(\mathbf{u}+\mathbf{v},\mathbf{w}) &= f(\mathbf{u},\mathbf{w}) + f(\mathbf{v},\mathbf{w})\\ f(\mathbf{w},\mathbf{u}+\mathbf{v}) &= f(\mathbf{w},\mathbf{u}) + f(\mathbf{w},\mathbf{v}).\end{aligned}$$

Les propriétés a), b), c) et d) sont l'écriture à l'aide du symbole $\wedge$ de ces 4 propriétés.

Solution de l'exercice 13.9

1 - Le produit scalaire $(\mathbf{u},\mathbf{v}) \in (\mathbb{R}^3)^2 \mapsto \mathbf{u}\cdot\mathbf{v} \in \mathbb{R}$ est une application bilinéaire, voir la définition 8.15 p. 479. Le produit vectoriel est lui aussi une application bilinéaire, voir l'exercice 13.8. On en déduit que pour tout $(\mathbf{u},\mathbf{v},\mathbf{w},\mathbf{z}) \in (\mathbb{R}^3)^4$ et pour tout $\lambda \in \mathbb{R}$ on a

$$\begin{aligned}g(\lambda\mathbf{u}+\mathbf{z},\mathbf{v},\mathbf{w}) &= ((\lambda\mathbf{u}+\mathbf{z})\wedge\mathbf{v})\cdot\mathbf{w} = (\lambda\mathbf{u}\wedge\mathbf{v}+\mathbf{z}\wedge\mathbf{v})\cdot\mathbf{w}\\ &= \lambda(\mathbf{u}\wedge\mathbf{v})\cdot\mathbf{w} + (\mathbf{z}\wedge\mathbf{v})\cdot\mathbf{w} = \lambda\, g(\mathbf{u},\mathbf{v},\mathbf{w}) + g(\mathbf{z},\mathbf{v},\mathbf{w}).\end{aligned}$$

De la même manière, on montre que

$$g(\mathbf{u}, \lambda\mathbf{v} + \mathbf{z}, \mathbf{w}) = \lambda\, g(\mathbf{u}, \mathbf{v}, \mathbf{w}) + g(\mathbf{u}, \mathbf{z}, \mathbf{w}),$$
$$g(\mathbf{u}, \mathbf{v}, \lambda\mathbf{w} + \mathbf{z}) = \lambda\, g(\mathbf{u}, \mathbf{v}, \mathbf{w}) + g(\mathbf{u}, \mathbf{v}, \mathbf{z}).$$

On en conclut que g est bien une application trilinéaire. Pour montrer que g est une application alternée, nous devons montrer que si deux des trois vecteurs $\mathbf{u}$, $\mathbf{v}$ et $\mathbf{w}$ sont égaux alors $g(\mathbf{u}, \mathbf{v}, \mathbf{w}) = 0$. En utilisant la propriété b) de l'exercice 13.8, on obtient pour tout $(\mathbf{u}, \mathbf{w}) \in (\mathbb{R}^3)^2$

$$g(\mathbf{u}, \mathbf{u}, \mathbf{w}) = 0. \tag{28}$$

Soit $(\mathbf{u}, \mathbf{v}) \in (\mathbb{R}^3)^2$. On a

$$g(\mathbf{u}, \mathbf{v}, \mathbf{u}) = (\mathbf{u} \wedge \mathbf{v}) \cdot \mathbf{u} = \begin{pmatrix} u_2v_3 - u_3v_2 \\ u_3v_1 - u_1v_3 \\ u_1v_2 - u_2v_1 \end{pmatrix} \cdot \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = 0. \tag{29}$$

D'après la propriété a) de l'exercice 13.8 et la linéarité de l'application g par rapport à sa première variable, on a aussi $g(\mathbf{u}, \mathbf{v}, \mathbf{v}) = g(-\mathbf{v}, \mathbf{u}, \mathbf{v}) = -g(\mathbf{v}, \mathbf{u}, \mathbf{v})$. Par conséquent, en utilisant la relation (29) on en déduit

$$g(\mathbf{u}, \mathbf{v}, \mathbf{v}) = 0. \tag{30}$$

Les trois résultats (28), (29) et (30) impliquent que l'application g est alternée.

2 - Les relations (29) et (30) impliquent que pour tout $(\mathbf{u}, \mathbf{v}) \in \mathbb{R}^3 \times \mathbb{R}^3$ le produit vectoriel $\mathbf{u} \wedge \mathbf{v}$ est un vecteur de $\mathbb{R}^3$ orthogonal aux vecteurs $\mathbf{u}$ et $\mathbf{v}$.

3 - En calculant explicitement la valeur de $\langle\!\langle \mathbf{u}, \mathbf{v}, \mathbf{w} \rangle\!\rangle$ en fonction des coordonnées des vecteurs $\mathbf{u}$, $\mathbf{v}$ et $\mathbf{w}$, on remarque que

$$\langle\!\langle \mathbf{u}, \mathbf{v}, \mathbf{w} \rangle\!\rangle = \det(\mathbf{u}, \mathbf{v}, \mathbf{w}),$$

autrement dit que $\langle\!\langle \mathbf{u}, \mathbf{v}, \mathbf{w} \rangle\!\rangle$ correspond au volume algébrique du parallélépipède construit à partir des trois vecteurs $\mathbf{u}$, $\mathbf{v}$ et $\mathbf{w}$.

4 - Soit $(\mathbf{u}, \mathbf{v}, \mathbf{w}) \in \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3$ de coordonnées respectives (u_1, u_2, u_3), (v_1, v_2, v_3) et (w_1, w_2, w_3). D'une part on a

$$\mathbf{u}\wedge(\mathbf{v}\wedge\mathbf{w}) = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} \wedge \begin{pmatrix} v_2w_3 - v_3w_2 \\ v_3w_1 - v_1w_3 \\ v_1w_2 - v_2w_1 \end{pmatrix} = \begin{pmatrix} u_2(v_1w_2 - v_2w_1) - u_3(v_3w_1 - v_1w_3) \\ u_3(v_2w_3 - v_3w_2) - u_1(v_1w_2 - v_2w_1) \\ u_1(v_3w_1 - v_1w_3) - u_2(v_2w_3 - v_3w_2) \end{pmatrix}$$

et d'autre part

$$(\mathbf{w}\cdot\mathbf{u})\mathbf{v} - (\mathbf{u}\cdot\mathbf{v})\mathbf{w} = (w_1u_1 + w_2u_2 + w_3u_3) \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} - (u_1v_1 + u_2v_2 + u_3v_3) \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix}$$

En développant ces deux quantités, on montre qu'elles sont égales, ce qui conclut l'exercice.

Solution de l'exercice 13.10

Par définition du produit extérieur, on a

$$\begin{aligned}\omega\wedge\omega &= (\mathrm{d}x_1\wedge\mathrm{d}x_2+\mathrm{d}x_3\wedge\mathrm{d}x_4)\wedge(\mathrm{d}x_1\wedge\mathrm{d}x_2+\mathrm{d}x_3\wedge\mathrm{d}x_4)\\ &= \mathrm{d}x_1\wedge\mathrm{d}x_2\wedge\mathrm{d}x_1\wedge\mathrm{d}x_2 \;+\; \mathrm{d}x_1\wedge\mathrm{d}x_2\wedge\mathrm{d}x_3\wedge\mathrm{d}x_4\\ &\quad+\; \mathrm{d}x_3\wedge\mathrm{d}x_4\wedge\mathrm{d}x_1\wedge\mathrm{d}x_2 \;+\; \mathrm{d}x_3\wedge\mathrm{d}x_4\wedge\mathrm{d}x_3\wedge\mathrm{d}x_4.\end{aligned}$$

Or, si deux 1-formes parmi les 1-formes a, b, c et d sont égales dans le produit extérieur $a\wedge b\wedge c\wedge d$ alors celui-ci est nul. On sait aussi que l'échange de deux 1-formes change le signe du résultat, par exemple $a\wedge b\wedge c\wedge d=-a\wedge b\wedge d\wedge c$. Ainsi, on a

$$\begin{aligned}&\mathrm{d}x_1\wedge\mathrm{d}x_2\wedge\mathrm{d}x_1\wedge\mathrm{d}x_2=0,\\ &\mathrm{d}x_3\wedge\mathrm{d}x_4\wedge\mathrm{d}x_1\wedge\mathrm{d}x_2=\mathrm{d}x_1\wedge\mathrm{d}x_2\wedge\mathrm{d}x_3\wedge\mathrm{d}x_4,\\ &\mathrm{d}x_3\wedge\mathrm{d}x_4\wedge\mathrm{d}x_3\wedge\mathrm{d}x_4=0.\end{aligned}$$

Par conséquent, on obtient $\omega\wedge\omega=2\mathrm{d}x_1\wedge\mathrm{d}x_2\wedge\mathrm{d}x_3\wedge\mathrm{d}x_4\neq 0$.

Solution de l'exercice 13.11

Un paramétrage de la surface S est $\varphi:(u,v)\in K\mapsto(u,v,1)$. L'intégrale de la 2-forme $x_1\mathrm{d}x_2\wedge\mathrm{d}x_3$ est égale à l'intégrale sur la surface S de l'application $\alpha:(x_1,x_2,x_3)\in\mathbb{R}^3\mapsto(x_1,0,0)\in\mathbb{R}^3$, c'est-à-dire

$$\int_S x_1\;\mathrm{d}x_2\wedge\mathrm{d}x_3=\int_K\alpha(\varphi(u,v))\cdot\mathbf{N}(u,v)\;\mathrm{d}u\mathrm{d}v,$$

où $\mathbf{N}(u,v)$ qui est la normale à la surface au point paramétré par $(u,v)\in K$ vaut $\mathbf{N}(u,v)=\frac{\partial\varphi}{\partial u}(u,v)\wedge\frac{\partial\varphi}{\partial v}(u,v)=(0,0,1)$. Puisque $\alpha(\varphi(u,v))=(u,0,0)$, on en déduit que

$$\int_S x_1\mathrm{d}x_2\wedge\mathrm{d}x_3=\int_K 0\;\mathrm{d}u\mathrm{d}v=0.$$

Solution de l'exercice 13.12

1 - En utilisant la différentielle d'une 1-forme donnée par la relation (21), p. 997, on a immédiatement dans le cas où $n=3$ la relation

$$\begin{aligned}\mathrm{d}\omega=\left(\frac{\partial\alpha_3}{\partial x_2}-\frac{\partial\alpha_2}{\partial x_3}\right)\;\mathrm{d}x_2\wedge\mathrm{d}x_3+\left(\frac{\partial\alpha_1}{\partial x_3}-\frac{\partial\alpha_3}{\partial x_1}\right)\;\mathrm{d}x_3\wedge\mathrm{d}x_1\\ +\left(\frac{\partial\alpha_2}{\partial x_1}-\frac{\partial\alpha_1}{\partial x_2}\right)\;\mathrm{d}x_1\wedge\mathrm{d}x_2.\end{aligned}$$

2 - La 1-forme $\alpha_1\mathrm{d}x_1+\alpha_2\mathrm{d}x_2+\alpha_3\mathrm{d}x_3$ peut s'obtenir en effectuant « formellement » le produit de $\alpha=(\alpha_1,\alpha_2,\alpha_3)$ par $\mathbf{dx}$ de la même manière qu'est calculé un produit scalaire. De même, on reconnaît dans l'expression de $\mathrm{d}\omega$ qu'il peut s'obtenir « formellement » comme le produit scalaire de $\mathrm{rot}(\alpha)$ par $\mathbf{ds}$.

Solution de l'exercice 13.13

1 - Un paramétrage de l'ellipse est[89] $\varphi : t \in [0, 2\pi] \mapsto (a\cos(t), b\sin(t))$.

2 - Selon le corollaire 13.3, en notant C l'ellipse et A le domaine du plan qu'elle délimite, on a

$$\begin{aligned}\text{Aire}(A) &= \frac{1}{2}\int_C (x_1 \mathrm{d}x_2 - x_2 \mathrm{d}x_1) = \frac{1}{2}\int_0^{2\pi} (\varphi_1(t)\varphi_2'(t) - \varphi_2(t)\varphi_1'(t))\ \mathrm{d}t \\ &= \frac{1}{2}\int_0^{2\pi} \left(ab\cos(t)^2 + ab\sin(t)^2\right)\ \mathrm{d}t = \frac{ab}{2}\int_0^{2\pi} \mathrm{d}t = ab\pi.\end{aligned}$$

Solution de l'exercice 13.14

1 - Puisque le réel α est non nul, la relation (25) permet d'établir que pour tout $\varphi \in]-\pi/2, \pi/2[$

$$\theta(\varphi) = \theta_1 + \frac{1}{\tan(\alpha)}\left(\ln(\tan(\frac{\varphi}{2} + \frac{\pi}{4})) - \ln(\tan(\frac{\varphi_1}{2} + \frac{\pi}{4}))\right).$$

En utilisant la représentation paramétrique d'une sphère[90] à l'aide des coordonnées sphériques, on en déduit que l'application F paramétrant la trajectoire de l'avion est

$$F : \varphi \in]-\pi/2, \pi/2[\longmapsto \begin{pmatrix} R\cos(\theta(\varphi))\cos(\varphi) \\ R\sin(\theta(\varphi))\cos(\varphi) \\ R\sin(\varphi) \end{pmatrix}.$$

2 - La longueur L de la courbe paramétrée par F est[91] égale à $\left|\int_{\varphi_1}^{\varphi_2} \|F'(\varphi)\|\ \mathrm{d}\varphi\right|$. Pour tout $\varphi \in]-\pi/2, \pi/2[$, la norme du vecteur $F'(\varphi)$ vaut $\frac{R}{|\sin(\alpha)|}$. Par conséquent

$$L = \left|\int_{\varphi_1}^{\varphi_2} \frac{R}{|\sin(\alpha)|}\mathrm{d}\varphi\right| = R\left|\frac{\varphi_2 - \varphi_1}{\sin(\alpha)}\right|.$$

3 - On déduit de la relation (25) que

$$\tan(\alpha) = \frac{1}{\theta_2 - \theta_1}\left(\ln(\tan(\frac{\varphi_2}{2} + \frac{\pi}{4})) - \ln(\tan(\frac{\varphi_1}{2} + \frac{\pi}{4}))\right).$$

On obtient[92] pour valeur de l'angle à suivre sur la boussole $-100,9^o$.

---

[89] Voir p. 724.

[90] Voir p. 726.

[91] La longueur d'une courbe paramétrée par une application $f : t \in I \mapsto f(t) \in \mathbb{R}^n$ vaut $\int_a^b \|f'(t)\|\ \mathrm{d}t$ lorsque $I = [a, b]$ et égale à $-\int_a^b \|f'(t)\|\ \mathrm{d}t$ lorsque $I = [b, a]$.

[92] Attention à ne pas confondre les mesures en radians et celles en degrés. On obtient d'abord $\alpha = -0,1906\ rad$, c'est-à-dire $\alpha = -10,9^o$ et donc sur la boussole un angle à suivre de $-90 - 10,9 = -100,9^o$.

4 - En utilisant la question 2, on montre que la distance que parcourt l'avion entre Lyon et Wuhan est d'environ 9389 km.

Solution de l'exercice 13.15

Partie 1 - Étude d'un exemple

1 - L'équation différentielle $y''(t) + y(t) = 0$ est une équation différentielle du second ordre linéaire et à coefficients constants. Son équation caractéristique qui est $X^2 + 1 = 0$ admet deux solutions imaginaires i et $-$i. L'ensemble des solutions est par conséquent[93] l'espace vectoriel de dimension 2 engendré par les deux applications $t \mapsto \cos(t)$ et $t \mapsto \sin(t)$. L'unique solution telle que $y(0) = x_{01}$ et $y'(0) = x_{02}$ est la solution

$$y(t) = x_{01}\cos(t) + x_{02}\sin(t).$$

2 - Pour une application y de $\mathbb{R}$ dans $\mathbb{R}$ de classe $\mathcal{C}^2$, considérons l'application $\mathbf{x}$ de $\mathbb{R}$ dans $\mathbb{R}^2$ par $\mathbf{x} : t \in \mathbb{R} \mapsto (y(t), y'(t))$. Cette application est différentiable sur $\mathbb{R}$ et pour tout $t \in \mathbb{R}$ on a $\mathbf{x}'(t) = (y'(t), y''(t))$. Si y est une solution de l'équation $y''(t)+y(t) = 0$ alors $\mathbf{x}$ satisfait $\mathbf{x}'(t) = f(\mathbf{x}(t))$ où f est l'application de $\mathbb{R}^2$ dans $\mathbb{R}^2$ définie par

$$f : (x_1, x_2) \in \mathbb{R}^2 \longmapsto (x_2, -x_1). \tag{31}$$

Réciproquement, si $\mathbf{x}$ est solution de l'équation $\mathbf{x}'(t) = f(\mathbf{x}(t))$ où f est donné par la relation (31) alors la première composante x_1 de $\mathbf{x}$ est solution de l'équation différentielle $x_1''(t) + x_1(t) = 0$.

3 - Le flot associé à cette application f est défini comme l'ensemble de applications g_t, $t \in \mathbb{R}$ qui à des conditions initiales $\mathbf{x}_0 = (x_{01}, x_{02}) \in \mathbb{R}^2$ associent $\mathbf{x}(t)$ où $\mathbf{x}$ est la solution du système différentiel $\mathbf{x}' = f(\mathbf{x})$ sous la condition initiale $\mathbf{x}(0) = \mathbf{x}_0$. D'après la question précédente, l'application g_t est l'application qui à $(x_{01}, x_{02}) \in \mathbb{R}^2$ associe le couple $(y(t), y'(t))$ où y est la solution de l'équation (27) vérifiant $y(0) = x_{01}$ et $y'(0) = x_{02}$. D'après la première question, on en conclut que pour tout $t \in \mathbb{R}$ on a

$$g_t : (x_{01}, x_{02}) \in \mathbb{R}^2 \longmapsto \big(x_{01}\cos(t) + x_{02}\sin(t), -x_{01}\sin(t) + x_{02}\cos(t)\big) \in \mathbb{R}^2.$$

4 - Pour $t \in \mathbb{R}$, on remarque que l'application g_t est une application linéaire de $\mathbb{R}^2$ dans $\mathbb{R}^2$. Sa matrice dans la base canonique de $\mathbb{R}^2$ s'écrit

$$\mathrm{Mat}(g_t) = \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix}.$$

On reconnaît[94] la matrice d'une rotation dans le plan d'angle $-t$. Par conséquent, l'application g_t correspond à la rotation d'angle $-t$ et l'ensemble $\{g_t \ ; \ t \in \mathbb{R}\}$ est le groupe $\mathrm{SO}_2(\mathbb{R})$ des rotations du plan.

---

[93] Voir le *Cours de première année* p. 993.

[94] Voir le théorème 8.7 p. 499.

Partie 2 - Étude du flot

5 - La solution $\mathbf{x}$ du système (26) étant de classe $\mathcal{C}^2$, on peut considérer son développement de Taylor à l'ordre 1 qui s'écrit :

$$\mathbf{x}(t) = \mathbf{x}(0) + t\mathbf{x}'(0) + \mathcal{O}_0(t^2).$$

Puisque $\mathbf{x}(0) = \mathbf{x}_0$, $\mathbf{x}'(0) = f(\mathbf{x}(0)) = f(\mathbf{x}_0)$ et $g_t(\mathbf{x}_0) = \mathbf{x}(t)$, on en déduit que :

$$g_t(\mathbf{x}_0) = \mathbf{x}_0 + tf(\mathbf{x}_0) + \mathcal{O}_0(t^2).$$

6 - Le volume $v(t)$ du sous-ensemble $D(t)$ vaut

$$v(t) = \int_{D(t)} 1 \, \mathrm{d}\mathbf{x} = \int_{g_t(D(0))} 1 \, \mathrm{d}\mathbf{x}.$$

Puisque g_t est une application de classe $\mathcal{C}^1$, bijective et que sa réciproque g_{-t} est aussi de classe $\mathcal{C}^1$, l'application g_t est un $\mathcal{C}^1$-difféomorphisme et la formule du changement de variables[(95)] permet d'écrire

$$v(t) = \int_{D(0)} |\mathcal{J}_{g_t}(\mathbf{x})| \, \mathrm{d}\mathbf{x}.$$

7 - Notons $\mathbf{c}_1, \ldots, \mathbf{c}_n$ les colonnes de la matrice A et $\mathbf{e}_1, \ldots, \mathbf{e}_n$ les vecteurs de la base canonique de $\mathbb{R}^n$. Le déterminant de la matrice $\mathrm{Id} + t\mathrm{A}$ vaut

$$\det(\mathrm{Id} + t\mathrm{A}) = \det(\mathbf{e}_1 + t\mathbf{c}_1, \ldots, \mathbf{e}_n + t\mathbf{c}_n).$$

Puisque l'application det est linéaire par rapport à chacune de ses colonnes, en particulier par rapport à sa première colonne, on a

$$\begin{aligned} \det(\mathrm{Id} + t\mathrm{A}) = {} & \det(\mathbf{e}_1, \mathbf{e}_2 + t\mathbf{c}_2, \ldots, \mathbf{e}_n + t\mathbf{c}_n) \\ & + t\det(\mathbf{c}_1, \mathbf{e}_2 + t\mathbf{c}_2, \ldots, \mathbf{e}_n + t\mathbf{c}_n). \end{aligned} \tag{32}$$

De même, on peut utiliser la linéarité du déterminant par rapport à la deuxième colonne dans chacun des deux déterminants du membre de droite de l'équation (32) ; on obtient :

$$\begin{aligned} \det(\mathrm{Id} + t\mathrm{A}) = {} & \det(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 + t\mathbf{c}_3, \ldots, \mathbf{e}_n + t\mathbf{c}_n) \\ & + t\det(\mathbf{e}_1, \mathbf{c}_2, \mathbf{e}_3 + t\mathbf{c}_3, \ldots, \mathbf{e}_n + t\mathbf{c}_n) \\ & + t\det(\mathbf{c}_1, \mathbf{e}_2, \mathbf{e}_3 + t\mathbf{c}_3, \ldots, \mathbf{e}_n + t\mathbf{c}_n) + t^2\det(\mathbf{c}_1, \mathbf{c}_2, \mathbf{e}_3 + t\mathbf{c}_3, \ldots, \mathbf{e}_n + t\mathbf{c}_n). \end{aligned}$$

Par récurrence, on obtient la relation suivante :

$$\begin{aligned} \det(\mathrm{Id} + t\mathrm{A}) = \det(\mathbf{e}_1, \ldots, \mathbf{e}_n) + t\big(\det(\mathbf{c}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n) + \det(\mathbf{e}_1, \mathbf{c}_2, \mathbf{e}_3, \ldots, \mathbf{e}_n) \\ + \cdots + \det(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_{n-1}, \mathbf{c}_n)\big) + \mathcal{O}_0(t^2). \end{aligned}$$

(95) Voir le théorème 12.2 p. 894.

La matrice dont les colonnes sont les vecteurs de la base canonique $\mathbf{e}_1, \ldots, \mathbf{e}_n$ est la matrice Id et a pour déterminant 1. En développant par rapport à la k–ième colonne le déterminant $\det(\mathbf{e}_1, \ldots, \mathbf{e}_{k-1}, \mathbf{c}_k, \mathbf{e}_{k+1}, \ldots, \mathbf{e}_n)$, on obtient la valeur a_{kk} correspondant au coefficient en position (k,k) de la matrice A. Finalement, on obtient la relation

$$\det(\mathrm{Id} + t\mathrm{A}) = 1 + t(a_{11} + \cdots + a_{nn}) + \mathcal{O}_0(t^2).$$

On en déduit la formule demandée puisque la trace de la matrice A correspond à la somme des éléments diagonaux de A : $\mathrm{Tr}(\mathrm{A}) = a_{11} + \cdots + a_{nn}$.

8 - Soit $\mathbf{x}_0 \in \mathbb{R}^n$. Le jacobien de l'application g_t en $\mathbf{x}_0$ est le déterminant de la matrice jacobienne de g_t en $\mathbf{x}_0$, c'est-à-dire $\mathcal{J}_{g_t}(\mathbf{x}_0) = \det(\mathrm{J}_{g_t}(\mathbf{x}_0))$. D'après la question 6, on a

$$\mathrm{J}_{g_t}(\mathbf{x}_0) = \mathrm{Id} + t\mathrm{J}_f(\mathbf{x}_0) + \mathcal{O}_0(t^2) = \mathrm{Id} + t\mathrm{A},$$

où $\mathrm{A} = \mathrm{J}_f(\mathbf{x}_0) + \mathcal{O}_0(t)$. En utilisant la question précédente et le fait que la trace est une application linéaire (c'est-à-dire que $\mathrm{Tr}(\mathrm{B} + \mathrm{C}) = \mathrm{Tr}(\mathrm{B}) + \mathrm{Tr}(\mathrm{C})$ pour toutes matrices carrées B et C), on obtient

$$\mathcal{J}_{g_t}(\mathbf{x}_0) = 1 + t\mathrm{Tr}\big(\mathrm{J}_f(\mathbf{x}_0)\big) + \mathcal{O}_0(t^2).$$

On en déduit la relation demandée puisque la trace de la matrice jacobienne de f au point $\mathbf{x}_0$ correspond[(96)] à la divergence de f en $\mathbf{x}_0$:

$$\mathcal{J}_{g_t}(\mathbf{x}_0) = 1 + t\ \mathrm{div} f(\mathbf{x}_0) + \mathcal{O}_0(t^2). \tag{33}$$

9 - Supposons que l'application f soit à divergence nulle. La relation (33) ainsi que le résultat de la question 7 montre que le volume de $D(t)$ vaut

$$v(t) = \int_{D(0)} |\mathcal{J}_{g_t}(\mathbf{x})|\ \mathrm{d}\mathbf{x} = \int_{D(0)} 1\ \mathrm{d}\mathbf{x} + \mathcal{O}_0(t^2) = v(0) + \mathcal{O}_0(t^2).$$

On en déduit que l'application v est dérivable en $t = 0$ et que $v'(0) = 0$. En utilisant le fait que le flot $\{g_t\ ;\ t \in \mathbb{R}\}$ est un groupe pour la loi de composition des applications, on va vérifier que ce que l'on vient de prouver pour $t = 0$ est encore vrai pour tout $t_0 \in \mathbb{R}$. En utilisant le changement de variable $g_{t_0 - t}$, on montre comme à la question 7 que

$$v(t) = \int_{D(t_0)} \left|\mathcal{J}_{g_{t-t_0}}(\mathbf{x})\right|\ \mathrm{d}\mathbf{x}.$$

Puisque l'on a, d'après la question 9,

$$\mathcal{J}_{g_{t-t_0}}(\mathbf{x}_0) = 1 + (t - t_0)\ \mathrm{div}(f(\mathbf{x}_0)) + \mathcal{O}_{t_0}((t - t_0)^2),$$

(96) Voir la remarque de la page 656.

on obtient, pour une application f à divergence nulle, la relation

$$v(t) = \int_{D(t_0)} 1 \, \mathrm{d}\mathbf{x} + \mathcal{O}_{t_0}((t-t_0)^2) = v(t_0) + \mathcal{O}_{t_0}((t-t_0)^2),$$

et par conséquent $v'(t_0) = 0$. Ainsi, si l'application f est à divergence nulle alors $v(t) = v(0)$ pour tout $t \in \mathbb{R}$.

Partie 3 - Théorème du retour de Poincaré

10 - Puisque g est une application de U dans U, tous les ensembles $g^n(V)$, $n \in \mathbb{N}$, sont inclus dans U. De plus, puisque l'application g conserve les volumes, tous les ensembles $g^n(V)$ pour $n \in \mathbb{N}$, possèdent le même volume et celui-ci est strictement[(97)] positif. Si ces ensembles étaient d'intersection vide, le volume de U serait infini ce qui est absurde puisque l'ensemble U est borné[(98)]. Donc, pour certains entiers naturels p et q avec $p > q$ on a $g^p(V) \cap g^q(V) \neq \varnothing$.

11 - Soit $\mathbf{z} \in g^p(V) \cap g^q(V)$; il existe $(\mathbf{x}, \mathbf{y}) \in V^2$ tel que $\mathbf{z} = g^p(\mathbf{x})$ et $\mathbf{z} = g^q(\mathbf{y})$. On a donc $\mathbf{y} = g^{p-q}(\mathbf{x})$, ce qui signifie qu'il existe $n \in \mathbb{N}^*$ $(n = p - q)$ tel que $g^n(\mathbf{x}) \in V$.

12 - D'après la première partie de cet exercice, l'ensemble des rotations du plan correspond au flot de l'application $f : (x_1, x_2) \in \mathbb{R}^2 \mapsto (x_2, -x_1)$. Puisque l'application f est à divergence nulle, on déduit de la seconde partie de l'exercice que les rotations conservent les volumes.
Pour $\mathbf{x} \in C$, on obtient en appliquant le théorème de retour de Poincaré (voir la question précédente) que pour tout voisinage V de $\mathbf{x}$, il existe $n \in \mathbb{N}^*$ tel que $g^n(\mathbf{x}) \in V$. Cela montre que l'on peut se rapprocher autant que l'on souhaite[(99)] du point initial $\mathbf{x}$.

13 - Notons $p_1(t), p_2(t), \ldots, p_N(t)$ les positions des particules au temps $t \in \mathbb{R}$ et $q_1(t), q_2(t), \ldots, q_N(t)$ leurs vitesses respectives. D'après le principe de moindre action de Hamilton, on a

$$\begin{cases} p_1'(t) = -\nabla_q H(p_1(t), q_1(t)) \\ \vdots \\ p_N'(t) = -\nabla_q H(p_N(t), q_N(t)) \end{cases} \quad \text{et} \quad \begin{cases} q_1'(t) = -\nabla_p H(p_1(t), q_1(t)) \\ \vdots \\ q_N'(t) = -\nabla_p H(p_N(t), q_N(t)) \end{cases}$$

Ces systèmes peuvent aussi s'écrire sous la forme d'un système différentiel de type (26), c'est-à-dire sous la forme $\mathbf{x}'(t) = f(\mathbf{x}(t))$ où

$$\mathbf{x} : t \in \mathbb{R} \longmapsto (p_1(t), q_1(t), \ldots, p_N(t), q_N(t)) \in \mathbb{R}^{6N}$$

---

(97) Le volume d'un ouvert est strictement positif puisque tout ouvert contient une boule de rayon strictement positif et que le volume d'une telle boule est strictement positif.

(98) L'ensemble U étant borné, il est contenu dans une boule et son volume est inférieur au volume de cette boule.

(99) On remarquera que lorsque l'angle α de la rotation est un nombre du type $\alpha = 2\pi\frac{p}{q}$, $(p, q) \in \mathbb{Z} \times \mathbb{N}^*$ alors on peut revenir exactement au point initial en effectuant par exemple q fois la rotation d'angle α.

où l'application f est définie par

$$f : (p_1, q_1, \ldots, p_N, q_N) \in \mathbb{R}^{6N} \longmapsto \Big(-\nabla_q H(p_1, q_1), \nabla_p H(p_1, q_1), \ldots \\ \ldots, -\nabla_q H(p_N, q_N), \nabla_p H(p_N, q_N) \Big) \in \mathbb{R}^{6N}.$$

Le flot de ce système est l'ensemble des applications g_t ($t \in \mathbb{R}$) qui, au vecteur $(p_1(0), q_1(0), \ldots, p_N(0), q_N(0))$ (c'est-à-dire aux positions et vitesses initiales) associe la solution $(p_1(t), q_1(t), \ldots, p_N(t), q_N(t))$ du système au temps t (c'est-à-dire les positions et vitesses au temps t).

Remarquons que pour tout $(p_1, q_1, \ldots, p_N, q_N) \in \mathbb{R}^{6N}$ on a

$$\operatorname{div} f(p_1, q_1, \ldots, p_N, q_N) = 0.$$

D'après la seconde partie, toutes les applications g_t du flot conservent les volumes et on peut par conséquent appliquer le théorème de retour de Poincaré à l'application g_1 : pour tout voisinage V de $(p_1(0), q_1(0), \ldots, p_N(0), q_N(0))$, il existe un entier $n \in \mathbb{N}^*$ tel que $g_1^n(p_1(0), q_1(0), \ldots, p_N(0), q_N(0)) \in V$. Puisque $g_1^n(p_1(0), q_1(0), \ldots, p_N(0), q_N(0)) = g_n(p_1(0), q_1(0), \ldots, p_N(0), q_N(0))$, c'est-à-dire $g_1^n(p_1(0), q_1(0), \ldots, p_N(0), q_N(0)) = (p_1(n), q_1(n), \ldots, p_N(n), q_N(n))$, on en déduit que pour tout voisinage V de $(p_1(0), q_1(0), \ldots, p_N(0), q_N(0))$ il existe un entier $n \in \mathbb{N}^*$ tel que $(p_1(n), q_1(n), \ldots, p_N(n), q_N(n)) \in V$.

Notons $C \subset \mathbb{R}^3$ la chambre dans laquelle se trouve toutes les particules initialement. L'ensemble $(C \times \mathbb{R}^3)^N$ est un voisinage de $(p_1(0), q_1(0), \ldots, p_N(0), q_N(0))$ et il existe donc $n \in \mathbb{N}^*$ tel que $(p_1(n), q_1(n), \ldots, p_N(n), q_N(n)) \in (C \times \mathbb{R}^3)^N$ c'est-à-dire tel que $(p_1(n), \ldots, p_N(n)) \in C^N$: toutes les particules seront dans la chambre C au temps n.

Bibliographie

[1] S. Balac et F. Sturm, *Algèbre et analyse, cours de mathématiques de première année*, Presses polytechnique et universitaires romandes, Lausanne, 2003.

[2] J. Borowczyk, J.L. Chabert, M. Guillemot, A. Djebbar, J.C. Martzloff et A. Michel-Pajus, *Histoire d'algorithmes. Du caillou à la puce*, Belin Coll. Pour la Science, Paris, 1995.

[3] J. Cornil et P. Testud, *Maple V. Introduction raisonnée à l'usage de l'étudiant, de l'ingénieur et du chercheur*, Spinger Verlag, Paris, 1997.

[4] A. Doneddu, *Cours de mathématiques (en 8 tomes)*, Vuibert, Paris, 1985.

[5] D. Hauchecorne et B. Suratteau, *Des mathématiciens de A à Z*, Ellipse, Paris, 1996.

[6] S. Lang, *Undergraduate analysis*, Springer Verlag, New York, 1983.

[7] S. Lang, *Undergraduate algebra*, Springer Verlag, New York, 1987.

[8] J. Lelong-Ferrand et J.M. Arnaudiès, *Cours de mathématiques (en 4 tomes)*, Dunod, Paris, 1977.

[9] A. Quarteroni, R. Sacco et F. Saleri, *Méthodes numériques pour le calcul scientifique*, Springer Verlag, Paris, 2000.

Index

Abel
  formule, 281
  lemme, 233
  règle, 218, 302
Abscisse, curviligne, 953
Absolue, convergence, 28, 118
Accroissements finis, 163, 201, 679
Adhérence
  d'un ensemble, 169, 378
  point, 169, 378
  valeur, 394
Adjacentes, suites, 7
Adjoint, endomorphisme, 496
Aire, 874, 979
Airy, équation d', 273
Algèbre, 544
Alternée
  application, 986
  critère spécial, 42, 130
  série, 41
  série harmonique, 4
Application
  alternée, 986
  bijective, 174
  bilinéaire, 438
  conforme, 202
  continue, 189, 546
  contractante, 558
  dérivable, 193
  de classe $\mathcal{C}^q$, 647
  différentiable, 609
  holomorphe, 193
  injective, 174
  lipschitzienne, 558
  partielle, 554
  surjective, 174
  trilinéaire, 985
Arc
  géométrique, 171, 736
  orientable, 740
  paramétré, 170, 714
  simple, 715
Automorphisme, 573

Banach
  espace, 403
  point fixe, 585
Base(s)
  cylindriques, 670
  orthogonale, 461
  orthonormale, 461
  polaire, 668
  sphérique, 674
Bertrand, série, 26
Bessel
  équation, 280
  fonction, 232, 244, 252, 278
Bijection, 174
Bilinéaire
  application, 438
  forme, 438
Borné
  application, 175
  ensemble, 168, 382
  suite, 392
Boule
  fermée, 164, 369
  ouverte, 164, 369
  unité, 369
Bouteille de Klein, 796

Cône, isotrope, 457
Cardioïde, 759
Cauchy
  critère, 9, 124
  produit, 38
  règle, 30, 241
  suite, 398

Cauchy-Schwarz, inégalité de, 481
Cellule, 856
Centre, d'inertie, 875
Cercle de convergence, 237
Changement de variables, 894
Chasles, relation de, 873
Chemin, 170, 385
Circulation
  d'un champ, 939
  d'une 1-forme, 969
Compact
  ensemble, 168
  intervalle, 413
Complet, espace métrique, 401
Conforme, transformation, 202
Congrue, matrice, 444
Conique, 507
Connexe, 384
Continuité, 160, 189, 546
  par morceaux, 321
  prolongement, 189
  uniforme, 556
Contractante, application, 558
Convergence
  absolue, 28, 118
  cercle, 237
  disque, 237
  normale, 125, 247
  rayon, 234
  simple, 2, 71, 114, 216, 217
  uniforme, 73, 120, 216
Coordonnées
  cylindriques, 670
  ligne, 802
  polaires, 668
  sphériques, 674
Cosinus
  complexe, 209
  hyperbolique complexe, 209
Courbe
  birégulière, 745
  cartésienne, 721
  de Lissajous, 715
  de niveau, 728
  fermée, 943
  implicite, 719
  longueur, 946
  paramétrée, 714
  plane, 778
  régulière, 745
  simple, 715
Critère
  de Cauchy, 9
  de Cauchy uniforme, 82, 124
  des séries alternées, 42, 130
Cylindre, 518

D'Alembert, règle de, 34, 242
Définie positive, 477
Dérivée, 193
  directionnelle, 627
  par morceaux, 320
  partielle, 630, 647
Développement, limité, 256
Densité, 380
Diagonalisable
  endomorphisme, 503
  matrice, 503
Difféomorphisme, 170, 663, 664
Différentiable, 609
Différentielle
  d'une application, 611
  forme, 956
Dilatation, 176
Dirichlet, théorème de, 322
Disque
  de convergence, 237
  fermé, 164
  ouvert, 164
Distance, 345
  équivalentes, 352
  d'un point à un ensemble, 386
  discrète, 346
  entre deux ensembles, 388
  euclidienne, 346, 356
  subordonnée à une norme, 361
Divergence, 2, 656
  grossière, 11
  intégrale généralisée, 918
Domaine, 384
Dual, 464

Ellipse, 509, 511, 725
Ellipsoïde, 516

Endomorphisme, 573
adjoint, 496
diagonalisable, 503
Ensemble
convexe, 678
fermé, 165, 371
orthogonal, 455
ouvert, 165, 371
Entière
fonction, 194
série, 231, 260
Equilatère, hyperbole, 725
Equivalence
de distances, 352
de nappes paramétrées, 790
de normes, 360
Equivalentes, matrices, 443
Espace
complet, 401
connexe, 384
de Banach, 403
euclidien, 484
isométriques, 561
métrique, 345
préhilbertien, 484
vectoriel, 358
Euclidien, espace, 484
Euclidienne
distance, 346, 356
norme, 362
Euler
constante, 27
formule, 281, 309
Exacte, forme, 960
Extraite, suite, 7, 395
Extremum, 687

Famille
liée, 459
libre, 459
orthogonale, 458
orthonormale, 458
Fermé
courbe, 943
ensemble, 165, 371
forme, 960
Flot, 1010
Flux, 973
Fonction
de Van der Waerden, 142
de Bessel, 278, 280
de Neumann, 280
en escalier, 857
entière, 194
zeta de Riemann, 53, 129, 140, 329
Forme
bilinéaire, 438
bilinéaire antisymétrique, 446
bilinéaire symétrique, 446
définie, 476
différentielle, 956, 988
exacte, 960
fermée, 960
linéaire, 463
polaire, 449
positive, 476
quadratique, 448
Formule(s)
d'Euler, 309
d'Euler-Abel, 281
d'Ostrogradski, 1006
de Green-Riemann, 1000
de Leibniz, 203
de Moivre, 53, 207
de polarisation, 449
de Stirling, 18, 96
de Stokes, 1005
de Taylor, 683, 684
de Taylor-Lagrange, 206
tigonométriques, 212
Fourier
coefficient, 306
série, 306
Fubini, théorème de, 876

Gauss
loi, 920
méthode de, 467
théorème de, 469
Gibbs, phénomène, 330
Gradient, 655
Green, formule de, 1000
Groupe

des isométries, 491
orthogonal, 493
spécial orthogonal, 497

Hadamard, inégalité d', 551
Hessienne, matrice, 653
Holomorphe, 193
Homéomorphisme, 664
Homographie, 176
Hyperbole, 509, 512, 725
Hyperboloïde, 516
Hyperplan, 466

Inégalité
triangulaire (première), 346
triangulaire (seconde), 351
de Cauchy-Schwarz, 481
de Hadamard, 551
des accroissements finis, 163, 201, 679
Inertie, 875
Injection, 174
Intégrale
convergente, 918
curviligne, 938
d'une 1-forme, 969
d'une 2-forme, 993
d'une fonction en escalier, 858
divergente, 918
généralisée, 918
multiple, 855
sur une surface, 973
Intérieur, 168, 377
Inversion, 176
Isométrie, 490, 560
directe, 498
groupe des, 491
indirecte, 498
Isométriques
espaces, 561
formes, 475
Isomorphisme, 573
Isotrope
cône, 457
vecteur, 457

Jacobienne, matrice, 636
Jordan, théorème, 323

Klein, bouteille de, 796
Kronecker, symbole de, 463

Lacunaire, 232, 243
Laplacien, 659
Lebesgue
lemme, 318
théorème, 91
Lemme
d'Abel, 233
de Lebesgue, 318
Lié, 459
Libre, 459
Ligne
coordonnées, 802
pêche à la, 658
Limite
d'une fonction, 160, 178, 536, 538
d'une suite, 388
Lipschitzienne, application, 558
Lissajous, courbe de, 715
Loi d'inertie de Sylvester, 472
Longueur, d'une courbe, 946
Loupe, règle de la, 15

Möbius, ruban de, 794
Méridien, 801
Méthode
de Gauss, 467
par éclatement, 48
Métrique, 345, 350
Mac-Laurin, série, 255, 265
Matrice(s)
équivalentes, 443
application bilinéaire, 441
application linéaire, 440, 492
congrue, 444
diagonalisable, 503
forme quadratique, 452
hessienne, 653
jacobienne, 636
orthogonale, 493
semblables, 443
Maximum

principe du, 159, 697
local, global, 687
Minimum, local, global, 687
Mixte, produit, 988
Moivre, formule de, 53
Moment, d'inertie, 875

Négligeable, 868
Nappe
géométrique, 789
orientable, 790
paramétrée, 716, 790
Normé
algèbre, 544
espace vectoriel, 358
Normale
convergence, 125, 247
vecteur, 731
Norme, 358
équivalence, 360
de la conv. en moyenne, 368
de la conv. quadratique, 368
de la conv. uniforme, 368
de matrice, 365
de Schur, 365
euclidienne, 362
semi, 358
Noyau, 457, 581

Orientable
arc, 740
nappe, 790
Orthogonal(e), 454
ensemble, 455
groupe, 493
matrice, 493
famille, 458
Orthonormale, famille, 458
Osculateur, plan, 772
Ostrogradski, formule d', 1006
Ouvert
étoilé, 961
ensemble, 165, 371

Périodique, 174, 302
Parabole, 509, 510
Paraboloïde, 517
Parallèle, 801
Paramétrage
d'une courbe, 714
d'une surface, 716
Parseval, égalité, 324
Partielle
application, 554
dérivée, 630
Partition, 384, 873
Pavé, 855
Permutation, 52
Plan
osculateur, 772
tangent, 731, 798, 799
Poincaré
théorème de, 963
théorème du retour, 1011
Point
adhérent, 169, 378
birégulier, 745
col, 693
critique, 689
d'accumulation, 169, 378
d'arrêt, 768
d'inflexion, 762
de rebroussement, 762
double, 743
fixe, 584
intérieur, 168, 377
isolé, 169, 379
multiple, 743
ordinaire, 762
régulier, 745, 791, 814, 826, 829
selle, 693
simple, 743
singulier, 745, 791, 814, 826
stationnaire, 745
Polynôme, trigonométrique, 301
Préhilbertien, espace, 484
Principe du maximum, 159, 697
Produit
cartésien, 346
extérieur, 992
mixte, 988
scalaire, 479
vectoriel, 987
Prolongement par continuité, 189

Pythagore, théorème de, 485

Quadrique, 514

Règle
  d'Abel, 218, 302
  de Cauchy, 30, 241
  de D'Alembert, 34, 242
  de la loupe, 15
Rang
  d'une application linéaire, 445
  d'une forme quadratique, 452
  théorème du, 458
Rayon
  de convergence, 234
  spectral, 366
Relation de Chasles, 873
Repère local, 759, 775
Reste, 8
Riemann
  fonction zeta de, 53, 129, 140, 329
  intégrale, 861
  série, 24
Rotation, 175, 498
Rotationnel, 657
Ruban, de Möbius, 794

Série
  alternée, 29, 41
  d'applications, 113
  de Bertrand, 26
  de fonctions, 113
  de Fourier, 306
  de Mac-Laurin, 255, 265
  de Riemann, 24
  entière, 231, 260
  géométrique, 3
  harmonique, 10
  harmonique alternée, 4
  lacunaire, 232, 243
  numérique, 2
  trigonométrique, 29, 47, 126, 133, 301
Schwarz
  inégalité, 481
  théorème de, 650
Segment, 201, 678, 961
Semblables, matrices, 443
Semi-norme, 358
Sesquilinéaire, 439
Signature, 474
Simple
  arc, 715
  convergence, 71, 114, 216, 217
  courbe, 715
  nappe paramétrée, 717
Singulier, point, 745, 814, 826
Sinus
  complexe, 209
  hyperbolique complexe, 209
Somme
  d'une série, 114, 217
  partielle, 1, 114
Sous-espace
  vectoriel normé, 358
  métrique, 350
Sous-suite, 7, 395
Spectral, rayon, 366
Stationnaire, point, 745
Stirling, formule, 18
Stokes
  formule, 1005
  théorème de, 998
Subdivision, 856
Suite(s), 388
  adjacentes, 7, 131
  bornée, 392
  d'applications, 70
  de Cauchy, 82, 398
  de fonctions, 70
  des sommes partielles, 114
  extraite, 7, 395
Support d'un arc géométrique, 737
Surface
  cartésienne, 721
  implicite, 719
  paramétrée, 716
  régulière, 791
  tangente, 731
Surjection, 174
Sylvester, loi d'inertie de, 472
Symétrie
  affine, 911

vectorielle, 911
Symétrique, matrice, 447
Symbole de Kronecker, 463

Tangent(e)
à une courbe, 754
plan, 731
Taylor
formule de, 683, 684
Théorème
d'inversion globale, 665
d'inversion locale, 665
de Banach, 585
de Dirichlet, 322
de Fubini, 876
de Gauss, 469
de Jordan, 323
de la convergence dominée, 91
de Lebesgue, 91
de Poincaré, 963
de Pythagore, 485
de Schwarz, 650
de Stokes, 998
des fonctions implicites, 810, 812, 821
du point fixe, 585
du rang, 458
Topologie, 374
Translation, 175
Travail, 938
Trigonométrique
polynôme, 301
série, 133, 301
Trilinéaire, application, 985

Uniforme
continuité, 556
convergence, 73, 120, 216
de Cauchy, 82
subdivision, 857
Unitaire, vecteur, 358

Valeur d'adhérence, 394
Van der Waerden, fonction de, 142
Variables, séparées, 880
Vecteur
champ de, 939
isotrope, 457
normal, 731
unitaire, 358
Vectoriel, produit, 987
Voisinage, 165, 371
Volume, 874

Zeta, fonction, 53, 129, 140, 329

Analyse et algèbre

Cours de mathématiques de deuxième année avec exercices corrigés et illustrations avec Maple

Cet ouvrage, réunissant en un tout cohérent analyse et algèbre, s'adresse de manière plus spécifique aux élèves de deuxième année des cycles préparatoires intégrés des écoles d'ingénieurs mais il peut être utilisé avec profit par tout étudiant se destinant à des études supérieures d'ingénieur. Il est la suite naturelle de l'ouvrage «Algèbre et analyse, Cours de mathématiques de première année» publié dans la même collection par S. Balac et F. Sturm. Il est issu de l'enseignement dispensé par les auteurs dans la filière de premier cycle international ASINSA de l'INSA de Lyon. À ce titre, il ne constitue pas seulement une somme de connaissances mathématiques de deuxième année de l'enseignement supérieur mais vise à présenter de manière précise les résultats essentiels à une formation d'ingénieur généraliste.

L'ouvrage est divisé en 13 chapitres regroupés en 4 grandes parties: suites et séries de fonctions, algèbre bilinéaire, calcul différentiel et calcul intégral pour les fonctions de plusieurs variables. Chaque chapitre contient de courts exercices visant à tester la bonne compréhension des notions introduites et se termine par quelques exercices de synthèse. Une correction détaillée et commentée de tous les exercices est fournie en fin de chapitre. Le logiciel de calcul formel Maple est largement utilisé dans tout l'ouvrage pour illustrer les notions introduites.

Stéphane Balac est Maître de conférence en mathématiques à l'Ecole Nationale Supérieure des Sciences Appliquées et de Technologie (ENSSAT) de Lannion. Agrégé de mathématiques, il a enseigné pendant plusieurs années les mathématiques en premier cycle à l'INSA de Lyon dans la filière classique et dans la filière internationale ASINSA. Docteur de l'Université de Rennes 1, ses recherches portent sur la modélisation mathématique et l'étude de méthodes numériques en électromagnétisme dans le cadre d'applications en imagerie médicale et plus récemment en optronique et télécommunications optiques.

Laurent Chupin, ancien élève de l'Ecole Normale Supérieure de Cachan, est Maître de conférence en mathématiques à l'INSA de Lyon. Il enseigne les mathématiques en premier cycle, principalement dans la filière internationale ASINSA. Il est Docteur de l'Université de Bordeaux 1 et chercheur à l'Institut Camille Jordan (Lyon). Ses recherches portent sur l'étude théorique et numérique de certaines équations aux dérivées partielles issues du domaine de la mécanique des fluides.

PRESSES POLYTECHNIQUES ET UNIVERSITAIRES ROMANDES