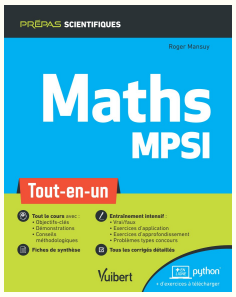
Ce document reprend la seconde édition de l'ouvrage **Mathématiques MPSI** (ISBN 978-2-311-40675-7), publiée par les éditions Vuibert en 2019.



Le livre est désormais épuisé et j'ai récupéré mes droits sur cet ouvrage ; je le laisse à disposition des étudiant·e·s mais je tiens à signaler deux réserves.

— Cet ouvrage répond à différents critères (souvent commerciaux) d'une collection et d'une maison d'édition ; il ne faut pas le considérer comme un cours « idéal ».
— Il est basé sur le programme avant la réforme entrée en vigueur à la rentrée 2021.

J'espère qu'il pourra être utile à quelques étudiant·e·s ou collègues.

Roger Mansuy
le 20/11/2021
roger.mansuy@gmail.com

PRÉPAS SCIENTIFIQUES

Maths
MPSI

Tout-en-un

Roger Mansuy est professeur en classe préparatoire scientifique
au lycée Saint-Louis à Paris.

Vuibert

SOMMAIRE

Chapitre 1 Bases mathématiques 13

Cours . . . 14
1. Écritures mathématiques . . . 14
2. Ensembles et applications . . . 19
3. Opérations entre parties . . . 25
4. Ensembles ordonnés . . . 30
5. L'ensemble des entiers naturels . . . 34
Fiche de synthèse . . . 37
Exercices . . . 39
Corrigés . . . 48

Chapitre 2 Nombres complexes 65

Cours . . . 66
1. Corps des nombres complexes . . . 66
2. Exponentielle complexe . . . 71
3. Équations algébriques . . . 83
Fiche de synthèse . . . 89
Exercices . . . 91
Corrigés . . . 96

Chapitre 3 Sommes et produits (finis) 109

Cours . . . 110
1. Règles de manipulation . . . 110
2. Quelques sommes classiques . . . 119
3. Formule du binôme . . . 123
Fiche de synthèse . . . 129
Exercices . . . 131
Corrigés . . . 135

Chapitre 4 Fonctions usuelles 149

Cours . . . 150
1. Propriétés des fonctions réelles . . . 150
2. Valeur absolue . . . 157

3. Exponentielles, logarithmes, puissances . . . 160
4. Exponentielle complexe, fonctions circulaires . . . 168
Fiche de synthèse . . . 174
Exercices . . . 176
Corrigés . . . 180

Chapitre 5 Équations différentielles 195

Cours . . . 196
1. Généralités . . . 196
2. Primitives . . . 198
3. Équations linéaires d'ordre 1 . . . 203
4. Équations linéaires d'ordre 2 . . . 206
Fiche de synthèse . . . 212
Exercices . . . 214
Corrigés . . . 218

Chapitre 6 Suites numériques 229

Cours . . . 230
1. Calculs explicites de suites . . . 230
2. Manipulations de suites . . . 234
3. Suites convergentes . . . 235
4. Relations de comparaison pour les suites . . . 244
5. Suites réelles . . . 250
6. Suites extraites . . . 257
Fiche de synthèse . . . 264
Exercices . . . 266
Corrigés . . . 271

Chapitre 7 Fonctions continues 283

Cours . . . 284
1. Études locales . . . 284
2. Fonctions continues sur un intervalle . . . 299
3. Fonctions continues sur un segment . . . 303
Fiche de synthèse . . . 309
Exercices . . . 311
Corrigés . . . 319

Chapitre 8 **Fonctions dérivables** **337**

Cours . . . 338
1. Fonctions une fois dérivables . . . 338
2. Fonctions de régularité supérieure . . . 344
3. Propriété des accroissements finis . . . 346
Fiche de synthèse . . . 357
Exercices . . . 359
Corrigés . . . 366

Chapitre 9 **Études locales et asymptotiques** **381**

Cours . . . 382
1. Relations de comparaison . . . 382
2. Développements limités . . . 385
3. Applications . . . 393
Fiche de synthèse . . . 396
Exercices . . . 398
Corrigés . . . 401

Chapitre 10 **Structures algébriques** **411**

Cours . . . 412
1. Groupes . . . 412
2. Anneaux et corps . . . 425
Fiche de synthèse . . . 430
Exercices . . . 431
Corrigés . . . 436

Chapitre 11 **Arithmétiques des entiers** **447**

Cours . . . 448
1. Divisibilité . . . 448
2. PGCD, PPCM . . . 452
3. Nombres premiers . . . 456
4. Systèmes congruentiels . . . 460
Fiche de synthèse . . . 463
Exercices . . . 465
Corrigés . . . 470

Chapitre 12 **Polynômes** **483**

Cours . . . 484
1. Définitions et premières propriétés . . . 484
2. Racines de polynômes . . . 488
3. Arithmétique des polynômes . . . 494
Fiche de synthèse . . . 502
Exercices . . . 504
Corrigés . . . 509

Chapitre 13 **Fractions rationnelles** **523**

Cours . . . 524
1. Définitions . . . 524
2. Décomposition en éléments simples (dénominateur scindé) . . . 528
3. Décomposition en éléments simples (cas général) . . . 532
Fiche de synthèse . . . 535
Exercices . . . 536
Corrigés . . . 540

Chapitre 14 **Espaces vectoriels** **549**

Cours . . . 550
1. Structure d'espace vectoriel . . . 550
2. Applications linéaires . . . 558
3. Endomorphismes remarquables . . . 564
Fiche de synthèse . . . 569
Exercices . . . 571
Corrigés . . . 575

Chapitre 15 **Dimension des espaces vectoriels** **585**

Cours . . . 586
1. Bases . . . 586
2. Dimension . . . 590
3. Applications linéaires en dimension finie . . . 597
Fiche de synthèse . . . 605
Exercices . . . 607
Corrigés . . . 614

Chapitre 16 **Calcul matriciel** **629**

Cours . . . 630
1. Matrices . . . 630
2. Produit matriciel . . . 637
3. Calcul de puissances . . . 642
Fiche de synthèse . . . 648
Exercices . . . 649
Corrigés . . . 653

Chapitre 17 **Équivalence des matrices et systèmes linéaires** **663**

Cours . . . 664
1. Matrices équivalentes . . . 664
2. Équivalence en pratique . . . 669
3. Applications aux systèmes linéaires . . . 676
4. Introduction à la similitude . . . 680
Fiche de synthèse . . . 684
Exercices . . . 686
Corrigés . . . 692

Chapitre 18 **Groupe symétrique** **705**

Cours . . . 706
1. Permutations . . . 706
2. Décomposition en cycles disjoints . . . 709
3. Signature . . . 712
Fiche de synthèse . . . 716
Exercices . . . 717
Corrigés . . . 720

Chapitre 19 **Déterminants** **729**

Cours . . . 730
1. Déterminant de n vecteurs dans une base . . . 730
2. Applications linéaires et matrices . . . 735
3. Quelques méthodes de calcul . . . 739
4. Applications . . . 745
Fiche de synthèse . . . 748
Exercices . . . 750
Corrigés . . . 758

Chapitre 20 **Espaces euclidiens** **773**

Cours . . . 774
1. Produit scalaire . . . 774
2. Orthogonalité . . . 780
3. Isométries vectorielles . . . 790
Fiche de synthèse . . . 798
Exercices . . . 800
Corrigés . . . 806

Chapitre 21 **Calcul intégral** **819**

Cours . . . 820
1. Intégrale sur un segment . . . 821
2. Sommes de Riemann . . . 826
3. Intégration et dérivation . . . 830
4. Formules de Taylor . . . 838
Fiche de synthèse . . . 841
Exercices . . . 843
Corrigés . . . 849

Chapitre 22 **Séries numériques** **867**

Cours . . . 868
1. Séries générales . . . 868
2. Séries à termes positifs . . . 872
3. Applications . . . 878
Fiche de synthèse . . . 882
Exercices . . . 883
Corrigés . . . 889

Chapitre 23 **Dénombrement** **905**

Cours . . . 906
1. Ensembles finis . . . 906
2. Exemples de dénombrements usuels . . . 912
Fiche de synthèse . . . 918
Exercices . . . 919
Corrigés . . . 924

Chapitre 24 **Probabilités sur un univers fini** **935**

Cours 936
1. Espace probabilisé fini 936
2. Conditionnement 941
3. Indépendance 946
Fiche de synthèse 949
Exercices 951
Corrigés 955

Chapitre 25 **Variables aléatoires** **963**

Cours 964
1. Généralités 964
2. Indépendance de variables aléatoires 967
3. Espérance, moments d'ordre supérieur 971
Fiche de synthèse 979
Exercices 981
Corrigés 986

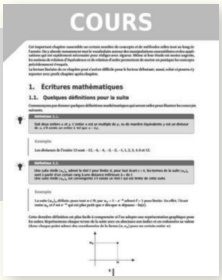
Index **1001**

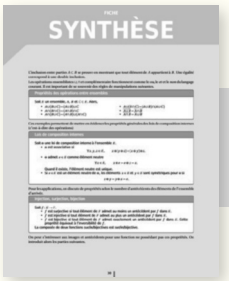
Mode d'emploi

Cet ouvrage, parfaitement conforme au programme de mathématiques en MPSI, vous propose les outils adaptés à la réussite de votre première année.

COURS

Tout le cours avec des rubriques claires pour un meilleur repérage de points importants à retenir, des conseils méthodologiques et de nombreuses démonstrations.





FICHES DE SYNTHÈSE

Des fiches de synthèse à la fin de chaque chapitre pour retenir l'essentiel

PICTOS DE REPÉRAGES

Pour un meilleur repérage des définitions

Conseils méthodologiques pour acquérir les bons réflexes et éviter les pièges

ENTRAÎNEMENT

Un entraînement intensif avec quatre typologies d'exercices : Vrai/Faux, application, approfondissement et problèmes types concours. Trois niveaux de difficulté clairement identifiés.

AIDES À LA RÉSOLUTION

Des aides à la résolution des exercices en cas de blocage à la lecture de l'énoncé

CORRIGÉS

Tous les corrigés détaillés pour comprendre chaque étape de résolution

Avant-propos

Partis pris de rédaction

Voici quelques-unes des contraintes qui ont guidé la rédaction de ce tout-en-un.

▹ Respecter l'**esprit du programme** en particulier la progression lorsque celle-ci est annoncée. Il ne s'agit pas de se limiter strictement au contenu explicitement mentionné dans le programme mais de fournir les moyens de comprendre en profondeur les notions mises en œuvre.

▹ Ajouter une **grande variété d'exemples** : l'importance des exemples est trop souvent sous-estimée ; ils permettent de voir les propositions et théorèmes en action. Pour les lecteurs, les exemples sont les premiers « exercices corrigés ».

▹ Illustrer les résultats avec des **schémas et dessins explicites**. Ces représentations nourrissent l'intuition.

▹ Structurer les chapitres pour faciliter l'assimilation et mettre en évidence les **articulations logiques**. L'organisation tient compte des attentes des lecteurs : certains chapitres ont été scindés pour améliorer la lecture en parallèle d'un cours en classe.

▹ Proposer des **exercices et problèmes** de difficultés variées permettant à chaque lecteur de progresser quelque soit son aisance initiale.

Recommandations aux lecteurs

Si ce livre est construit pour être utile, sa seule possession ne garantit en rien la réussite. Une étudiante ou un étudiant doit aussi apprendre à l'exploiter efficacement.

▹ Lire **le cours** sans passer les démonstrations ou les exemples ; au besoin, refaire un calcul ou un schéma sur un brouillon. J'approuve le mathématicien Paul Halmos qui explique « *Don't just read it ; fight it !* »

▹ Retenir les énoncés et l'importance de chaque hypothèse ; pour cela on peut notamment utiliser les questionnaires de type **Vrai/Faux** (à la fin du cours de chaque chapitre).

▹ Faire une **fiche de synthèse** (celles proposées dans ce livre sont purement indicatives et devraient en théorie servir de « vérification » après le travail personnel de synthèse) en faisant apparaître les points de cours importants, leurs articulations voire des méthodes de calcul ou de résolution.

▹ S'attaquer aux **exercices** avec ténacité et ne pas lire la correction proposée sans s'être posé les trois questions suivantes : quels sont les points de cours concernés ? quels sont les énoncés proches que je connais ? pourquoi mes tentatives n'aboutissent pas ?

▹ Travailler un **problème** dans la durée en appréciant bien les enchaînements de questions, la logique interne à l'énoncé. À la demande des lecteurs, j'ai ajouté des problèmes de niveau variable mais demandant toujours un investissement dans la durée.

Remerciements

Même si un seul nom figure sur la couverture de cet ouvrage, il ne faudrait pas oublier qu'un livre est toujours le fruit de nombreuses rencontres, d'échanges et de critiques. Merci aux étudiantes et étudiants, aux collègues des lycées Louis-le-Grand et Saint-Louis, aux lecteurs et à tous les amis qui m'ont permis de transformer mes premiers essais en cet ouvrage qui, je l'espère, vous conviendra.

Chapitre 1

Bases mathématiques

1. Écritures mathématiques — **2.** Ensembles et applications — **3.** Opérations entre parties — **4.** Ensembles ordonnés — **5.** L'ensemble des entiers naturels

Objectifs et compétences du programme

- Comprendre et savoir utiliser les notations d'un énoncé mathématique.
- Connaître les méthodes usuelles de raisonnement.
- Appréhender les manipulations entre ensembles, entre parties.
- Maîtriser les propriétés des applications.
- Rédiger un raisonnement par récurrence.

Guiseppe Peano, mathématicien italien (1848-1932)
Il aborde de nombreux domaines mathématiques :
- le théorème d'existence de solutions pour certaines équations différentielles non linéaires ;
- les axiomes définissant les entiers naturels à partir de la théorie des ensembles (qu'on citera dans ce chapitre) ;
- la première utilisation de nombreuses notations ensemblistes comme, par exemple, les symboles $\in$ ou $\subset$.

Tout au long de sa carrière, il a fait preuve d'une minutie extrême, très efficace pour la recherche, mais qui est devenue un grand travers pour enseigner. Retenons qu'il convient d'être rigoureux mais de ne pas trop l'être si cela interdit de communiquer les idées d'une démarche.

COURS

Cet important chapitre rassemble un certain nombre de concepts et de méthodes utiles tout au long de l'année. On y aborde notamment tout le vocabulaire autour des manipulations ensemblistes et des applications qui est rapidement nécessaire pour rédiger avec rigueur. Même si leur étude est moins urgente, les notions de relation d'équivalence et de relation d'ordre permettent de mettre en pratique les concepts précédemment évoqués.
La lecture linéaire de ce chapitre peut s'avérer difficile pour le lecteur débutant ; aussi, celui-ci pourra s'y reporter avec profit chapitre après chapitre.

1. Écritures mathématiques

1.1. Quelques définitions pour la suite

Commençons par donner quelques définitions mathématiques qui seront utiles pour illustrer les concepts suivants.

Définition 1.1.

Soit deux entiers n et p. L'entier n est un multiple de p, ou de manière équivalente p est un diviseur de n, s'il existe un entier k tel que $n = kp$.

Exemple

Les diviseurs de l'entier 12 sont $-12, -6, -4, -3, -2, -1, 1, 2, 3, 4, 6$ et 12.

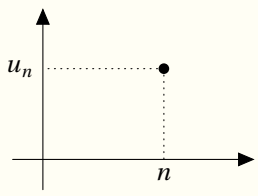
Définition 1.2.

Une suite réelle $(u_n)_n$ admet le réel ℓ pour limite si, pour tout écart $\varepsilon > 0$, les termes de la suite $(u_n)_n$ sont à partir d'un certain rang à une distance inférieure à ε de ℓ.
Une suite réelle $(u_n)_n$ est convergente s'il existe un réel ℓ qui est limite de cette suite.

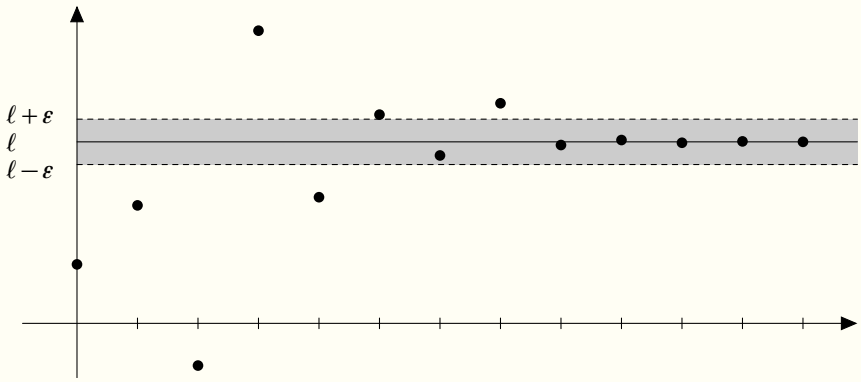
Exemple

La suite $(u_n)_n$ définie, pour tout $n \in \mathbb{N}$, par $u_n = 1 - e^{-n}$ admet $\ell = 1$ pour limite. En effet, l'écart entre u_n et ℓ est e^{-n} qui est plus petit que ε dès que n dépasse $-\ln(\varepsilon)$.

Cette dernière définition est plus facile à comprendre si l'on adopte une représentation graphique pour les suites. Représentons chaque terme de la suite avec en abscisses son indice et en ordonnées sa valeur (donc chaque point admet des coordonnées de la forme (n, u_n) pour un certain entier n).



Avec cette convention, la définition se représente alors de la façon suivante : la zone grisée indique les valeurs dont l'écart à ℓ est inférieur à ε et on observe qu'à partir d'un certain rang, les valeurs de la suite s'y trouvent.



1.2. Écriture avec quantificateurs

Il y a deux symboles mathématiques, appelés quantificateurs, pour écrire définitions et affirmations en mode mathématique.

- Le quantificateur universel $\forall$ signifie « quel que soit » ou « pour tout ».
- Le quantificateur existentiel $\exists$ signifie « il existe ».

Ces quantificateurs servent à introduire (au sens de définir) les variables que l'on va manipuler pour la suite.

À titre d'exemples, réécrivons les définitions de la section précédente.

Définition 1.3.

Soit deux entiers n et p. L'entier n est un multiple de p si

$$\exists k \in \mathbb{Z}, \qquad n = kp.$$

Définition 1.4.

Une suite réelle $(u_n)_n$ admet le réel ℓ pour limite si

$$\forall \varepsilon > 0, \quad \exists N \in \mathbb{N}, \quad \forall n \in \mathbb{N}, \qquad n \geq N \quad \Rightarrow \quad |u_n - \ell| \leq \varepsilon.$$

Une suite réelle $(u_n)_n$ est convergente si

$$\exists \ell \in \mathbb{R}, \quad \forall \varepsilon > 0, \quad \exists N \in \mathbb{N}, \quad \forall n \in \mathbb{N}, \qquad n \geq N \quad \Rightarrow \quad |u_n - \ell| \leq \varepsilon.$$

Dans cette écriture quantifiée, on a utilisé le symbole $\Rightarrow$ pour l'implication. On lit donc que si n est plus grand que N, alors on a $|u_n - \ell| \leq \varepsilon$.

Attention, les quantificateurs apparaissent dans les phrases mathématiques mais pas dans les phrases « de texte ». Ils sont placés avant le corps de l'affirmation (en particulier, jamais en fin de phrase) !

Conseils méthodologiques

L'ordre des quantificateurs est essentiel : on ne peut intervertir, sans justification précise, que deux quantificateurs de même nature (deux universels ou deux existentiels).
Dit autrement, toute variable définie par un quantificateur $\exists$ dépend *a priori* de toutes les variables définies auparavant.

Exemple

Une seule des deux phrases suivantes est correcte.

$$\forall x \in \mathbb{R}, \quad \exists y \in \mathbb{R}_+, \qquad y = x^2.$$
$$\exists y \in \mathbb{R}_+, \quad \forall x \in \mathbb{R}, \qquad y = x^2.$$

La première indique l'existence du carré d'un réel ; la seconde signifie qu'il existe un réel positif qui est le carré de tous les réels ce qui est faux puisque 0 et 1, par exemple, admettent des carrés différents.

Conseils méthodologiques

Pour nier une phrase avec quantificateurs, il faut appliquer les deux étapes suivantes :
- on remplace tout quantificateur existentiel par un quantificateur universel et réciproquement (sans en changer l'ordre) ;
- on nie la conclusion.

Ceci est déjà l'usage dans le langage courant ; la négation de l'affirmation « tous les élèves sont bruns » est « il existe un élève non brun ».

Exemple

Avec les deux définitions initiales, on obtient l'écriture des négations.
▹ L'entier n n'est pas un multiple de p si

$$\forall q \in \mathbb{N}, \qquad n \neq q.p,$$

▹ La suite réelle $(u_n)_n \in \mathbb{R}^{\mathbb{N}}$ n'est pas convergente si

$$\forall \ell \in \mathbb{R}, \quad \exists \varepsilon > 0, \quad \forall N \in \mathbb{N}, \quad \exists n \in \mathbb{N}, \qquad n \geq N \quad \text{et} \quad |u_n - \ell| > \varepsilon.$$

On remarque que l'on n'a utilisé que la négation de l'implication « $P \Rightarrow Q$ » est « P et non $-Q$ ».

Remarque

On utilise parfois la combinaison de caractères $\exists!$ pour indiquer « il existe un unique ». Par exemple, la phrase suivante indique que tout réel positif admet une unique racine carrée positive :

$$\forall x \in \mathbb{R}_+, \exists! y \in \mathbb{R}_+, \qquad x = y^2.$$

1.3. Écriture d'ensembles

Il est facile de noter et donc de manipuler un ensemble lorsque l'on dispose d'une notation pour celui-ci. Par exemple, on dispose des notations suivantes :
- $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ respectivement pour les ensembles des nombres entiers naturels, entiers relatifs, rationnels, réels et complexes.
- $[a, b]$ (respectivement $[a, b[$) l'intervalle formé des réels supérieurs ou égaux à a et inférieurs ou égaux (respectivement strictement inférieurs) à b.
- $[\![a, b]\!]$ l'ensemble des entiers supérieurs ou égaux à a et inférieurs ou égaux à b. On note quelquefois $\mathbb{N}_n$ l'intervalle d'entiers $[\![1, n]\!]$.
- B^A pour l'ensemble des applications de A dans B ; par exemple, $\mathbb{R}^{\mathbb{R}}$ désigne l'ensemble des fonctions de $\mathbb{R}$ dans $\mathbb{R}$ et $\mathbb{R}^{\mathbb{N}}$ l'ensemble des fonctions de $\mathbb{N}$ dans $\mathbb{R}$ c'est-à-dire des suites réelles.
- $\mathbb{R}[X]$ l'ensemble des polynômes à coefficients réels.

Tout au long du cours, on introduira de nouvelles notations pour les ensembles particulièrement remarquables ou fréquemment utilisés. Toutefois, certains ensembles sont d'usage trop rare pour mériter une notation standardisée et il faut alors recourir à une notation plus explicite. Il en existe deux qu'il convient de savoir lire.

▹ La notation en compréhension consiste à décrire notre ensemble comme une partie d'un ensemble plus grand formée des éléments vérifiant une ou plusieurs propriété(s). Par exemple, avec les écritures

$$A = \{n \in \mathbb{N},\ \exists k \in \mathbb{N},\ n = 2k\},$$
$$B = \{x \in \mathbb{R},\ \exists k \in \mathbb{Z},\ x = 10^k\}.$$

L'ensemble A contient exactement les entiers naturels n tels qu'il existe un entier k vérifiant $n = 2k$; B est l'ensemble des réels (en fait des rationnels) qui sont une puissance de 10. Dans cette notation, la virgule se lit « tel que » : A est l'ensemble des entiers naturels n tels qu'il existe un entier k dont n est le double.

▹ La notation en extension consiste à décrire notre ensemble par la liste de ses éléments (souvent paramétrés par d'autres variables). Par exemple, avec cette notation, les ensembles A et B précédents s'écrivent

$$A = \{2k,\ k \in \mathbb{N}\}$$
$$B = \{10^k,\ k \in \mathbb{Z}\}$$

La virgule ici se lit « avec » : A est l'ensemble des éléments de la forme $2k$ avec k un entier naturel.
Il ne faut pas confondre ces deux notations et bien identifier la nature des objets de l'ensemble.

Exemple

Soit q un entier. L'ensemble des multiples de q est noté $q\mathbb{Z}$; il est décrit des deux manières suivantes

$$q\mathbb{Z} = \{kq,\ k \in \mathbb{Z}\} = \{n \in \mathbb{Z},\ \exists k \in \mathbb{Z},\ n = kq\}.$$

1.4. Modes de raisonnement usuels

On se pose maintenant la question de savoir comment démontrer l'implication « $P \Rightarrow Q$ » entre deux propositions P et Q et on passe en revue quelques méthodes courantes. Avant tout, précisons un point de langage.

Définition 1.5.

Si « $P \Rightarrow Q$ », alors P est une condition suffisante pour Q ou de manière équivalente Q est une condition nécessaire pour P.

1.4.1. Déduction

C'est la méthode naturelle : on part de l'hypothèse et on aboutit à la conclusion par une suite d'implications simples.

Exemple

Soit n, k deux entiers. Montrons que si l'entier n est un multiple de k alors n^2 l'est également.

$$\begin{aligned} n \text{ est un multiple de } k &\Rightarrow \exists p \in \mathbb{N}, \quad n = kp \\ &\Rightarrow \exists p \in \mathbb{N}, \quad n^2 = k^2p^2 \\ &\Rightarrow \exists p \in \mathbb{N}, \quad n^2 = k(kp^2) \\ &\Rightarrow n^2 \text{ est un multiple de } k. \end{aligned}$$

Remarquons ici la suite d'implications qui traduit les différentes étapes élémentaires de notre raisonnement.

1.4.2. Disjonction de cas

Pour montrer l'implication « $(P_1$ ou P_2 ou … ou $P_n) \Rightarrow Q$ », on montre successivement les différentes implications « $P_k \Rightarrow Q$ » pour chaque $k \in [\![1, n]\!]$.

Exemple

Montrons que pour tout $n \in \mathbb{N}$, $n(n+1)(2n+1)$ est une multiple de 3. Pour cela, on sépare en trois cas selon les valeurs possibles 0, 1 ou 2 du reste r de la division de n par 3.

$$\begin{array}{lllll}
r=0 & \Rightarrow & \exists p \in \mathbb{N}, & n=3.p \\
 & \Rightarrow & \exists p \in \mathbb{N}, & n(n+1)(2n+1)=3.p(n+1)(2n+1) \\
 & \Rightarrow & n(n+1)(2n+1) \text{ est un multiple de 3.} & \\
 & & & \\
r=1 & \Rightarrow & \exists p \in \mathbb{N}, & n-1=3.p \\
 & \Rightarrow & \exists p \in \mathbb{N}, & 2n+1=3.(2p+1) \\
 & \Rightarrow & \exists p \in \mathbb{N}, & n(n+1)(2n+1)=3.(2p+1)n(n+1) \\
 & \Rightarrow & n(n+1)(2n+1) \text{ est un multiple de 3.} & \\
 & & & \\
r=2 & \Rightarrow & \exists p \in \mathbb{N}, & n-2=3.p \\
 & \Rightarrow & \exists p \in \mathbb{N}, & n+1=3.(p+1) \\
 & \Rightarrow & \exists p \in \mathbb{N}, & n(n+1)(2n+1)=3.n(p+1)(2n+1) \\
 & \Rightarrow & n(n+1)(2n+1) \text{ est un multiple de 3.} &
\end{array}$$

On conclut par disjonction des cas que la propriété est vraie pour tout entier n.

1.4.3. Contraposition et raisonnement par l'absurde

L'idée de la contraposition est de remplacer l'implication « $P \Rightarrow Q$ » par l'implication logiquement équivalente « non$-Q \Rightarrow$ non$-P$ ». C'est souvent intéressant lorsque la négation d'une proposition est plus simple à manipuler que la proposition de départ.

L'idée du raisonnement par l'absurde consiste à montrer que l'on peut déduire une contradiction de l'hypothèse P et non$-Q$.

Ces deux modes de raisonnement sont proches dans la rédaction d'une copie. Il convient de faire attention à ne pas les confondre.

Exemple

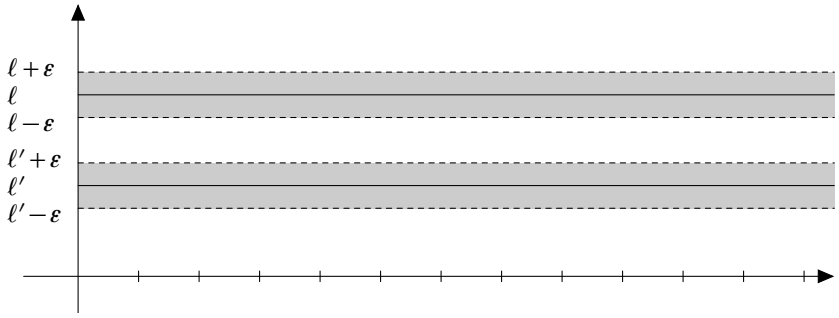
Montrons que si p est un nombre premier alors $\sqrt{p}$ est un irrationnel. Par l'absurde,

$$\begin{array}{lllll}
\sqrt{p} \in \mathbb{Q} & \Rightarrow & \exists a \in \mathbb{Z}, & \exists b \in \mathbb{N} \setminus \{0\}, & \sqrt{p} = \frac{a}{b} \\
 & \Rightarrow & \exists a \in \mathbb{Z}, & \exists b \in \mathbb{N} \setminus \{0\}, & p = \frac{a^2}{b^2} \\
 & \Rightarrow & \exists a \in \mathbb{Z}, & \exists b \in \mathbb{N} \setminus \{0\}, & p\,b^2 = a^2.
\end{array}$$

Les exposants de p dans la décomposition en facteurs premiers de a^2 et de b^2 sont pairs donc les exposants de p dans chacun des membres sont différents (pair d'un côté, impair de l'autre) : contradiction.

Exemple

Montrons par l'absurde qu'une suite convergente admet une unique limite.
Heuristiquement, voici le raisonnement à partir de notre représentation graphique : supposons que la suite $(u_n)_n$ admette deux limites distinctes ℓ et ℓ'. Pour toute « bande » autour de ℓ (respectivement autour de ℓ'), il existe un entier N (respectivement N') à partir duquel les termes de la suite définissent des points dans cette bande autour de ℓ (respectivement autour de ℓ'). Ainsi, à partir d'un certain rang, les termes de la suite sont dans les deux bandes. Comme on peut choisir les deux bandes disjointes (car $\ell \neq \ell'$), on obtient une contradiction.



Formalisons notre raisonnement. Soit ℓ et ℓ' deux réels distincts tels que

$$\forall \varepsilon > 0, \quad \exists N \in \mathbb{N}, \forall n \in \mathbb{N}, \qquad n \geq N \quad \Rightarrow \quad |u_n - \ell| \leq \varepsilon,$$

$$\forall \varepsilon > 0, \quad \exists N \in \mathbb{N}, \forall n \in \mathbb{N}, \qquad n \geq N \quad \Rightarrow \quad |u_n - \ell'| \leq \varepsilon.$$

Considérons $\varepsilon = \frac{1}{4}|\ell - \ell'|$ (de sorte à avoir notamment $2\varepsilon < |\ell - \ell'|$).
D'après la première condition, on fixe N tel que

$$\forall n \in \mathbb{N}, \qquad n \geq N \quad \Rightarrow \quad |u_n - \ell| \leq \varepsilon.$$

D'après la deuxième condition, on fixe N' tel que

$$\forall n \in \mathbb{N}, \qquad n \geq N' \quad \Rightarrow \quad |u_n - \ell'| \leq \varepsilon.$$

Donc, pour tout $n \geq \max(N, N')$, on a (la première inégalité s'appelle inégalité triangulaire)

$$|\ell - \ell'| \leq |u_n - \ell| + |u_n - \ell'| \leq 2\varepsilon \leq \frac{1}{2}|\ell - \ell'|.$$

D'où $|\ell - \ell'| = 0$ et donc $\ell = \ell'$. On a donc montré l'unicité de la limite par l'absurde.

2. Ensembles et applications

2.1. Égalité, inclusion

On suppose connue la relation d'appartenance $\in$ et on réécrit les définitions élémentaires.

Définition 1.6.

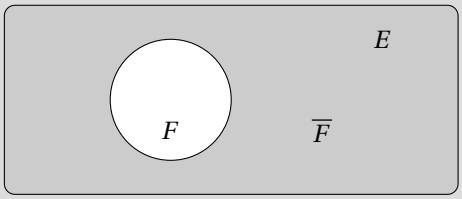
- Un ensemble F est inclus dans l'ensemble E si

$$\forall x \in F, \qquad x \in E.$$

  On note alors $F \subset E$ et on dit que F est une partie de E.
  L'ensemble des parties de E est noté $\mathscr{P}(E)$.
- Si $F \subset E$, alors le complémentaire de F dans E est l'ensemble $\overline{F}$ défini par

$$\forall x \in E, \qquad x \in \overline{F} \quad \Leftrightarrow \quad x \notin F.$$

  Sur la figure ci-dessous, le complémentaire de F correspond à la zone grisée.



D'autres notations sont possibles pour le complémentaire de F dans E : $F^{\complement}$, $E \setminus F$ (notation à privilégier s'il y a une ambiguïté sur l'ensemble dans lequel on prend le complémentaire) ou encore $\complement_E^F$.

- Deux parties F et G d'un même ensemble E sont égales si, et seulement si,

$$\forall x \in E, \qquad x \in F \quad \Leftrightarrow \quad x \in G.$$

Conseils méthodologiques

Ces définitions indiquent directement les méthodes correspondantes de démonstration.

- Pour établir l'inclusion $F \subset E$, on part d'un élément quelconque de F et on montre qu'il est dans E.
- Pour établir l'égalité entre ensembles F et G, il suffit de montrer séparément les deux inclusions $F \subset G$ et $G \subset F$.

Exemple

Les parties de $\{0,1\}$ sont l'ensemble vide, les deux singletons et la partie à deux éléments. Ainsi, l'ensemble des parties de $\{0,1\}$ est $\mathscr{P}(\{0,1\}) = \{\varnothing, \{0\}, \{1\}, \{0,1\}\}$.

Exemple

L'ensemble des suites convergeant vers 0 est une partie de l'ensemble des suites réelles.

Exemple

Soit deux entiers a et b. L'ensemble $a\mathbb{Z}$ des multiples de a est inclus dans l'ensemble $b\mathbb{Z}$ des multiples de b si, et seulement si, a est un multiple de b.
En effet, pour le sens direct, remarquons que $a \in b\mathbb{Z}$ donc a est un multiple de b. Réciproquement, si a est un multiple de b, alors tout multiple de a est un multiple de b.

2.2. Applications

Commençons par quelques définitions générales.

Définition 1.7.

Le produit cartésien de deux ensembles E et F est l'ensemble $E \times F$ des couples formés d'un élément de E puis d'un élément de F, c'est-à-dire : $\{(x,y),\ x \in E,\ y \in F\}$.
Si $E = F$, on note $E^2 = E \times E$.

Remarque

Le produit cartésien $E \times F$ est un ensemble de couples d'éléments donc, en particulier, ne contient ni E ni F.

On peut définir de même le produit de n ensembles comme l'ensemble des n-uplets dont les différentes coordonnées appartiennent aux ensembles de départ. La définition de E^n pour tout $n \in \mathbb{N} \setminus \{0\}$ est donc

Définition 1.8.

Une application d'un ensemble E dans un ensemble F est un ensemble de couples $G \subset E \times F$ tel que, pour tout $x \in E$, il existe un unique $y \in F$ vérifiant $(x, y) \in G$. L'ensemble des applications de E dans F est noté F^E ou $\mathscr{F}(E, F)$.

Cette définition est assez peu maniable. Aussi note-t-on souvent $f(x)$, l'unique $y \in F$ tel que $(x, y) \in G$ et on adopte la notation plus simple pour une application

$$f : \left\{ \begin{array}{ccc} E & \longrightarrow & F \\ x & \longmapsto & f(x). \end{array} \right.$$

Remarque

Pour montrer qu'un objet est bien une application, il faut démontrer que chaque élément de l'ensemble de départ est associé à un unique élément de l'ensemble d'arrivée. Par exemple, ce qui suit ne définit pas une application

$$n \longmapsto \left\{ \begin{array}{ll} 0 & \text{si 2 divise } n, \\ 1 & \text{si 3 divise } n, \\ 2 & \text{sinon.} \end{array} \right.$$

car l'élément $n = 6$ serait « envoyé » sur deux valeurs distinctes 0 ou 1 dans l'ensemble d'arrivée selon que l'on considère que 2 divise 6 ou que 3 divise 6.

Conseils méthodologiques

Pour vérifier qu'une application est définie « par morceaux », il faut vérifier que dans l'intersection de plusieurs cas, les différentes définitions coïncident.

Définition 1.9.

Soit $f : E \longrightarrow F$ une application et E' une partie de E. La restriction de f à E' est l'application

$$f_{|E'} : \left\{ \begin{array}{ccc} E' & \longrightarrow & F \\ x & \longmapsto & f(x). \end{array} \right.$$

On définit de même, lorsque cela est possible (c'est-à-dire quand, pour tout $x \in E$, $f(x) \in F'$), la corestriction $f^{|F'} : E \longrightarrow F'$ d'une application $f : E \longrightarrow F$ à une partie F' de l'ensemble d'arrivée.

Définition 1.10.

Soit $f : E \longrightarrow F$ et $g : F \longrightarrow G$ deux applications. La composée $g \circ f$ est l'application de E dans G qui associe, à tout $x \in E$, l'élément $g(f(x)) \in G$.

On vérifie rapidement les propriétés suivantes.

Proposition 1.11.

Soit f, g et $h \in E^E$. Alors,

$$f \circ (g \circ h) = (f \circ g) \circ h, \qquad f \circ Id_E = Id_E \circ f = f.$$

Remarque
En général, on ne peut pas permuter les applications, ou en d'autres termes, la composition n'est pas commutative : par exemple, avec $f : x \mapsto 2x$ et $g : x \mapsto x+1$, on observe pour tout $x \in \mathbb{R}$ que

$$f \circ g(x) = 2x+2 \neq 2x+1 = g \circ f(x).$$

2.3. Image, image réciproque

Définition 1.12.

Soit $f : E \longrightarrow F$ une application.
▹ L'image d'un élément $x \in E$ est l'élément $f(x) \in F$.
▹ L'image (directe) d'une partie $E' \subset E$ est la partie de F définie par

$$f(E') = \{f(x),\ x \in E'\},$$

c'est-à-dire l'ensemble des images des éléments de E'.

Exemple

Soit $f : \left\{ \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & x^2 \end{array} \right.$. Alors, on obtient

- $f(0) = 0$
- $f([0,1]) = [0,1]$
- $f([-1,1]) = [0,1]$

Définition 1.13.

Soit $f : E \longrightarrow F$ une application.
▹ Un antécédent d'un élément $y \in F$ est un élément $x \in E$ tel que $f(x) = y$.
▹ L'image réciproque d'une partie $F' \subset F$ est la partie de E définie par

$$f^{-1}(F') = \{x \in E,\ f(x) \in F'\},$$

c'est-à-dire l'ensemble de tous les antécédents des éléments de F'.

Remarque
L'ensemble des antécédents d'un élément $y \in F$ est par définition $f^{-1}(\{y\})$. Il peut être vide, réduit à un élément, ou beaucoup plus rempli !

Poursuivons l'exemple précédent.

Exemple

Soit $f : \left\{ \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & x^2 \end{array} \right.$. Alors, on calcule les résultats suivants

- $f^{-1}(\mathbb{R}_+) = \mathbb{R}$
- $f^{-1}(\mathbb{R}) = \mathbb{R}$
- $f^{-1}(\{0\}) = \{0\}$
- $f^{-1}(\{1\}) = \{-1,1\}$
- $f^{-1}([0,4]) = [-2,2]$

Remarque
Il faut faire attention à ne pas « simplifier » abusivement les expressions avec images directes et réciproques.
▹ On dispose des relations suivantes pour une application $f : E \to F$:

$$\forall A \in \mathscr{P}(E), \quad f^{-1}(f(A)) \supset A, \qquad \forall B \in \mathscr{P}(F), \quad f(f^{-1}(B)) \subset B.$$

En effet, tout élément de A est un antécédent de son image et tout élément de B est l'image d'un de ses antécédents.
▹ En revanche, les inclusions réciproques sont généralement fausses. Par exemple, avec l'application f de l'exemple précédent, on vérifie que

$$f(f^{-1}([-1,1])) = [0,1], \text{ et } f^{-1}(f([0,1])) = [-1,1].$$

2.4. Injection, surjection, bijection

Définition 1.14.

▹ Une application $f : E \longrightarrow F$ est injective si tout élément $y \in F$ a au plus un antécédent $x \in E$

$$\forall (x,x') \in E^2, \qquad f(x) = f(x') \quad \Rightarrow \quad x = x'.$$

▹ Une application $f : E \longrightarrow F$ est surjective si tout élément $y \in F$ a au moins un antécédent $x \in E$

$$\forall y \in F, \quad \exists x \in E, \qquad f(x) = y,$$

ou de manière équivalente $f(E) = F$.
▹ Une application $f : E \longrightarrow F$ est bijective si tout élément $y \in F$ a exactement un antécédent $x \in E$. En d'autres termes, f est bijective si, et seulement si, f est injective et surjective.

Conseils méthodologiques

Les définitions d'injectivité et de surjectivité donnent en même temps le moyen de montrer ces propriétés.

- Pour démontrer la surjectivité, on cherche si l'équation $y = f(x)$ admet une solution x.
- Pour l'injectivité, on étudie l'équation $f(x) = f(x')$ et, dans le cas d'une application injective, on doit en déduire $x = x'$.

Exemple

Les caractères injectifs, surjectifs et donc bijectifs dépendent des ensembles de départ et d'arrivée :

- $f : \left\{ \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & x^2 \end{array} \right.$ n'est ni injective, ni surjective.
- $g : \left\{ \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R}_+ \\ x & \longmapsto & x^2 \end{array} \right.$ est surjective mais pas injective.
- $h : \left\{ \begin{array}{ccc} \mathbb{R}_+ & \longrightarrow & \mathbb{R} \\ x & \longmapsto & x^2 \end{array} \right.$ est injective mais pas surjective.

L'application h est la restriction de f à $\mathbb{R}_+$ ou, de manière équivalente, f est un prolongement de h à $\mathbb{R}$.

Exemple

Soit E un ensemble et A une partie non vide de E. L'application

$$f : \left\{ \begin{array}{ccc} \mathscr{P}(E) & \longrightarrow & \mathscr{P}(E) \\ X & \longmapsto & X \cup A \end{array} \right.$$

n'est pas injective car $f(\varnothing) = f(A)$; elle n'est pas surjective car il n'existe pas de parties X telles que $f(X) = \varnothing$. En effet, pour tout $X \in \mathscr{P}(E)$, $A \subset f(X)$.

Exemple

L'application $(p,q) \mapsto (2p+1)2^q$ réalise une bijection de $\mathbb{N}^2$ dans $\mathbb{N} \setminus \{0\}$. En effet, tout entier $n \in \mathbb{N} \setminus \{0\}$ admet une décomposition en facteurs premiers uniques : en isolant le facteur 2 et en rassemblant les facteurs premiers impairs, on obtient l'existence et l'unicité de l'écriture de n comme $(2p+1)2^q$ donc l'existence et l'unicité de l'antécédent de n par f.

Définition 1.15.

Une application $f : E \to F$ est inversible s'il existe une application $g : F \to E$ qui satisfait à la fois $f \circ g = Id_F$, et $g \circ f = Id_E$.
Cette application g est appelée inverse de f.

On relie cette définition aux précédentes par la proposition suivante.

Proposition 1.16.

Une application est bijective si, et seulement si, elle est inversible.
L'inverse d'une application f bijective est appelé bijection réciproque de f et noté f^{-1}.

Démonstration

Soit une application $f : E \to F$.

($\Rightarrow$) Supposons f bijective. Considérons l'application $g : F \to E$ qui associe à un élément de F son unique antécédent par f. Alors, pour tout $x \in E$, $g(f(x))$ est l'unique antécédent de $f(x)$ donc est égal à x. De même, pour tout $y \in F$, $f(g(y))$ est l'image de l'antécédent de y donc est égal à y.

Par conséquent, $f \circ g = \mathrm{Id}_F$ et $g \circ f = \mathrm{Id}_E$.

($\Leftarrow$) Supposons f inversible d'inverse noté g. Pour tout $y \in F$, la condition $y = f(x)$ entraîne

$$g(y) = g(f(x)) = x$$

donc le seul antécédent de y est $g(y)$. En conclusion, tout élément de F admet un unique antécédent par f : $f : E \to F$ est bijective. ■

Remarque

Il convient de ne pas confondre $f^{-1}(\{y\})$ l'ensemble des antécédents de y (qui est toujours défini) et $f^{-1}(y)$ l'élément image de y par la bijection réciproque (qui n'est définie que si f est bijective).

Proposition 1.17.

La composée de deux applications injectives (respectivement surjectives, bijectives) est injective (respectivement surjective, bijective).

Démonstration

▹ Soit $f : E \to F$ et $g : F \to G$ deux applications injectives. Pour tout $x, x' \in E$, $g(f(x)) = g(f(x'))$ implique $f(x) = f(x')$ par injectivité de g puis $x = x'$ par injectivité de f.

▹ Soit $f : E \to F$ et $g : F \to G$ deux applications surjectives. Pour tout $y \in G$, il existe $z \in F$ tel que $g(z) = y$ car g est surjective. Ce z étant fixé, il existe $x \in E$ tel que $f(x) = z$ car f est surjective.

Par conséquent, il existe $x \in E$ tel que $g(f(x)) = y$. ■

3. Opérations entre parties

3.1. Lois de composition interne

Définition 1.18.

Une loi de composition interne sur un ensemble E est une application de E^2 dans E.

Une loi de composition interne est donc le terme mathématique pour décrire une opération sur un ensemble.

Exemples

Les opérations $+$, $\times$ définissent des lois de composition interne sur les ensembles de nombres usuels ($\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$) ; la composition $\circ$ est une loi de composition interne sur $\mathbb{R}^{\mathbb{R}}$, l'ensemble des applications de $\mathbb{R}$ dans $\mathbb{R}$.

Définissons quelques propriétés des lois de composition interne.

Définition 1.19.

Soit $\star$ une loi de composition interne sur un ensemble E.

▹ La loi $\star$ est commutative si

$$\forall (x,y) \in E^2, \qquad x \star y = y \star x.$$

▹ La loi $\star$ est associative si

$$\forall (x,y,z) \in E^3, \qquad x \star (y \star z) = (x \star y) \star z.$$

▹ La loi $\star$ admet un élément neutre si

$$\exists e \in E, \quad \forall x \in E, \qquad x \star e = e \star x = x.$$

▹ Si la loi $\star$ admet l'élément neutre e, un élément $x \in E$ admet un symétrique (dans E) pour $\star$ si

$$\exists y \in E, \qquad x \star y = y \star x = e.$$

Remarque

▹ Toutes les lois ne sont pas commutatives : la composition des fonctions est un contre-exemple.

▹ L'associativité indique qu'il n'y a pas d'ordre de priorité dans les calculs qui n'impliquent que $\star$: on peut donc les mener dans l'ordre que l'on veut.

▹ L'élément neutre, lorsqu'il existe, est unique. Il suffit de calculer $e \star e'$ lorsque e et e' sont des éléments neutres pour montrer que $e = e'$ en utilisant tour à tour les propriétés d'élément neutre de chaque élément.

▹ Lorsqu'une loi associative $\star$ admet un élément neutre e et qu'un élément x admet un symétrique, ce symétrique est unique. Soit y et y' des symétriques de x ; alors,

$$y = y \star e = y \star (x \star y') = (y \star x) \star y' = e \star y' = y'.$$

En fait, nous utiliserons ces propriétés pour traiter de nombreuses situations simultanément : l'étude des ensembles munis d'une loi de composition interne, associative, ayant un élément neutre et telle

que tout élément admette un symétrique se fera indépendamment de la nature des éléments (nombres, ensembles, applications, suites...). Cette riche idée d'abstraction sera développée avec l'introduction des structures algébriques.
On aura quelquefois besoin d'une propriété qui lie deux lois de composition interne.

Définition 1.20.

Soit E un ensemble muni de deux lois de composition interne $\diamond$ et $\star$. La loi $\star$ est distributive sur la loi $\diamond$ si

$$\forall (x,y,z) \in E^3, \qquad x \star (y \diamond z) = (x \star y) \diamond (x \star z),$$
$$\forall (x,y,z) \in E^3, \qquad (y \diamond z) \star x = (y \star x) \diamond (z \star x).$$

Par exemple, sur les ensembles de nombres usuels, la multiplication des nombres est distributive sur l'addition.

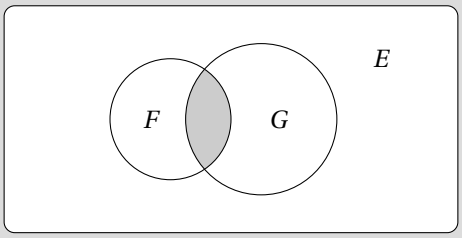
3.2. Intersection et réunion

Définition 1.21.

Soit F et G deux parties d'un ensemble E.
▹ L'intersection de F et G est la partie de E, notée $F \cap G$, formée des éléments communs à F et G, c'est-à-dire la partie définie par

$$\forall x \in E, \qquad x \in F \cap G \quad \Leftrightarrow \quad (x \in F \text{ et } x \in G).$$

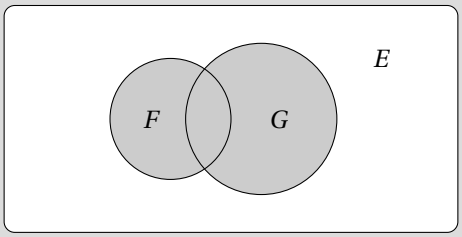
Sur la figure ci-dessous, l'intersection de F et G correspond à la zone grisée.



▹ La réunion de F et G est la partie de E, notée $F \cup G$, formée des éléments appartenant à F ou à G, c'est-à-dire la partie définie par

$$\forall x \in E, \qquad x \in F \cup G \quad \Leftrightarrow \quad (x \in F \text{ ou } x \in G).$$

Sur la figure ci-dessous, la réunion de F et G correspond à la zone grisée.



On vérifie que $F \cup G$ contient F et G, alors que $F \cap G$ est contenue dans F et G.

Exemple

Déterminons les parties F et G d'un ensemble E telles que $F \cap G = F \cup G$. On remarque que

$$F \subset F \cup G = F \cap G \subset G.$$

Par symétrie, $G \subset F$ et donc, par double-inclusion, $F = G$.

Proposition 1.22.

Les opérations $\cup$ et $\cap$ sont des lois de composition interne commutatives et associatives de $\mathscr{P}(E)$, d'éléments neutres $\varnothing$ et E respectivement.

L'associativité permet de définir les intersections et les réunions d'un nombre quelconque d'ensembles et on observe les propriétés de distributivité.

Proposition 1.23.

Soit F une partie de E et $(F_i)_{i\in I}$ une famille de parties de E. Alors

$$F \cap \Big(\bigcup_{i\in I} F_i\Big) = \bigcup_{i\in I}(F \cap F_i), \qquad F \cup \Big(\bigcap_{i\in I} F_i\Big) = \bigcap_{i\in I}(F \cup F_i).$$

On a évidemment, d'après la commutativité, les mêmes propriétés de distributivité à gauche.

Démonstration

Montrons la première égalité par double inclusion. La seconde se montre de manière analogue.

($\subset$) Soit $x \in F \cap \Big(\bigcup_{i\in I} F_i\Big)$. Comme $x \in \bigcup_{i\in I} F_i$, il existe $i \in I$ tel que $x \in F_i$. Ainsi, $x \in F \cap F_i \subset \bigcup_{i\in I}(F \cap F_i)$.

($\supset$) Soit $x \in \bigcup_{i\in I}(F \cap F_i)$. Par définition, il existe $i \in I$ tel que $x \in F \cap F_i \subset F_i$ donc $x \in \Big(\bigcup_{i\in I} F_i\Big)$.

En conclusion, $x \in F \cap \Big(\bigcup_{i\in I} F_i\Big)$. ■

3.3. Partitions

Définition 1.24.

Une famille $(F_i)_{i\in I}$ de parties de E est une partition de E si

- Aucune partie n'est vide
$$\forall i \in I, \qquad F_i \neq \varnothing.$$
- Deux parties distinctes sont disjointes
$$\forall (i,j) \in I^2, \qquad i \neq j \quad \Rightarrow \quad F_i \cap F_j = \varnothing.$$
- La réunion est égale à E
$$\bigcup_{i\in I} F_i = E.$$

Exemple

La famille $([n, n+1[)_{n\in\mathbb{Z}}$ est une partition de $\mathbb{R}$: les parties sont non vides, disjointes et de réunion $\mathbb{R}$. Il s'agit de partitionner l'ensemble des réels selon leur partie entière.

Exemple

Les parties $2\mathbb{Z}$ et $\{2k+1,\ k\in\mathbb{Z}\}$ forment une partition de $\mathbb{Z}$ (selon la parité).

En fait, ces exemples relèvent d'un cadre plus général que nous détaillons maintenant.

Définition 1.25.

Une relation d'équivalence sur un ensemble E est une relation binaire $\sim$ qui vérifie les propriétés suivantes :

- $\sim$ est réflexive, c'est-à-dire, tout élément $x\in E$ est en relation avec lui-même pour $\sim$:
$$\forall x\in E,\qquad x\sim x.$$
- $\sim$ est symétrique, c'est-à-dire pour tous les éléments x et $y\in E$ tels que x est en relation avec y, on a aussi y en relation avec x
$$\forall x,y\in E,\qquad x\sim y\qquad\Leftrightarrow\qquad y\sim x.$$
- $\sim$ est transitive, c'est-à-dire pour tous les éléments x, y et $z\in E$ tels que x est en relation avec y et y est en relation avec z, on a aussi x en relation avec z
$$\forall x,y,z\in E,\qquad x\sim y\ \text{ et }\ y\sim z\qquad\Rightarrow\qquad x\sim z.$$

Deux éléments en relation sont dits équivalents. La classe d'équivalence d'un élément x pour la relation $\sim$ est l'ensemble des éléments de E équivalents à x pour $\sim$, à savoir
$$\{y\in E,\ y\sim x\}.$$

Exemple Congruence sur les entiers

Soit $p\in\mathbb{N}\setminus\{0\}$. Deux entiers m et n sont congruents modulo p si p divise $m-n$. On note alors $m\equiv n[p]$ ou plus simplement $m=n[p]$ en gardant à l'esprit que ce signe d'égalité n'est pas une véritable égalité.
On vérifie immédiatement que la congruence modulo p est une relation d'équivalence sur $\mathbb{Z}$ et que la classe d'équivalence de a est
$$\{a+bp,\ b\in\mathbb{Z}\}.$$

Généralisons l'exemple précédent aux réels (où, rappelons-le, la notion de divisibilité est peu pertinente puisque tout réel non nul divise tous les réels).

Exemple

Deux réels x et y sont congruents modulo 2π si la différence $x-y$ est un multiple entier de 2π. On définit ainsi une relation d'équivalence sur $\mathbb{R}$ (le choix de 2π correspond à une utilisation courante pour les fonctions trigonométriques, on peut bien évidemment choisir n'importe quel autre réel non nul).

Proposition 1.26.

Soit $\sim$ une relation d'équivalence sur un ensemble E. Les classes d'équivalence pour $\sim$ forment une partition de E.

Démonstration
Il est clair qu'une classe d'équivalence est non vide et que E est inclus dans la réunion des classes d'équivalence (car chaque élément est inclus dans sa propre classe d'équivalence). Montrons que deux classes d'équivalence distinctes sont disjointes. Pour cela, considérons x et $y \in E$ tels que les classes d'équivalence de x et y ne soient pas disjointes et $z \in E$ à la fois équivalent à x et y. Alors, par transitivité, la classe de x est la classe de z ; de même, la classe de y est la classe de z. Ainsi, les classes de x et de y sont égales. ■

3.4. Fonctions indicatrices

Définition 1.27.

La fonction indicatrice d'une partie F de E est l'application $\mathbb{1}_F : E \to \{0,1\}$ définie, pour tout $x \in E$, par

$$\mathbb{1}_F(x) = \begin{cases} 1 & \text{si } x \in F \\ 0 & \text{sinon.} \end{cases}$$

Exemple

Les fonctions indicatrices de E et de $\varnothing$ sont respectivement les fonctions constantes sur E et égales à 1 et à 0 respectivement.

On vérifie immédiatement avec cette définition que $F = \{x \in E,\ \mathbb{1}_F(x) = 1\}$. On en déduit que deux parties ayant la même fonction indicatrice sont égales.

Conseils méthodologiques

Pour montrer une égalité entre ensembles, il suffit de montrer l'égalité des fonctions indicatrices.

Proposition 1.28.

Soit F et G des parties de E. Alors,

$$\mathbb{1}_{\overline{F}} = 1 - \mathbb{1}_F, \qquad \mathbb{1}_{F\cap G} = \mathbb{1}_F \mathbb{1}_G, \qquad \mathbb{1}_{F\cup G} = \mathbb{1}_F + \mathbb{1}_G - \mathbb{1}_F \mathbb{1}_G.$$

Démonstration
▹ Pour tout $x \in E$,

$$\begin{aligned} x \in \overline{F} &\Leftrightarrow x \notin F \\ &\Leftrightarrow \mathbb{1}_F(x) = 0 \\ &\Leftrightarrow (1 - \mathbb{1}_F)(x) = 1. \end{aligned}$$

Ainsi, $\mathbb{1}_{\overline{F}} = 1 - \mathbb{1}_F$.
▹ Pour tout $x \in E$,

$$\begin{aligned} x \in F \cap G &\Leftrightarrow x \in F \text{ et } x \in G \\ &\Leftrightarrow \mathbb{1}_F(x) = 1 \text{ et } \mathbb{1}_G(x) = 1 \\ &\Leftrightarrow (\mathbb{1}_F \mathbb{1}_G)(x) = 1. \end{aligned}$$

Ainsi, $\mathbb{1}_{F\cap G} = \mathbb{1}_F \mathbb{1}_G$.

▷ Pour tout $x \in E$,

$$\begin{aligned} x \in F \cup G &\Leftrightarrow x \in F \text{ ou } x \in G \\ &\Leftrightarrow \mathbb{1}_F(x) = 1 \text{ ou } \mathbb{1}_G(x) = 1 \\ &\Leftrightarrow (\mathbb{1}_F + \mathbb{1}_G - \mathbb{1}_F \mathbb{1}_G)(x) = 1. \end{aligned}$$

Ainsi, $\mathbb{1}_{F\cup G} = \mathbb{1}_F + \mathbb{1}_G - \mathbb{1}_F \mathbb{1}_G$. ■

Proposition 1.29. Lois de De Morgan

Soit F et G deux parties de E.

$$\overline{F \cap G} = \overline{F} \cup \overline{G}, \qquad \overline{F \cup G} = \overline{F} \cap \overline{G}.$$

Démonstration

Les deux égalités sont équivalentes en utilisant que le complémentaire du complémentaire d'une partie est égal à la partie. Il suffit d'en établir une, ce que nous allons faire en comparant les fonctions caractéristiques.

$$\mathbb{1}_{\overline{F\cap G}} = 1 - \mathbb{1}_{F\cap G} = 1 - \mathbb{1}_F \mathbb{1}_G = 1 - (1 - \mathbb{1}_{\overline{F}})(1 - \mathbb{1}_{\overline{G}}) = \mathbb{1}_{\overline{F}} + \mathbb{1}_{\overline{G}} - \mathbb{1}_{\overline{F}} \mathbb{1}_{\overline{G}} = \mathbb{1}_{\overline{F}\cup\overline{G}}.$$

■

4. Ensembles ordonnés

4.1. Relation d'ordre

Définition 1.30.

Une relation d'ordre sur un ensemble E non vide est une relation binaire $\prec$ qui vérifie les trois propriétés suivantes

- $\prec$ est réflexive
$$\forall x \in E, \qquad x \prec x.$$
- $\prec$ est antisymétrique
$$\forall x, y \in E, \qquad x \prec y \text{ et } y \prec x \quad \Rightarrow \quad x = y.$$
- $\prec$ est transitive
$$\forall x, y, z \in E, \qquad x \prec y \text{ et } y \prec z \quad \Rightarrow \quad x \prec z.$$

Le couple $(E, \prec)$ est alors appelé ensemble ordonné.

Voici quelques exemples de relations d'ordre usuelles.

Exemple

La relation de comparaison « inférieur ou égal » $\leq$ est une relation d'ordre sur $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $[a, b]$ En revanche, la relation « inférieur strictement » $<$ ne l'est pas puisqu'il lui manque la réflexivité.

Exemple

L'inclusion $\subset$ est une relation d'ordre sur $\mathscr{P}(E)$ avec E non vide (l'antisymétrie correspond à la définition de l'égalité entre parties *via* la double-inclusion).

Exemple

La divisibilité définit une relation d'ordre sur $\mathbb{N} \setminus \{0\}$.

Exemple Ordre lexicographique

L'ordre lexicographique (ou ordre du dictionnaire) est une relation d'ordre sur $\mathbb{N}^p$.
Rappelons que cet ordre consiste à comparer les p-uplets en comparant les premières composantes puis en cas d'égalité en passant à la suivante comme on le fait pour classer les mots dans un dictionnaire ; ainsi $(1,7,8,9)$ est plus petit pour cet ordre que $(1,7,9,3)$: les deux premières coordonnées sont égales dans les deux 4-uplet donc on considère la troisième : on conclut en remarquant que 8 est plus petit que 9.

Parmi ces relations d'ordre, toutes n'ont pas les mêmes propriétés.

Définition 1.31.

Une relation d'ordre $\prec$ sur E est totale si l'on peut comparer tous les éléments de E

$$\forall x, y \in E, \qquad x \prec y \text{ ou } y \prec x.$$

Une relation d'ordre qui n'est pas totale est partielle.

Parmi les exemples précédents, l'inclusion et la divisibilité sont des relations d'ordre partielles.

Exemple

Si $(E, \prec)$ est un ensemble ordonné et F une partie de E, alors $\prec$ définit une relation d'ordre sur F, appelée relation d'ordre induite.
Si la relation $\prec$ sur E est totale, alors la relation induite sur une partie F est elle aussi totale (en revanche, la réciproque est fausse).

4.2. Plus petit élément, minorant

Définition 1.32.

Soit $(E, \prec)$ un ensemble ordonné et A une partie de E.
▹ Un majorant de A est un élément $m \in E$ tel que

$$\forall x \in A, \qquad x \prec m.$$

▹ Un minorant de A est un élément $m \in E$ tel que

$$\forall x \in A, \qquad m \prec x.$$

▹ Si A admet un majorant (respectivement un minorant), alors A est majorée (respectivement minorée). Si A est majorée et minorée, alors A est bornée.

Exemples

▹ La partie $]-\infty, \pi]$ est majorée par 4 ou 1789 mais aussi par π ; en revanche, elle n'est pas minorée.
▹ La partie $\{1, 2, 4, 10^{100}\}$ est bornée.
▹ L'ensemble $\mathbb{Z}$ n'est ni majoré, ni minoré.
▹ Toutes les parties de $\mathscr{P}(E)$ sont bornées pour l'inclusion (minorées par $\varnothing$, majorées par E).

Définition 1.33.

Soit $(E, \prec)$ un ensemble ordonné et A une partie de E.
▹ Un élément $m \in A$ est un plus grand élément de A si

$$\forall x \in A, \qquad x \prec m.$$

Autrement dit, un plus grand élément de A est un majorant appartenant à A.
▹ Un élément $m \in A$ est un plus petit élément de A si

$$\forall x \in A, \qquad m \prec x.$$

Autrement dit, un plus petit élément de A est un minorant appartenant à A.

Remarque

Les plus petits éléments et plus grands éléments n'existent pas forcément. Par exemple, l'ensemble $]0,1[$ (muni de la relation $\leq$) n'admet ni plus grand ni plus petit éléments.
Supposons que a est un plus petit élément de $]0,1[$; alors $\frac{a}{2}$ est un élément de $]0,1[$ strictement plus petit que a : contradiction.
De même, supposons que a est un plus grand élément de $]0,1[$; alors $\frac{1+a}{2}$ est un élément de $]0,1[$ strictement plus grand que a : contradiction.

Proposition 1.34.

Quand ils existent, les plus grands et plus petits éléments d'une partie A sont uniques.
On les note respectivement $\max A$ et $\min A$.

Démonstration

Supposons qu'il existe m et m' deux plus grands éléments. Alors, comme m' est un plus grand élément, $m \prec m'$; de même, $m' \prec m$ puis, par antisymétrie de $\prec$, $m = m'$. ■

Proposition 1.35.

Soit A et B deux parties d'un ensemble ordonné $(E, \prec)$ admettant un plus grand élément telles que $A \subset B$. Alors, $\max A \prec \max B$.

Démonstration

Par définition du plus grand élément de B,

$$\forall x \in B, \qquad x \prec \max B.$$

Or, $\max A \in B$ donc $\max A \prec \max B$. ■

4.3. Borne supérieure, borne inférieure

Définition 1.36.

Soit A un partie d'un ensemble ordonné $(E, \prec)$.
▹ Si A est majorée, la borne supérieure de A est, lorsqu'elle existe, le plus petit majorant de A. On la note $\sup A$.
▹ Si A est minorée, la borne inférieure de A est, lorsqu'elle existe, le plus grand minorant de A. On la note $\inf A$.

On remarque que contrairement à $\max A$ et $\min A$, $\sup A$ et $\inf A$ ne sont *a priori* pas des éléments de A. En fait, les bornes supérieure et inférieure appartiennent à l'ensemble si, et seulement si, elles coïncident avec les plus grand et plus petit éléments.

Exemples

Avec $A = \{x \in \mathbb{Q},\ x^2 < 2\}$, la borne supérieure dans $\mathbb{R}$ est $\sup A = \sqrt{2} \notin A$. Il n'y a pas de borne supérieure dans $\mathbb{Q}$.
Avec $B = \{x \in \mathbb{Z},\ x^2 < 2\}$, la borne supérieure dans $\mathbb{Z}$, $\mathbb{Q}$ ou $\mathbb{R}$ est $\sup B = 1 \in B$.

Proposition 1.37.

Soit A et B deux parties d'un ensemble ordonné $(E, \prec)$ admettant une borne supérieure telles que $A \subset B$. Alors, $\sup A \prec \sup B$.

Démonstration
Par définition de la borne supérieure de B,

$$\forall x \in B, \qquad x \prec \sup B.$$

Ainsi, comme $A \subset B$, les éléments de A appartiennent à B donc

$$\forall x \in A, \qquad x \prec \sup B.$$

Par conséquent, $\sup B$ est un majorant de A donc plus grand que le plus petit des majorants $\sup A$. ■

4.4. Applications croissantes

Définition 1.38.

Soit $f : (E, \prec_E) \to (F, \prec_F)$ une application entre deux ensembles ordonnés.
▹ L'application f est croissante si

$$\forall x, y \in E, \qquad x \prec_E y \quad \Rightarrow \quad f(x) \prec_F f(y).$$

▹ L'application f est strictement croissante si

$$\forall x, y \in E, \qquad x \prec_E y \text{ et } x \neq y \quad \Rightarrow \quad f(x) \prec_F f(y) \text{ et } f(x) \neq f(y).$$

On définit de même les propriétés de décroissance et de décroissance stricte.

Exemple

Soit E un ensemble et A une partie non vide de E. L'application

$$f : \left\{ \begin{array}{ccc} \mathscr{P}(E) & \to & \mathscr{P}(E) \\ X & \mapsto & X \cup A \end{array} \right.$$

est croissante mais pas strictement croissante (car $f(\varnothing) = f(A)$ alors que $\varnothing \subsetneq A$).

Proposition 1.39.

Soit $f : (E, \prec_E) \to (F, \prec_F)$ où $\prec_E$ est un ordre total. Si f est strictement croissante, alors f est injective.

Démonstration
Supposons, par l'absurde, qu'il existe $x \neq y$ tels que $f(x) = f(y)$.
Si $x \prec_E y$, alors, par croissance stricte de f, $f(x) \prec_E f(y)$ et $f(x) \neq f(y)$ ce qui contredit l'hypothèse $f(x) = f(y)$.
On procède de même pour le cas $y \prec_E x$. ■

5. L'ensemble des entiers naturels

5.1. Propriétés de Peano

L'idée de cette partie n'est pas de définir $\mathbb{N}$ (ce qui relève de la théorie générale des ensembles et qui est délicat) mais de montrer quelles propriétés seront utiles. Voilà les propriétés que nous admettrons.

Proposition 1.40.

L'ensemble $\mathbb{N}$ des entiers est un ensemble non vide ordonné tel que :
- toute partie non vide de $\mathbb{N}$ admet un plus petit élément ;
- toute partie non vide majorée admet un plus grand élément ;
- $\mathbb{N}$ n'admet pas de plus grand élément.

Remarque
En fait, on peut montrer que si A est un ensemble vérifiant ces propriétés, alors il existe une bijection croissante de A vers $\mathbb{N}$.

D'après ces propriétés, il est évident que $\mathbb{N}$ est alors totalement ordonné (car un ensemble à deux éléments admet un plus petit élément), que $\mathbb{N}$ admet un plus petit élément (car c'est une partie non vide de $\mathbb{N}$) que l'on note 0.
Pour tout $n \in \mathbb{N}$, on définit $n+1$ comme le plus petit élément strictement supérieur à n : ceci est possible car
- n n'est pas le plus grand élément de $\mathbb{N}$ donc il existe un élément strictement plus grand ;
- l'ensemble des éléments strictement plus grands que n est non vide donc admet un plus petit élément.

De même, on définit $n-1$ pour tout élément $n \in \mathbb{N}$ différent de 0.

5.2. Principe de récurrence

Définition 1.41.

Un prédicat sur $\mathbb{N}$ est une application de $\mathbb{N}$ dans l'ensemble des booléens : vrai et faux.

Proposition 1.42. Principe de récurrence

Soit $\mathscr{P}$ un prédicat sur $\mathbb{N}$ tel que :
- $\mathscr{P}(0)$ est vrai ;
- pour tout $n \in \mathbb{N}$, si $\mathscr{P}(n)$ est vrai, alors $\mathscr{P}(n+1)$ est vrai.

Alors, $\mathscr{P}(n)$ est vrai pour tout $n \in \mathbb{N}$.

En remplaçant la première condition par $\mathscr{P}(m)$ est vrai, on obtient la conclusion $\mathscr{P}(n)$ est vrai pour tout entier $n \geq m$.

Démonstration
Considérons l'ensemble $A = \{n \in \mathbb{N},\ \mathscr{P}(n) \text{ est faux}\}$. Si A est non vide, il admet un plus petit élément $m = \min A$ qui est différent de 0. Alors, $m-1 \notin A$ et donc par la seconde propriété, $m \notin A$: contradiction donc A est vide. ■

Proposition 1.43. Principe de récurrence forte

Soit $\mathscr{P}$ un prédicat sur $\mathbb{N}$ tel que :
- $\mathscr{P}(0)$ est vrai ;
- pour tout $n \in \mathbb{N}$, si $\mathscr{P}(k)$ est vrai pour tout $k \leq n$, alors $\mathscr{P}(n+1)$ est vrai.

Alors, $\mathscr{P}(n)$ est vrai pour tout $n \in \mathbb{N}$.

La preuve est identique à la précédente avec le prédicat $\mathscr{Q}$ tel que $\mathscr{Q}(n)$ est vrai si, et seulement si, $\mathscr{P}(k)$ est vrai pour tout $k \leq n$.
Regardons quelques exemples.

Exemple

Montrons que tout entier supérieur à 2 admet une décomposition comme produit de nombres premiers.
Posons, pour tout $n \geq 2$, $\mathscr{P}(n)$: « n admet une décomposition en produit de nombres premiers ».
- $\mathscr{P}(2)$ est vrai car 2 est un nombre premier.
- Soit $n \geq 2$. Supposons $\mathscr{P}(k)$ vrai pour tout $k \leq n$ et montrons $\mathscr{P}(n+1)$ par disjonction de cas :
  - Si $n+1$ est premier, alors $\mathscr{P}(n+1)$ est vrai.
  - Si $n+1$ est composé, alors il existe p et q entre 2 et n tels que $n+1 = p.q$. Or, par hypothèse de récurrence, $\mathscr{P}(p)$ et $\mathscr{P}(q)$ sont vrais donc $n+1$ est un produit de nombres premiers et $\mathscr{P}(n+1)$ est vrai.

On conclut avec le principe fort de récurrence.

Exemple

Notons, pour tout $n \in \mathbb{N} \setminus \{0\}$, S_n la somme des cubes des entiers entre 1 et n et montrons par récurrence sur $n \in \mathbb{N} \setminus \{0\}$ que $S_n = \left(\frac{n(n+1)}{2}\right)^2$.
Définissons, pour tout $n \in \mathbb{N} \setminus \{0\}$, $\mathscr{P}(n)$: « $S_n = \left(\frac{n(n+1)}{2}\right)^2$ ».
- $\mathscr{P}(1)$ est vrai car $1 = \left(\frac{1(1+1)}{2}\right)^2$.
- Soit $n \in \mathbb{N} \setminus \{0\}$. Supposons $\mathscr{P}(n)$ vrai et montrons $\mathscr{P}(n+1)$

$$\begin{aligned} S_{n+1} &= S_n + (n+1)^3 \\ &= \left(\frac{n(n+1)}{2}\right)^2 + (n+1)^3 \\ &= (n+1)^2\left(\frac{n^2}{4} + n + 1\right) = \frac{(n+1)^2(n+2)^2}{4}. \end{aligned}$$

D'où le résultat.

On conclut avec le principe de récurrence.

Exemple

Soit $(F_n)_n$ la suite de Fibonacci définie par $F_0 = F_1 = 1$ et, pour tout $n \in \mathbb{N}$, $F_{n+2} = F_{n+1} + F_n$. Définissons, pour tout $n \in \mathbb{N}$, S_n la somme des termes de cette suite d'indice inférieur ou égal à n. Montrons par récurrence sur $n \in \mathbb{N}$, la propriété $\mathscr{P}(n)$: « $S_n = F_{n+2} - 1$ ».

- $\mathscr{P}(0)$ est vrai car $F_0 = F_2 - F_1 = F_2 - 1$.
- Soit $n \in \mathbb{N}$. Supposons $\mathscr{P}(n)$ vrai et montrons $\mathscr{P}(n+1)$.

$$\begin{aligned} S_{n+1} &= S_n + F_{n+1} \\ &= F_{n+2} - 1 + F_{n+1} = F_{n+3} - 1. \end{aligned}$$

On conclut avec le principe de récurrence.

En fait, on remarque que l'on peut procéder « par récurrence » sur d'autres ensembles que $\mathbb{N}$ et que ce principe d'induction sera fort utile en informatique.

FICHE

SYNTHÈSE

L'inclusion entre parties $A \subset B$ se prouve en montrant que tout élément de A appartient à B. Une égalité correspond à une double inclusion.

Les opérations ensemblistes $\cup$, $\cap$ et complémentaire fonctionnent comme le *ou*, le *et* et le *non* du langage courant. Il est important de se souvenir des règles de manipulations suivantes.

Propriétés des opérations entre ensembles

Soit E un ensemble, A, B et $C \subset E$. Alors,

- $A \cup (B \cup C) = (A \cup B) \cup C$
- $A \cap (B \cap C) = (A \cap B) \cap C$
- $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
- $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
- $\overline{A \cup B} = \overline{A} \cap \overline{B}$
- $\overline{A \cap B} = \overline{A} \cup \overline{B}$

Ces exemples permettent de mettre en évidence les propriétés générales des lois de composition internes (c'est-à-dire des opérations).

Lois de composition internes

Soit $\star$ une loi de composition interne à l'ensemble E.
▹ $\star$ est associative si

$$\forall x, y, z \in E, \qquad x \star (y \star z) = (x \star y) \star z.$$

▹ $\star$ admet $e \in E$ comme élément neutre

$$\forall x \in E, \qquad x \star e = e \star x = x.$$

Quand il existe, l'élément neutre est unique.
▹ Si $e \in E$ est un élément neutre de $\star$, les éléments $x \in E$ et $y \in E$ sont symétriques pour $\star$ si

$$x \star y = y \star x = e.$$

Pour les applications, on discute de propriétés selon le nombre d'antécédents des éléments de l'ensemble d'arrivée.

Injection, surjection, bijection

Soit $f : E \to F$ une application.
▹ f est surjective si tout élément de F admet au moins un antécédent par f dans E.
▹ f est injective si tout élément de F admet au plus un antécédent par f dans E.
▹ f est bijective si tout élément de F admet exactement un antécédent par f dans E. Cette propriété équivaut à l'inversibilité de f.
La composée de deux fonctions sur/in/bijectives est sur/in/bijective.

On peut s'intéresser aux images et antécédents pour une fonction ne possédant pas ces propriétés. On introduit alors les parties suivantes.

Image directe et image réciproque

Soit $f : E \longrightarrow F$ une application.
▹ L'image directe par f d'une partie $A \subset E$ est l'ensemble des images des éléments de A par f. On la note $f(A)$.
▹ L'image réciproque par f d'une partie $B \subset F$ est l'ensemble des antécédents des éléments de B par f. On la note $f^{-1}(B)$ (et ceci est cohérent avec le cas où f est bijective).

Généralisant les relations de congruence sur les entiers, on obtient le concept suivant pour les relations d'équivalence.

Relation d'équivalence

▹ Une relation d'équivalence sur E est une relation binaire réflexive, symétrique et transitive.
▹ L'ensemble des éléments équivalent à un élément donné est appelé classe d'équivalence de cet élément. Les classes d'équivalence forment une partition de E.

Généralisant la relation $\leq$ sur $\mathbb{R}$, on obtient le concept suivant pour les relations d'ordre.

Relation d'ordre

▹ Une relation d'ordre sur E est une relation binaire réflexive, antisymétrique et transitive.
▹ La borne supérieure d'une partie $A \subset E$ est, quand il existe, le plus petit des majorants de A. On le note $\sup A$.
▹ Le plus grand élément d'une partie $A \subset E$ est, lorsqu'il existe un élément de la partie qui est un majorant. On le note $\max A$. Un plus grand élément de A est nécessairement la borne supérieure de A.

Récurrence

Pour montrer une propriété par récurrence, il faut établir à la fois l'initialisation et le caractère héréditaire. La récurrence forte n'est qu'un cas particulier de récurrence.

EXERCICES

Vrai ou faux ?

		Vrai	Faux
a)	$6\mathbb{Z}\cap 4\mathbb{Z}=12\mathbb{Z}$.	☐	☐
b)	$6\mathbb{Z}\cup 4\mathbb{Z}=2\mathbb{Z}$.	☐	☐
c)	$\forall f:E\to F,\ \forall A\subset B\subset E,\ f(A)\subset f(B)$.	☐	☐
d)	$\forall f:E\to F,\ \forall B\in\mathscr{P}(F),\ \overline{f^{-1}(B)}=f^{-1}(\overline{B})$.	☐	☐
e)	$\forall f:E\to F,\ \forall A,A'\in\mathscr{P}(E),\ f(A)\cap f(A')=f(A\cap A')$.	☐	☐
f)	$\forall f:E\to F,\ \forall B\in\mathscr{P}(F),\ f(f^{-1}(B))\subset B$.	☐	☐
g)	L'application $f:E\to F$ est surjective si, et seulement si, tout élément de F admet au moins un antécédent par f.	☐	☐
h)	Si $f:E\to F$ est injective, alors $f:E\to f(E)$ est bijective.	☐	☐
i)	L'application $f:\left\{\begin{array}{ccc}\mathbb{N}&\to&\mathbb{N}\\ n&\mapsto&2n\end{array}\right.$ est surjective.	☐	☐
j)	L'application $f:\left\{\begin{array}{ccc}\mathbb{N}&\to&\mathbb{N}\\ n&\mapsto&2n\end{array}\right.$ est injective.	☐	☐
k)	Si $f:E\to F$ est surjective, alors $f^{-1}(f(A))=A$ pour tout $A\in\mathscr{P}(E)$.	☐	☐
l)	Si f et g sont surjectives, alors $f\circ g$ est surjective.	☐	☐
m)	Si $f\circ g$ est injective, alors f est injective.	☐	☐
n)	Si $f\circ g$ est surjective, alors f est surjective.	☐	☐
o)	Si une application $f:\mathbb{N}\to\mathbb{N}$ est surjective, alors elle est injective.	☐	☐
p)	Si la borne supérieure d'un ensemble A existe, alors A admet un plus grand élément.	☐	☐

Exercices d'application du chapitre 1

→ Commençons par quelques manipulations ensemblistes.

●○○ **Exercice 1**

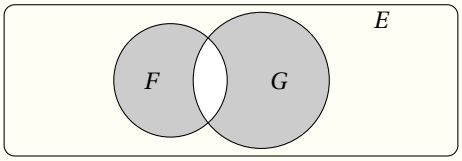
Soit A et B deux parties d'un ensemble E. Résoudre l'équation (d'inconnue $X\in\mathscr{P}(E)$) $A\cap X=B$ puis l'équation $A\cup X=B$.

●○○ **Exercice 2**

Soit E un ensemble. La différence symétrique de deux parties F et G de E est la partie

$$F\Delta G=(\overline{F}\cap G)\cup(F\cap\overline{G}).$$

Sur la figure ci-dessous, la différence symétrique de F et G correspond à la zone grisée.



1. Montrer que, pour toutes parties F et G de E, $F\Delta G = (F\cup G)\cap\overline{F\cap G}$.
2. Vérifier que, pour toutes parties F et G de E, la fonction indicatrice de $F\Delta G$ est égale à
$$\mathbb{1}_F + \mathbb{1}_G - 2\mathbb{1}_F\mathbb{1}_G.$$
Retrouver le résultat de la question précédente.
3. Montrer que Δ est une loi de composition interne sur $\mathscr{P}(E)$ associative, commutative, d'élément neutre $\varnothing$ et telle que toute partie admette un symétrique.
4. Montrer que la loi $\cap$ est distributive sur Δ.

→ **Explorons désormais les propriétés des applications.**

●○○ **Exercice 3**

Soit $f:\mathbb{N}\to\mathbb{Z}$ définie pour tout $n\in\mathbb{N}$ par
$$f(n)=\begin{cases}\frac{n}{2} & \text{si } n=0[2],\\ -\frac{n+1}{2} & \text{sinon.}\end{cases}$$
Montrer que f est bijective et déterminer sa bijection réciproque.

●○○ **Exercice 4**

1. Soit $f:E\to F$ et $g:F\to G$ des applications telles que $g\circ f$ est surjective. Montrer que g est surjective.
2. Soit $f:E\to F$ et $g:F\to G$ des applications telles que $g\circ f$ est injective. Montrer que f est injective.
3. Soit $f:E\to F$, $g:F\to G$ et $h:G\to H$ trois applications telles que $g\circ f$ et $h\circ g$ sont bijectives. Montrer que f, g et h sont bijectives.

●○○ **Exercice 5**

Soit $f:E\to F$, $A\subset E$ et $B\subset F$. Montrer que $f(A\cap f^{-1}(B))=f(A)\cap B$.

●●● **Exercice 6**

Soit $f:E\to F$ une application. Définissons les applications suivantes
$$\Phi:\begin{cases}\mathscr{P}(E) & \to & \mathscr{P}(F)\\ A & \mapsto & f(A)\end{cases}\qquad \Psi:\begin{cases}\mathscr{P}(F) & \to & \mathscr{P}(E)\\ B & \mapsto & f^{-1}(B)\end{cases}$$

1. Montrer que f est injective si, et seulement si, Φ est injective si, et seulement si, Ψ est surjective.
2. Trouver un analogue pour f surjective.

●●● **Exercice 7 Lemmes de factorisation**

1. Soit $f:F\to E$ et $g:G\to E$ deux applications. Montrer qu'il existe une application $h:G\to F$ telle que $g=f\circ h$ si, et seulement si, $g(G)\subset f(F)$.
À quelle condition cette application h est-elle unique ?
2. Soit $f:E\to F$ et $g:E\to G$ deux applications. Montrer qu'il existe une application $h:F\to G$ telle que $g=h\circ f$ si, et seulement si,
$$\forall\, x,y\in E,\qquad f(x)=f(y)\quad\Rightarrow\quad g(x)=g(y).$$
À quelle condition h est-elle unique ?

→ Les exercices suivants permettent de s'approprier les propriétés des relations binaires.

●○○ **Exercice 8**

Déterminer les réels α tels que la relation binaire définie par

$$\forall x, y \in \mathbb{R}, \qquad x \star y = xy + \alpha(x+y)$$

soit associative.

●○○ **Exercice 9**

Soit E un ensemble et A une partie de E. On définit une relation $\mathscr{R}$ sur $\mathscr{P}(E)$ par :

$$X \mathscr{R} Y \Leftrightarrow X \cup A = Y \cup A.$$

1. Montrer que $\mathscr{R}$ est une relation d'équivalence.
2. Décrire la classe d'équivalence de $X \in \mathscr{P}(E)$.

●○○ **Exercice 10**

On définit une relation $\mathscr{R}$ sur $\mathbb{R}$ par $x \mathscr{R} y$ si

$$\cos^2 x + \sin^2 y = 1.$$

Montrer que $\mathscr{R}$ est une relation d'équivalence puis déterminer ses classes d'équivalence.

●○○ **Exercice 11**

On définit une relation binaire $\mathscr{R}$ sur $\mathbb{Z}$ par $x \mathscr{R} y$ si, et seulement si, $x+y$ est pair. Montrer que $\mathscr{R}$ est une relation d'équivalence et préciser ses classes d'équivalence.

●●● **Exercice 12**

On définit une relation $\sim$ sur $\mathbb{R}^{\mathbb{N}}$ par $u \sim v$ si

$$\forall n \in \mathbb{N}, \exists p, q \geq n, \qquad u_p \leq v_n \text{ et } v_q \leq u_n.$$

Montrer que $\sim$ est une relation d'équivalence.

●○○ **Exercice 13**

1. Soit A et B deux parties non vides de $\mathbb{R}$ telles que

$$\forall (a,b) \in A \times B \qquad a \leq b.$$

   Montrer que $\sup A$ et $\inf B$ existent et que $\sup A \leq \inf B$.
2. Exhiber un exemple de parties A et B telles que $\sup A = \inf B$ et

$$\forall (a,b) \in A \times B \qquad a < b.$$

●●○ **Exercice 14**

1. Considérons la relation binaire sur $\mathbb{C}$ définie par

$$z \vdash z' \quad \Leftrightarrow \quad \exists n \in \mathbb{N},\ z' = z^{2^n}.$$

   $\vdash$ est-elle une relation d'ordre ?
2. Reprendre la question précédente en définissant $\vdash$ sur $\mathbb{R}$.

→ Les trois exercices suivants permettent de vérifier la maîtrise du raisonnement par récurrence.

●○○ **Exercice 15**

Soit $x_1, x_2, \ldots, x_n \geq 1$. Montrer que $x_1 + x_2 + \ldots + x_n \leq n - 1 + x_1 x_2 \ldots x_n$.

EXERCICES

●○○ **Exercice 16**

Notons, pour tout $k \in \mathbb{N} \setminus \{0\}$, u_k le plus grand diviseur impair de k.
Montrer que, pour tout $n \in \mathbb{N} \setminus \{0\}$,

$$u_{n+1} + u_{n+2} + \ldots + u_{2n} = n^2.$$

●●○ **Exercice 17**

1. Calculer $s_k = k^2 - (k+1)^2 - (k+2)^2 + (k+3)^2$ pour tout $k \in \mathbb{N}$.
2. Remarquer la valeur de $s_k - s_{k+4}$ pour tout $k \in \mathbb{N}$.
3. En déduire que, pour tout entier n, il existe une infinité d'entiers p et des suites $(\varepsilon_k)_{k \leq p} \in \{\pm 1\}^p$ tels que $n = \varepsilon_1 1^2 + \varepsilon_2 2^2 + \ldots + \varepsilon_p p^2$.

Exercices d'approfondissement du chapitre 1

●○○ **Exercice A**

Soit f et g deux bijections de $\mathbb{Z}$ dans $\mathbb{Z}$. Montrer que l'application $k \mapsto f(k)g(k)$ n'est pas une bijection de $\mathbb{Z}$ dans $\mathbb{Z}$.

●●○ **Exercice B**

Soit $f : E \to E$ une application. Pour tout $n \in \mathbb{N} \setminus \{0\}$, f^n désigne la composée $n^{\text{ième}}$ de f définie par réccurence par $f^0 = \mathrm{Id}_E$ et pour tout $k \in \mathbb{N}$, $f^{k+1} = f \circ f^k$.
Soit $A \subset E$, $A_n = f^n(A)$, et $B = \bigcup_{n \in \mathbb{N}} A_n$. exo

1. Montrer que $f(B) \subset B$.
2. Montrer que B est la plus petite partie de E stable par f et contenant A.

●●● **Exercice C Théorème du point fixe de Knaster-Tarsky**

Soit $f : \mathscr{P}(E) \to \mathscr{P}(E)$ croissante pour l'inclusion. Montrer qu'il existe une partie X de E qui vérifie $f(X) = X$.

●●○ **Exercice D**

1. Soit $(E_i)_{1 \leq i \leq n}$ une famille finie d'ensembles vérifiant
$$\forall i, j \leq n, \exists k \leq n, \qquad E_i \cup E_j \subset E_k.$$
Montrer qu'il existe $k \leq n$ tel que, pour tout $i \leq n$, $E_i \subset E_k$.
2. Le résultat subsiste-t-il pour une famille infinie d'ensembles ?

●●● **Exercice E**

Soit $E = \{x_1, \ldots, x_p\}$. Prouver que l'on peut ordonner les éléments de $\mathscr{P}(E)$ en respectant les quatre règles suivantes :

- on commence par $\varnothing$;
- on termine par un singleton ;
- chaque élément de $\mathscr{P}(E)$ apparaît une et une seule fois ;
- chaque terme de la suite est obtenu par l'ajout ou le retrait d'un seul élément du terme précédent.

●●○ **Exercice F**

Soit E un ensemble et $f : E \to E$. Montrer que f est injective si, et seulement si, pour toutes parties A et B de E, $f(A \cap B) = f(A) \cap f(B)$.

●●○ **Exercice G**

On définit une relation binaire $\prec$ sur $\mathbb{R}^2$ définie par

$$(x, y) \prec (x', y') \qquad \Leftrightarrow \qquad |x' - x| \leq y' - y.$$

1. Montrer que $\prec$ est une relation d'ordre. Est-elle totale ?
2. Montrer que la borne supérieure du disque $\mathscr{D} = \{(x,y) \in \mathbb{R}^2,\ x^2+y^2 \leq 1\}$ est le point $S = (0, \sqrt{2})$.

Problèmes du chapitre 1

●●○ **Problème 1 Lemme de classe monotone**

Dans ce problème, on fixe un ensemble E et on considère des ensembles de parties de E. Plus précisément, on s'intéresse aux notions de la définition suivante.

Définition 1.44.

Soit $\mathscr{A}$ un ensemble de parties de E.
- $\mathscr{A}$ est stable par intersection si pour tous A, $B \in \mathscr{A}$, $A \cap B \in \mathscr{A}$.
- $\mathscr{A}$ est stable par réunion si pour tous A, $B \in \mathscr{A}$, $A \cup B \in \mathscr{A}$.
- $\mathscr{A}$ est une classe monotone si
  - $\mathscr{A}$ contient E,
  - pour tous A, $B \in \mathscr{A}$ tels que $B \subset A$, $A \cap \overline{B} \in \mathscr{A}$,
  - pour toute suite $(A_n)_n \in \mathscr{A}^{\mathbb{N}}$ croissante pour l'inclusion
$$\bigcup_{n\in\mathbb{N}} A_n \in \mathscr{A}.$$
- $\mathscr{A}$ est une tribu si
  - $\mathscr{A}$ est non vide,
  - pour tout $A \in \mathscr{A}$, $\overline{A} \in \mathscr{A}$,
  - pour toute suite $(A_n)_n \in \mathscr{A}^{\mathbb{N}}$
$$\bigcup_{n\in\mathbb{N}} A_n \in \mathscr{A}.$$

Partie A – Exemples et premières propriétés

1. Montrer que les ensembles suivants sont des tribus
   a. $\mathscr{P}(E)$
   b. $\{\varnothing, E\}$
   c. $\{\varnothing, A, \overline{A}, E\}$ avec $A \subset E$ fixé
2. Montrer qu'une tribu est stable par réunion et par intersection.
3. Montrer qu'une tribu est une classe monotone.
4. Soit $\mathscr{A}$ une classe monotone stable par réunion.
   a. Montrer que, pour tous $A_1, \ldots, A_n \in \mathscr{A}$, $A_1 \cup \ldots \cup A_n \in \mathscr{A}$.
   b. Soit $(A_n)_n \in \mathscr{A}^{\mathbb{N}}$ et $(B_n)_n$ la suite de parties de E définie par
$$\forall n \in \mathbb{N}, \qquad B_n = \bigcup_{k=0}^{n} A_k.$$
   Exprimer $\bigcup_{k=0}^{n} B_k$ et $\bigcup_{k\in\mathbb{N}} B_k$ en fonction des termes de la suite $(A_n)_n$.
   c. En déduire qu'une classe monotone stable par réunion est une tribu.
5. En déduire qu'une classe monotone stable par intersection est une tribu.

Partie B – Classes et tribus engendrées

1. Montrer qu'une intersection d'une famille quelconque de classes monotones de E est encore une classe monotone.
2. Énoncer et justifier un résultat équivalent pour les tribus.
3. Soit $\mathscr{A}$ un ensemble de parties de E. Montrer qu'il existe une plus petite (au sens de l'inclusion) classe monotone contenant $\mathscr{A}$ et une plus petite tribu contenant $\mathscr{A}$.

On note ces $\lambda(\mathscr{A})$ et $\sigma(\mathscr{A})$ cette classe monotone et cette tribu et on les appelle classe engendrée par $\mathscr{A}$ et tribu engendrée par $\mathscr{A}$.

Partie C – Lemme de classe monotone

Le but de cette partie est d'établir le résultat suivant :

Théorème 1.44.

Soit $\mathscr{A}$ un ensemble de parties de E stable par intersection. La classe monotone engendrée par $\mathscr{A}$ est la tribu engendrée par $\mathscr{A}$.

On fixe $\mathscr{A}$ un ensemble de parties de E stable par intersection.

1. Pour tout $A \in \lambda(\mathscr{A})$, notons $\mathscr{C}_A = \{B \in \lambda(\mathscr{A}),\ A \cap B \in \lambda(\mathscr{A})\}$.
   a. Montrer que, pour tout $A \in \lambda(\mathscr{A})$, $\mathscr{C}_A$ est une classe monotone.
   b. Montrer que si $A \in \mathscr{A}$, alors $\mathscr{A} \subset \mathscr{C}_A$ puis $\lambda(\mathscr{A}) \subset \mathscr{C}_A$.
   c. Soit $A \in \mathscr{A}$, $B \in \lambda(\mathscr{A})$. Justifier successivement que $B \in \mathscr{C}_A$ puis $A \in \mathscr{C}_B$.
   d. En déduire que, pour tout $B \in \lambda(\mathscr{A})$, $\lambda(\mathscr{A}) \subset \mathscr{C}_B$.
   e. Montrer que $\lambda(\mathscr{A})$ est stable par intersection.
2. Conclure.

●○○ Problème 2 Saturation pour une relation d'équivalence

Étant donnés un ensemble E et une relation d'équivalence $\sim$ sur E, la classe d'équivalence de $x \in E$ est la partie de E, notée $\overline{x}$, constituée des éléments en relation avec x ; en d'autres termes, pour tout $x \in E$,

$$\overline{x} = \{y \in E,\ x \sim y\}.$$

Afin d'éviter les confusions avec cette écriture d'une classe d'équivalence, dans ce problème on notera $A^{\complement}$ le complémentaire d'une partie A.

Partie A – Partition en classes d'équivalence

1. Soit E un ensemble muni d'une relation d'équivalence $\sim$.
   Montrer que deux classes d'équivalence distinctes sont disjointes ; en déduire que les classes d'équivalence forment une partition de E.
2. Dans cette question, on considère E, F deux ensembles et $f : E \to F$. Pour tous $x, y \in E$, on note $x \sim y$ si $f(x) = f(y)$.
   a. Vérifier que $\sim$ est une relation d'équivalence sur E.
   b. Montrer que, pour tout $x \in E$, $\overline{x} = f^{-1}(\{f(x)\})$.
   c. Déterminer les classes d'équivalence sur $\mathbb{Z}$ pour la relation obtenue avec l'application $f : \mathbb{Z} \to [\![0,4]\!]$ qui, a un entier n, associe le reste de la division de n^2 par 5 (que l'on pourra noter abusivement $n^2[5]$).
3. Dans cette question, on considère E un ensemble fini et A une partie de E. Pour tous X, $Y \in \mathscr{P}(E)$, on note $X \sim Y$ si $X \cup A = Y \cup A$.
   a. Montrer que $\sim$ est une relation d'équivalence sur $\mathscr{P}(E)$.
   b. Déterminer les classes d'équivalence ; en préciser le nombre.
   c. Montrer que toutes les classes d'équivalence ont le même cardinal.

Partie B – Saturation d'une partie

Soit E un ensemble muni d'une relation d'équivalence $\sim$. La saturation d'une partie $A \subset E$ est la partie

$$\mathrm{sat}(A) = \bigcup_{x \in A} \overline{x}.$$

1. Montrer que, pour toute partie $A \subset E$, $A \subset \mathrm{sat}(A)$.

2. Comparer, pour toute partie $A \subset E$, $\mathrm{sat}(A)$ et $\mathrm{sat}(\mathrm{sat}(A))$.
3. Soit $A \subset E$ et $x \in E$. Montrer l'équivalence entre les propriétés
   - $x \in \mathrm{sat}(A)$;
   - $\overline{x} \cap \mathrm{sat}(A) \neq \varnothing$;

   En déduire $\mathrm{sat}(\mathrm{sat}(A)^{\complement})$.
4. Montrer que la saturation d'une réunion de parties est la réunion des saturations de ces parties.
5. Comparer la saturation d'une intersection de parties et l'intersection des saturations de ces parties.

Partie C – Parties saturées

Soit E un ensemble muni d'une relation d'équivalence $\sim$. Une partie $A \subset E$ est saturée si elle est égale à sa saturation. On note $\mathscr{S}$ l'ensemble des parties saturées de E.

1. a. Montrer qu'une partie est saturée si, et seulement si, elle est réunion de classes d'équivalence.
   b. Déterminer toutes les parties saturées de $\mathbb{Z}$ pour la relation d'équivalence A.c..
2. Montrer que $\mathscr{S}$ est stable par réunion, intersection et passage au complémentaire, c'est-à-dire que, pour tous $A, B \in \mathscr{S}$
$$A \cup B \in \mathscr{S}, \qquad A \cap B \in \mathscr{S}, \qquad A^{\complement} \in \mathscr{S}.$$
3. a. Soit A une partie saturée et $B \subset E$ tels que $A \cap B = \varnothing$.
   Montrer que $A \cap \mathrm{sat}(B) = \varnothing$.
   b. L'hypothèse A saturée est-elle nécessaire à la question précédente ?
4. Dans cette question, on reprend les notations de la question A.2., à savoir deux ensembles E, F, une application $f : E \to F$ et la relation d'équivalence qui en découle.
   Montrer que $\mathscr{S}$ est en bijection avec $\mathscr{P}(f(E))$.

●●● Problème 3 Théorème de Sprague-Grundy

Un jeu à deux joueurs est un couple (S, V) où S est un ensemble fini non vide d'éléments appelés configurations et V une application de S dans $\mathscr{P}(S)$. Une partie du jeu issue de la configuration x_0 entre les joueurs A et B est constituée d'étapes :

- le joueur A choisit une configuration $x_1 \in V(x_0)$,
- ensuite, si cela est possible, le joueur B choisit une configuration $x_2 \in V(x_1)$,
- ensuite, si cela est possible, le joueur A choisit une configuration $x_3 \in V(x_2)$,
- …

Le joueur perdant est celui qui ne peut choisir une configuration.

Par exemple, dans une partie d'échecs, chaque configuration correspond à une disposition (d'une partie) des pièces sur l'échiquier et l'ensemble $V(x)$ associé à une configuration x est l'ensemble des configurations associées à un mouvement licite d'une pièce depuis la configuration x.

Dans ce problème, on ne considère que des jeux pour lesquels toute partie est finie (donc pas les échecs dans les règles simples).

Voici quelques notations utilisées.

Notation

▹ Pour toute partie $E \subsetneq \mathbb{N}$, $\mathrm{mex}(E)$ (pour minimum exclu) désigne le plus petit entier naturel n'appartenant par à E, c'est-à-dire $\min E^{\complement}$. Par exemple
$$\mathrm{mex}\{2^a 3^b 5^c,\ a, b, c \in \mathbb{N}\} = 0, \qquad \mathrm{mex}[\![0, 7]\!] = 8.$$

▹ Pour tous entiers x et y, l'entier $x \oplus y$ est défini comme la somme des puissances de 2 qui n'apparaissent que dans l'une des écritures binaires de x et y. Par exemple,
$$13 \oplus 20 = (2^3 + 2^2 + 2^0) \oplus (2^4 + 2^2) = (2^4 + 2^3 + 2^0) = 25.$$

On admet que la loi de composition interne $\oplus$ sur $\mathbb{N}$ est associative.

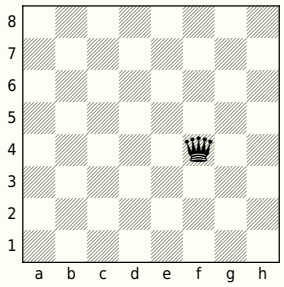
PROBLÈMES

Partie A – Fonction de Grundy

Soit (S, V) un jeu fini. On définit la fonction de Grundy f de ce jeu par $f(x) = 0$ si $V(x) = \varnothing$ et $f(x) = \text{mex}\{f(y),\ y \in V(x)\}$ sinon.

1. Justifier qu'il existe au moins une configuration x telle que $V(x) = \varnothing$.
2. Montrer que l'application f est correctement définie.
   Une configuration de départ x est gagnante s'il existe une stratégie permettant au joueur A de gagner indépendamment des choix du joueur B.
3. Montrer qu'une configuration x est gagnante si, et seulement si, $f(x) > 0$. On précisera la stratégie gagnante pour le joueur démarrant en une configuration x tel que $f(x) > 0$.
4. Dans cette question, on considère le jeu de Wythoff. Sur un échiquier, on pose une seule pièce, une dame ; à tour de rôle, les joueurs déplacent cette dame soit vers le bas, soit vers la gauche, soit en diagonale en bas et à gauche. Le joueur ne pouvant plus bouger la dame a perdu.

   a. Décrire le jeu avec la définition de ce problème.
   b. Déterminer la fonction de Grundy associée à ce jeu sur un échiquier 8×8.
      On pourra disposer les valeurs de cette fonction directement sur un échiquier.
   c. Acceptez-vous de jouer en premier dans la configuration ci-contre ?



Partie B – Jeux composés

La somme des jeux $J_1 = (S_1, V_1)$ et $J_2 = (S_2, V_2)$, de fonctions de Grundy f_1 et f_2, est le jeu, noté $J_1 + J_2$, défini par $(S_1 \times S_2, V)$ avec

$$V : (x, y) \mapsto \{(s, y),\ s \in V_1(x)\} \cup \{(x, s),\ s \in V_2(y)\}.$$

Intuitivement, les deux jeux sont disposés côté à côte et, à chaque étape, le joueur est libre de jouer dans l'un ou dans l'autre. On admet que toute partie du jeu $J_1 + J_2$ est finie lorsque toutes les parties de J_1 et J_2 le sont.

1. Définissons $(\mathscr{V}_n)_n$ une suite de parties de $S_1 \times S_2$ définie par récurrence avec $\mathscr{V}_0 = \varnothing$ et

$$\forall n \in \mathbb{N}, \qquad \mathscr{V}_{n+1} = \{(x, y) \in S_1 \times S_2,\ V(x, y) \subset \mathscr{V}_n\}.$$

   Montrer que $S_1 \times S_2$ est la réunion des parties $\mathscr{V}_n$.
2. On cherche désormais à établir le théorème de Sprague-Grundy.

Théorème 1.45.

Soit f la fonction de Grundy du jeu $J_1 + J_2$. Alors, pour tout $n \in \mathbb{N}$,

$$\forall (x, y) \in \mathscr{V}_n, \qquad f(x, y) = f_1(x) \oplus f_2(y).$$

   a. Établir le résultat pour $n = 1$.
      Dorénavant, on fixe $n \in \mathbb{N}$ et on suppose le résultat établi pour tout élément de $\mathscr{V}_n$.
      Soit $(x, y) \in \mathscr{V}_{n+1}$.
   b. Montrer, en raisonnant par l'absurde et en utilisant la définition de $f(x, y)$ comme mex, que $f(x, y) \leq f_1(x) \oplus f_2(y)$.
   c. Montrer que $f(x, y) = f_1(x) \oplus f_2(y)$.
      On admettra le lemme arithmétique suivant :

Lemme 1.46.

Soit a, b et $c \in \mathbb{N}$ tels que $a < b \oplus c$. Alors $a \oplus b < c$ ou $a \oplus c < b$.

3. On considère $J = (S, V)$ un jeu. À quelle condition sur n et $x \in S$ la position $(x, x, \ldots, x)$ est gagnante dans la somme $J + J + \ldots + J$ de n occurrences de J ?

Partie C – Jeu de Nim

Dans cette question, on considère le jeu de Nim. Une configuration du jeu est un ensemble de k tas de jetons et les joueurs enlèvent à tour de rôle un nombre quelconque de jetons de l'un des tas. Le joueur prenant le dernier jeton a gagné. Dans la configuration initiale, les tas comportent $n_1, n_2, \ldots, n_k$ jetons.

1. Décrire le jeu en précisant S et V.
2. Déterminer la fonction de Grundy associée au jeu de Nim à un seul tas.
3. Déterminer la fonction de Grundy associée au jeu de Nim à k tas.
4. Acceptez-vous de jouer en premier dans une partie de jeu de Nim avec quatre tas à 31, 17, 11 et 5 jetons ?

●●● Problème 4 Confluence

Partie A – Préliminaire

Définition 1.47.

Soit E un ensemble muni d'une relation binaire $\longrightarrow$. ▹ La clôture transitive de $\longrightarrow$ est la relation $\xrightarrow{*}$ définie, pour tous $x, y \in E$, par $x \xrightarrow{*} y$ s'il existe un entier $n \in \mathbb{N}$ et des éléments $x_0, x_1, \ldots, x_n \in E$ tels que $x = x_0$, $y = x_n$ et, pour tout $k \in [\![0, n-1]\!]$, $x_k \longrightarrow x_{k+1}$; *en d'autres termes, $x \xrightarrow{*} y$ si l'on peut passer de x à y par une succession de relations $\longrightarrow$.*

▹ Un élément $x \in E$ est irréductible pour $\longrightarrow$ s'il n'existe pas de y différent de x tels que $x \xrightarrow{*} y$.
▹ La relation $\longrightarrow$ termine s'il n'existe pas de suite (x_n) d'éléments de E telle que $x_n \longrightarrow x_{n+1}$ pour tout n.

1. Rappeler la définition de transitivité pour une relation binaire.
2. Montrer que la relation $\xrightarrow{*}$ est réflexive et transitive. Est-ce une relation d'équivalence ?
3. Donner un exemple de relation qui termine et un exemple de relation qui ne termine pas.

Partie B – Réécriture

Cette partie est consacrée à un exemple très particulier de relations, appelée relation de réécriture.

Considérons E l'ensemble des mots formés avec les lettres a et b.

1. Munissons E de la relation $\longrightarrow$ définie par $m \longrightarrow m'$ si l'on peut obtenir le mot m' en supprimant dans m un bloc de deux lettres consécutives ab.
   a. Déterminer tous les éléments irréductibles de E.
   b. En considérant la longueur des mots, montrer que la relation termine.
   Posons, pour tout mot m, $\delta(m)$ la différence du nombre de b de m moins le nombre de a de m et $\Delta(m)$ la plus grande valeur de δ pour un préfixe du mot m.
   c. Montrer que si $m \longrightarrow m'$, alors $\delta(m) = \delta(m')$ et $\Delta(m) = \Delta(m')$.
   d. En déduire que, pour tout mot m, il existe un unique couple d'entiers (p, q) satisfaisant la propriété $m \xrightarrow{*} b^p a^q$.
2. Munissons E de la relation $\longrightarrow$ définie par $m \longrightarrow m'$ si l'on peut obtenir le mot m' en remplaçant dans m un bloc de deux lettres consécutives ab par un bloc de trois lettres consécutives bba. Par exemple, $abab \longrightarrow abbba$ (en considérant la réécriture du bloc formé des deux dernières lettres) et $abab \longrightarrow bbaab$ (en considérant la réécriture du bloc formé des deux premières lettres).
   a. Déterminer tous les éléments irréductibles de E.
   Posons, pour tout m, $f(m)$ la somme des 3^k où k désigne le nombre de a à gauche d'une occurrence de b.

PROBLÈMES

Par exemple, la quantité associée au mot $baabbab$ vaut $3^0+3^2+3^2+3^3=46$.

b. Montrer que si $m \to m'$, alors $f(m) > f(m')$.

c. En déduire que la relation $\to$ termine.

3. Munissons E de la relation $\to$ définie par $m \to m'$ si l'on peut obtenir le mot m' en remplaçant dans m un bloc de deux lettres consécutives ab par un bloc de quatre lettres consécutives $bbaa$. Montrer à l'aide du mot aab que la relation ne termine pas.

Partie C – Propriétés de confluence

Définition 1.48.

Soit E un ensemble muni d'une relation $\to$.

▹ La relation $\to$ est localement confluente si, pour tout $(x,y,z) \in E^3$ tel que $x \to y$ et $x \to z$, il existe $t \in E$ tel que $y \xrightarrow{*} t$ et $z \xrightarrow{*} t$.

▹ La relation $\to$ est confluente si, pour tout $(x,y,z) \in E^3$ tel que $x \xrightarrow{*} y$ et $x \xrightarrow{*} z$, il existe $t \in E$ tel que $y \xrightarrow{*} t$ et $z \xrightarrow{*} t$.

1. Soit E un ensemble muni d'une relation $\to$ confluente. Montrer que, pour tout $x \in E$, il existe au plus un élément irréductible $y \in E$ tel que $x \xrightarrow{*} y$.
2. Montrer qu'une relation confluente est localement confluente.
3. Définissons, dans cette question, la relation $\to$ sur $[\![0,3]\!]$ par $1 \to 0$, $1 \to 2$, $2 \to 1$ et $2 \to 3$. La relation $\to$ est-elle confluente ? localement confluente ?
4. Définissons, dans cette question, la relation $\to$ sur $\mathbb{N} \cup \{\circ, \bullet\}$ (on ajoute deux nouveaux symboles à l'ensemble des entiers naturels $\mathbb{N}$) par
   - pour tout $k \in \mathbb{N}$, $k \to k+1$,
   - pour tout $k \in \mathbb{N}$, $2k \to \circ$,
   - pour tout $k \in \mathbb{N}$, $2k+1 \to \bullet$.

   On peut représenter la relation par le schéma (partiel) suivant

$$\begin{tikzcd}
 & & \circ & & & & \\
0 \arrow[r] \arrow[urr] & 1 \arrow[r] \arrow[dr] & 2 \arrow[r] \arrow[u] & 3 \arrow[r] \arrow[d] & 4 \arrow[r] \arrow[ul] & 5 \arrow[r] \arrow[dl] & 6 \arrow[r, dotted] \arrow[ull] \\
 & & & \bullet & & &
\end{tikzcd}$$

   Montrer que $\to$ est localement confluente mais pas confluente.

Partie D – Lemme de Newman

Soit E un ensemble muni d'une relation $\to$ localement confluente qui termine.

L'objectif de cette partie est alors de montrer que la relation $\to$ est confluente (lemme de Newman).

1. Montrer qu'il existe au moins un élément irréductible.
2. Considérons $\mathscr{P} : E \to \{\text{Vrai}, \text{Faux}\}$ telle que
   - $\mathscr{P}(x) = \text{Vrai}$ pour tout élément irréductible $x \in E$;
   - pour tout $y \in E$, si $\mathscr{P}(x) = \text{Vrai}$ pour tout $x \in E$ tel que $y \to x$, alors $\mathscr{P}(y) = \text{Vrai}$.

   Montrer que $\mathscr{P}(x) = \text{Vrai}$ pour tout $x \in E$.
3. Montrer que la relation $\to$ est confluente.

CORRIGÉS

Corrigés des Vrai/Faux

a) Vrai car les entiers multiples à la fois de 4 et de 6 sont les multiples de 12.

b) Faux, $2 \notin 6\mathbb{Z} \cup 4\mathbb{Z}$ par exemple.

c) Vrai.

d) Vrai. En effet, $x \notin f^{-1}(B)$ si, et seulement si, $f(x) \notin B$.

e) Faux avec, par exemple, $f : x \in \mathbb{R} \mapsto 1$, $A = \{0\}$, $A' = \{1\}$.

f) Vrai car l'image d'un antécédent d'un élément $y \in B$ est y.

g) Vrai, c'est la définition.

h) Vrai, elle est injective car l'application de départ l'est ; elle est surjective car on a « limité » l'ensemble d'arrivée aux seuls éléments atteints.

i) Faux, un entier impair n'admet pas d'antécédent.

j) Vrai, un entier pair admet comme unique antécédent sa moitié et un entier impair n'admet pas d'antécédent.

k) Faux avec, par exemple, $f : x \in \mathbb{R} \mapsto x^2$ surjective dans $\mathbb{R}_+$ et $A = \{1\}$; alors, $f^{-1}(f(A)) = \{\pm 1\}$.

l) Vrai.

m) Faux, la seule information dont on dispose est que la restriction de f à $g(E)$ est injective.

n) Vrai.

o) Faux en considérant, par exemple, l'application qui associe à un entier n la partie entière de $n/2$.

p) Faux, l'intervalle $[0, 1[$ admet 1 comme borne supérieure mais n'admet pas de plus grand élément.

Corrigés des exercices d'application du chapitre 1

Exercice 1

▹ Pour la première équation, remarquons que pour avoir des solutions, il est nécessaire que $B \subset A$ puisque $A \cap X \subset A$ pour tout $X \subset E$. Dans ce cas, l'intersection d'une solution de X avec A est B et l'on ne dispose d'aucune information sur l'intersection de X avec le complémentaire $\overline{A}$.

Réciproquement, vérifions que les parties $X = B \cup C$ où $C \subset \overline{A}$ sont bien solutions :

$$\cup C) \cap A = (B \cap A) \cup (C \cap A) = B \cap A \cup \varnothing = B.$$

▹ On procède de même et on montre qu'il est nécessaire que $A \subset B$ puis que les solutions sont réunions d'une partie de A et de $B \cap \overline{A}$.

Exercice 2

1. Utilisons successivement la loi de De Morgan puis la distributivité :

$$\begin{aligned}(F\cup G)\cap\overline{F\cap G} &= (F\cup G)\cap(\overline{F}\cup\overline{G})\\ &= (F\cap\overline{F})\cup(F\cap\overline{G})\cup(G\cap\overline{F})\cup(G\cap\overline{G})\\ &= (F\cap\overline{G})\cup(G\cap\overline{F})\\ &= F\Delta G.\end{aligned}$$

2. Appliquons les règles de calcul pour la réunion, l'intersection et le passage au complémentaire.

$$\begin{aligned}\mathbb{1}_{F\Delta G} &= \mathbb{1}_{(\overline{F}\cap G)\cup(F\cap\overline{G})}\\ &= \mathbb{1}_{\overline{F}\cap G}+\mathbb{1}_{F\cap\overline{G}}-\mathbb{1}_{\overline{F}\cap G}\mathbb{1}_{F\cap\overline{G}}\\ &= (1-\mathbb{1}_F)\mathbb{1}_G+(1-\mathbb{1}_G)\mathbb{1}_F-(1-\mathbb{1}_F)\mathbb{1}_G(1-\mathbb{1}_G)\mathbb{1}_F\\ &= \mathbb{1}_F+\mathbb{1}_G-2\mathbb{1}_F\mathbb{1}_G.\end{aligned}$$

On retrouve le résultat de la question précédente en calculant la fonction indicatrice de $(F\cup G)\cap\overline{F\cap G}$ et en montrant que c'est la fonction indicatrice de $F\Delta G$.

$$\begin{aligned}\mathbb{1}_{(F\cup G)\cap\overline{F\cap G}} &= \mathbb{1}_{F\cup G}\mathbb{1}_{\overline{F\cap G}}\\ &= (\mathbb{1}_F+\mathbb{1}_G-\mathbb{1}_F\mathbb{1}_G)(1-\mathbb{1}_F\mathbb{1}_G)\\ &= \mathbb{1}_F+\mathbb{1}_G-2\mathbb{1}_F\mathbb{1}_G.\end{aligned}$$

3. ▹ La commutativité de Δ est évidente d'après la définition de Δ et la commutativité de $\cap$ et $\cup$.
▹ Pour tout $F\subset E$, $F\Delta\varnothing=(F\cup\varnothing)\cap F\cap\overline{\varnothing}=F$. Le calcul de $\varnothing\Delta F$ s'obtient avec la propriété de commutativité.
▹ Pour tout $F\subset E$, $\delta F=(F\cup F)\cap\overline{F\cap F}=\varnothing$. Ainsi, chaque partie est son propre symétrique pour Δ.
▹ Pour toutes parties F, G, $H\subset E$,

$$\begin{aligned}\mathbb{1}_{F\Delta(G\Delta H)} &= \mathbb{1}_F+\mathbb{1}_{G\Delta H}-2\mathbb{1}_F\mathbb{1}_{G\Delta H}\\ &= \mathbb{1}_F+\mathbb{1}_G+\mathbb{1}_H-2\mathbb{1}_G\mathbb{1}_H-2\mathbb{1}_F(\mathbb{1}_G+\mathbb{1}_H-2\mathbb{1}_G\mathbb{1}_H)\\ &= \mathbb{1}_F+\mathbb{1}_G+\mathbb{1}_H-2\mathbb{1}_F\mathbb{1}_G-2\mathbb{1}_F\mathbb{1}_H-2\mathbb{1}_G\mathbb{1}_H+4\mathbb{1}_F\mathbb{1}_G\mathbb{1}_H.\end{aligned}$$

Comme ce résultat est symétrique en F, G et H, on en déduit que

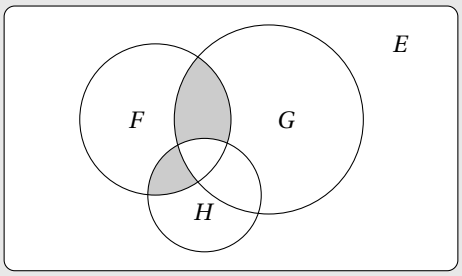
$$\mathbb{1}_{F\Delta(G\Delta H)}=\mathbb{1}_{(F\Delta G)\Delta H}$$

donc $F\Delta(G\Delta H)=(F\Delta G)\Delta H$.

4. Pour toutes parties F, G, $H\subset E$, calculons encore la fonction caractéristique de $F\cap(G\Delta H)$ et de $(F\Delta G)\cap(F\Delta H)$:

$$\begin{aligned}\mathbb{1}_{F\cap(G\Delta H)} &= \mathbb{1}_F(\mathbb{1}_G+\mathbb{1}_H-2\mathbb{1}_G\mathbb{1}_H)\\ \mathbb{1}_{(F\cap G)\Delta(F\cap H)} &= \mathbb{1}_{F\cap G}+\mathbb{1}_{F\cap H}-2\mathbb{1}_{F\cap G}\mathbb{1}_{F\cap H}\\ &= \mathbb{1}_F(\mathbb{1}_G+\mathbb{1}_H-2\mathbb{1}_G\mathbb{1}_H).\end{aligned}$$

Ainsi, $F\cap(G\Delta H)=(F\cap G)\Delta(F\cap H)$.
Sur la figure ci-dessous, la partie $F\cap(G\Delta H)$ correspond à la zone grisée.



Exercice 3

Remarquons d'abord que la fonction est bien définie. De plus,

- pour tout $k \in \mathbb{N}$, $k = f(n)$ est équivalent à $n = 0[2]$ et $k = \frac{n}{2}$ donc $n = 2k$;
- pour tout $k < 0$, $k = f(n)$ est équivalent à $n = 1[2]$ et $k = -\frac{n+1}{2}$ donc $n = -2k-1$.

Dans tous les cas, un élément de $\mathbb{Z}$ admet un unique antécédent par f donc f est une bijection et la bijection réciproque est

$$f^{-1} : \left\{ \begin{array}{ccccc} \mathbb{Z} & \to & & \mathbb{N} & \\ k & \mapsto & & 2k & \text{si } k \geq 0 \\ & & & -2k-1 & \text{sinon.} \end{array} \right.$$

Exercice 4

1. Si $g \circ f$ est surjective, alors pour tout $y \in G$, il existe $x \in E$ tel que $g \circ f(x) = y$ donc y admet au moins un antécédent par g (en l'occurrence $f(x)$) : donc g est surjective.
2. Supposons que $g \circ f$ est injective et montrons que f est injective. Soit x et x' tels que $f(x) = f(x')$ et donc $g(f(x)) = g(f(x'))$; comme $g \circ f$ est injective, on en déduit $x = x'$.
3. D'après les questions précédentes, g est à la fois injective et surjective donc bijective. En composant g^{-1} et $g \circ f$ (respectivement $h \circ g$ et g^{-1}), on obtient que f (respectivement h) est une bijection comme composée de deux bijections.

Exercice 5

Séparons les deux inclusions.

($\subset$) Soit $x \in A \cap f^{-1}(B)$. Alors, $f(x) \in f(A)$ et $f(x) \in B$ donc $f(x) \in f(A) \cap B$.

($\supset$) Soit $y \in f(A) \cap B$. Par définition de $f(A)$, il existe $x \in A$ tel que $y = f(x)$. Comme $y \in B$, $x \in f^{-1}(B)$. Ainsi, $y \in f(A \cap f^{-1}(B))$.

Exercice 6

1. ▹ Si f est injective, alors, pour tout $A \in \mathscr{P}(E)$, $f^{-1}(f(A)) = A$; en effet, chaque élément de A est un antécédent de leur propre image donc $A \subset f^{-1}(f(A))$. Soit $x \in f^{-1}(f(A))$. Alors, x est l'antécédent d'un élément $f(a)$ avec $a \in A$. Comme f est injective, $f(a)$ admet un unique antécédent : a. Ainsi, $x = a \in A$.
En revenant à la question originale, on obtient $\Psi \circ \Phi = \mathrm{Id}_{\mathscr{P}(E)}$ donc Ψ est surjective et Φ est injective.
▹ Supposons Φ injective. Soit $x, y \in E$ tels que $f(x) = f(y)$. Alors, $\Phi(\{x\}) = \Phi(\{y\})$ donc $\{x\} = \{y\}$ par injectivité de Φ donc $x = y$. En conclusion, f est injective.
▹ Supposons Ψ surjective. Soit $x, y \in E$ tels que $f(x) = f(y)$. Soit $B \in \mathscr{P}(F)$ tel que $\{x\} = \Psi(B)$. Par définition, $f(x) = f(y) \in B$ donc $y \in \{x\}$: $x = y$. En conclusion, f est injective.
2. On démontre comme dans les questions précédentes que f est surjective si, et seulement si, Φ est surjective si, et seulement si, Ψ est injective.

Exercice 7

1. ($\Rightarrow$) S'il existe une application $h : G \to F$ telle que $g = f \circ h$, alors $g(G) = f(h(G)) \subset f(F)$.
($\Leftarrow$) Pour tout $x \in G$, $g(x) \in g(G) \subset f(F)$. Par conséquent, $f^{-1}(\{g(x)\}) \neq \varnothing$. En considérant $h(x) \in f^{-1}(\{g(x)\})$, on définit bien une application h convenant. Il y a unicité si, et seulement si, pour tout $x \in G$, $f^{-1}(\{g(x)\})$ est un singleton.
2. ($\Rightarrow$) S'il existe une application $h : F \to G$ telle que $g = h \circ f$, alors, pour tout $x, y \in E$ tel que $f(x) = f(y)$, on a

$$g(x) = h(f(x)) = h(f(y)) = g(y).$$

($\Leftarrow$) Définissons l'application h sur l'ensemble $f(E)$ en associant $g(x)$ à $f(x)$ (ceci définit bien une application car si $f(x)=f(y)$ alors $g(x)=g(y)$). Sur $\overline{f(E)}$, h est définie de manière quelconque. Il y a unicité si, et seulement si, $\overline{f(E)}=\varnothing$, c'est-à-dire si f est surjective.

Exercice 8

Soit $x, y, z \in \mathbb{R}$. Après calculs, on obtient

$$x \star (y \star z) = \alpha^2(y+z)+\alpha(x+\sigma_2)+\sigma_3$$
$$(x \star y) \star z = \alpha^2(y+x)+\alpha(z+\sigma_2)+\sigma_3$$

où $\sigma_2 = xy+yz+zx$, $\sigma_3 = xyz$ désignent les « fonctions symétriques élémentaires » de x, y et z.
Ainsi, la relation est associative si, et seulement si, pour tous x, y, $z \in \mathbb{R}$,

$$\alpha^2(y+z)+\alpha(x+\sigma_2)+\sigma_3 = \alpha^2(y+x)+\alpha(z+\sigma_2)+\sigma_3,$$

soit encore $(\alpha^2-\alpha)(x-z)=0$ c'est-à-dire $\alpha \in \{0,1\}$.

Exercice 9

1. La relation $\mathscr{R}$ est réflexive, symétrique et transitive, donc elle définit une relation d'équivalence.
2. Soit Y une partie de E équivalente à X. $Y \cup A = X \cup A \Rightarrow X \setminus A = Y \setminus A$. En remarquant que $Y = (Y \setminus A) \cup (Y \cap A)$, Y s'écrit sous la forme $Y = (X \setminus A) \cup B$, avec $B \in \mathscr{P}(A)$.
Réciproquement, si $Y = (X \setminus A) \cup B$, avec $B \in \mathscr{P}(A)$, $Y \cup A = (X \setminus A) \cup A = X \cup A$, donc $X \mathscr{R} Y$.

Exercice 10

Remarquons que la relation $x \mathscr{R} y$ se réécrit $\cos^2 x = \cos^2 y$. Avec cette expression, il est immédiat que la relation est réflexive, symétrique et transitive.
La classe d'un réel x est l'ensemble des y tels que $\cos^2 x = \cos^2 y$, c'est-à-dire $\{\pm x + k\pi,\ k \in \mathbb{Z}\}$.

Exercice 11

La relation $\mathscr{R}$ est réflexive car, pour tout $x \in \mathbb{Z}$, $x+x=2x$ est pair, symétrique et transitive. En effet, si $x \mathscr{R} y$ et $y \mathscr{R} z$, il existe des entiers k et l tels que $x+y=2k$ et $y+z=2l$. Donc $x+z=2(k+l-y)$ est pair.
Tous les entiers pairs sont équivalents à 0 et tous les entiers impairs sont équivalents à 1. Comme les entiers pairs et impairs forment une partition, ce sont les deux seuls classes d'équivalence pour $\mathscr{R}$.

Exercice 12

Détaillons les trois points de la définition de relation d'équivalence.
▹ Soit $u \in \mathbb{R}^{\mathbb{N}}$. Alors, pour tout $n \in \mathbb{N}$, on a avec $p=q=n$, $u_p \leq u_n$ et $u_q \leq u_n$. Ainsi, $u \sim u$ et la relation $\sim$ est réflexive.
▹ Soit u, $v \in \mathbb{R}^{\mathbb{N}}$ des suites telles que $u \sim v$. Pour tout $n \in \mathbb{N}$, il existe p, $q \geq n$ tels que $u_p \leq v_n$ et $v_q \leq u_n$, c'est-à-dire tels que $v_q \leq u_n$ et $u_p \leq v_n$. Par conséquent, $v \sim u$ et la relation $\sim$ est symétrique.
▹ Soit u, v, $w \in \mathbb{R}^{\mathbb{N}}$ des suites telles que $u \sim v$ et $v \sim w$ et $n \in \mathbb{N}$.
- Comme $v \sim w$, il existe $p \geq n$ tel que $v_p \leq w_n$. Puis comme $u \sim v$, il existe $p' \geq p$ tel que $u_{p'} \leq v_p$ (en appliquant la propriété au rang p et non au rang n) : ainsi, il existe $p' \geq n$ tel que $u_{p'} \leq w_n$.
- Comme $u \sim v$, il existe $q \geq n$ tel que $v_q \leq u_n$. Puis comme $v \sim w$, il existe $q' \geq q$ tel que $w_{q'} \leq v_q$: ainsi, il existe $q' \geq n$ tel que $w_{q'} \leq u_n$.

Par conséquent, $u \sim w$ et la relation $\sim$ est transitive.

Exercice 13

1. Tout élément $a \in A$ est un minorant de B : la partie B est non vide minorée donc admet une borne inférieure et, comme le plus grand minorant $\inf B$ est supérieur à tous les minorants,

$$\forall a \in A, \qquad a \leq \inf B.$$

Ainsi, $\inf B$ est un majorant de A : la partie A est non vide majorée donc admet une borne inférieure et $\sup A \leq \inf B$ (car le plus petit majorant est inférieur au majorant $\inf B$).
2. Avec $A = \mathbb{R}_-^*$ et $B = \mathbb{R}_+^*$, on a l'inégalité stricte de l'énoncé et pourtant $\sup A = \inf B$.

Exercice 14

1. ▹ La relation est évidemment réflexive puisque $z = z^{2^0}$ pour tout $z \in \mathbb{C}$.
▹ Soit $x, y, z \in \mathbb{C}$ tels que $x \vdash y$ et $y \vdash z$. Par définition, il existe $n, n' \in \mathbb{N}$ tels que $x = y^{2^n}$ et $y = z^{2^{n'}}$ donc $x = (y^{2^n})^{2^{n'}} = z^{2^{n+n'}}$ ce qui entraîne donc $x \vdash z$: la transitivité est établie.
▹ L'antisymétrie est fausse : pour $x = j$ et $y = j^2$, on a $y \vdash x$ (car $y = x^2$) et $x \vdash y$ (car $x = y^2$).
En conclusion, sur $\mathbb{C}$, $\vdash$ n'est pas une relation d'ordre.
2. Il suffit d'étudier l'antisymétrie car le reste est déjà établi. Soit x, $y \in \mathbb{R}$ tels que $x \vdash y$ et $y \vdash x$. Par définition, il existe $n, n' \in \mathbb{N}$ tels que $x = y^{2^n}$ et $y = x^{2^{n'}}$ donc $x = (x^{2^n})^{2^{n'}} = x^{2^{n+n'}}$ ce qui entraîne donc $x = 0$ ou $x^{2^{n+n'}-1} = 1$ donc $x \in \{0, 1\}$. Dans ces deux cas $x = y$. En conclusion, sur $\mathbb{R}$, $\vdash$ est une relation d'ordre.

Exercice 15

Montrons le résultat par récurrence sur $n \geq 1$.
▹ Pour $n = 1$, on a immédiatement que, pour tout $x_1 \geq 1$, $x_1 \leq 1 - 1 + x_1$.
▹ Montrons la propriété pour $n = 2$ (même si ce n'est *a priori* pas nécessaire pour la récurrence). Soit x_1, $x_2 \geq 1$. Alors,

$$1 + x_1 x_2 - x_1 - x_2 = (x_1 - 1)(x_2 - 1) \geq 0.$$

▹ Soit $n \geq 1$ tel que le résultat est établi au rang n. Considérons $x_1, \ldots, x_{n+1} \geq 1$. Alors, par hypothèse de récurrence sur les n premiers termes de la somme,

$$\begin{aligned} x_1 + x_2 + \ldots + x_{n+1} &= x_1 + x_2 + \ldots + x_n + x_{n+1} \\ &\leq n - 1 + x_1 x_2 \ldots x_n + x_{n+1} \\ &\leq n - 1 + 1 + x_1 x_2 \ldots x_n x_{n+1} \end{aligned}$$

en appliquant la propriété au rang 2 avec les réels $x_1 x_2 \ldots x_n$, x_{n+1} (qui sont bien plus grands que 1). Ainsi,

$$x_1 + x_2 + \ldots + x_{n+1} \leq (n+1) - 1 + x_1 x_2 \ldots x_n x_{n+1}$$

ce qui achève la récurrence.

Exercice 16

Montrons le résultat par récurrence sur $n \geq 1$.
▹ Pour $n = 1$, la somme est $u_2 = 1 = 1^2$: la propriété est établie.
▹ Soit $n \geq 1$ tel que $u_{n+1} + u_{n+2} + \ldots + u_{2n} = n^2$. Alors, appliquons l'hypothèse de récurrence et les résultats suivants $u_{2n+2} = u_{n+1}$ et $u_{2n+1} = 2n + 1$ (immédiat avec la définition de la suite u),

$$\begin{aligned} u_{n+2} + u_{n+3} + \ldots + u_{2(n+1)} &= u_{n+1} + u_{n+2} + \ldots + u_{2n} + u_{2n+2} + u_{2n+1} - u_{n+1} \\ &= n^2 + u_{n+1} + 2n + 1 - u_{2n+1} \\ &= (n+1)^2. \end{aligned}$$

Ainsi, le raisonnement par récurrence se termine.

Exercice 17

1. En développant cette expression, on obtient que $s_k = 4$ pour tout $k \in \mathbb{N}$.
2. D'après le résultat de la première question, $s_k - s_{k+4} = 0$ pour tout $k \in \mathbb{N}$.
3. ▹ Montrons tout d'abord par récurrence sur $n \geq 1$ qu'il existe bien une telle décomposition.

• Pour les premiers entiers, on trouve

$$\begin{aligned} 1 &= 1^2 \\ 2 &= -1^2 - 2^2 - 3^2 + 4^2 \\ 3 &= -1^2 + 2^2 \\ 4 &= -1^2 - 2^2 + 3^2 (= s_0) \end{aligned}$$

• Soit $n \geq 4$ tel que tout entier inférieur ou égal à n admet une décomposition de la forme de l'énoncé. Par hypothèse de récurrence, $(n+1)-4$ admet donc une décomposition de la forme

$$n+1-4 = \varepsilon_1 1^2 + \varepsilon_2 2^2 + \ldots + \varepsilon_p p^2.$$

Alors, comme $s_{p+1} = 4$ (d'après la première question),

$$n+1 = \varepsilon_1 1^2 + \varepsilon_2 2^2 + \ldots + \varepsilon_p p^2 + s_{p+1}$$

est bien une décomposition de $n+1$, ce qui achève la récurrence.
▹ Soit n un entier qui admet une décomposition s'arrêtant à l'entier p. Alors, il est facile de construire une infinité de telle décomposition en ajoutant les termes (nuls d'après la deuxième question) $s_{p+1} - s_{p+5}$, $s_{p+1} - s_{p+5} + s_{p+9} - s_{p+13}$, ...

Corrigés des exercices d'approfondissement du chapitre 1

Exercice A

Raisonnons par l'absurde et supposons que l'application produit $h : k \mapsto f(k)g(k)$ réalise une bijection de $\mathbb{Z}$ dans $\mathbb{Z}$.

Par surjectivité de h, il existe des entiers k et l distincts tels que $h(k) = 1$ et $h(l) = -1$, c'est-à-dire $f(k)g(k) = 1$ et $f(l)g(l) = -1$. Comme f et g sont à valeurs entières, on obtient que $f(k) = g(k) \in \{\pm 1\}$ et $f(l) = -g(l) \in \{\pm 1\}$. Comme f est injective, $f(k) \neq f(l)$ soit, avec les conditions ci-dessus, $f(k) = -f(l)$. Par conséquent, $g(k) = g(l)$ ce qui contredit l'injectivité de g.

On a en fait établi un résultat plus précis que celui de l'énoncé : le produit de deux injections de $\mathbb{Z}$ dans $\mathbb{Z}$ ne pouvait être une surjection de $\mathbb{Z}$ dans $\mathbb{Z}$.

Exercice B

1. Soit $x \in B$; montrons que $f(x) \in B$. Par définition, il existe $n \in \mathbb{N}$ tel que $x \in A_n$ et donc $t \in A$ tel que $x = f^n(t)$. Par conséquent, $f(x) = f^{n+1}(t) \in A_{n+1} \subset B$.
2. Soit C une partie de E stable par f et contenant A. Pour tout $n \in \mathbb{N}$, $A_n = f^n(A) \subset C$; par conséquent, $B \subset C$ donc B est effectivement la plus petite partie de E stable par f et contenant A.

Exercice C

Considérons $\mathscr{A} = \{A \in \mathscr{P}(E),\ A \subset f(A)\}$. Remarquons que $\mathscr{A}$ est non vide puisque $\varnothing \in \mathscr{A}$.
Notons X la réunion de toutes les parties de $\mathscr{A}$ et montrons que cette partie convient. Remarquons tout d'abord que

$$f(X) = f\Big(\bigcup_{A\in\mathscr{A}} A\Big) = \bigcup_{A\in\mathscr{A}} f(A) \supset \bigcup_{A\in\mathscr{A}} A = X.$$

Ensuite, comme f est croissante, $f(X) \subset f(f(X))$ donc $f(X) \in \mathscr{A}$ et par conséquent, $f(X) \subset \bigcup_{A\in\mathscr{A}} A = X$.
Cette double-inclusion donne bien $f(X) = X$.

Exercice D

1. Montrons le résultat par récurrence sur $n \in \mathbb{N} \setminus \{0\}$.
- L'initialisation avec un ensemble est évidente.
- Soit $n \in \mathbb{N} \setminus \{0\}$ tel que la propriété soit vraie au rang n. Considérons $n+1$ ensembles $E_1, E_2, \dots, E_{n+1}$ tels que

$$\forall i, j \leq n+1,\ \exists k \leq n+1, \qquad E_i \cup E_j \subset E_k.$$

Notons pour tout $k \leq n$, $E'_k = E_k \cap \overline{E_{n+1}}$. Pour tous $i, j \leq n$,

$$E'_i \cup E'_j \subset (E_i \cup E_j) \cap \overline{E_{n+1}},$$

donc, il existe $k \leq n$ tel que $E'_i \cup E'_j \subset E'_k$. Par hypothèse de récurrence, il existe $k \leq n$ tel que, pour tout $i \leq n$, $E'_i \subset E'_k$. Alors, pour tout $i \leq n+1$,

$$E_i \subset E_k \cup E_{n+1}.$$

Un tel k étant fixé, il existe $l \leq n+1$ tel que $E_k \cup E_{n+1} \subset E_l$ et l'ensemble E_l contient donc tous les autres.
2. Non, par exemple avec $E_i = \{1, 2, \dots, i\}$ pour tout $i \in \mathbb{N} \setminus \{0\}$.

Exercice E

On prouve le résultat par récurrence sur p.
- Pour $p = 1$, on a $E = \{x_1\}$, et l'ordre $\varnothing, \{x_1\}$ convient.
- Supposons que le résultat soit assuré pour un certain $p \geq 1$. On considère alors $E = \{x_1, \dots, x_p, x_{p+1}\}$.

D'après l'hypothèse de récurrence, il existe un ordre convenable pour les parties de $\{x_1, \dots, x_p\}$ qui, sans perte de généralité, se termine par $\{x_p\}$, que l'on note $\varnothing = e_1, e_2, \dots, e_{2^n} = \{x_p\}$. Il est alors facile de vérifier que l'ordre

$$e_1, e_2, \dots, e_{2^n}, e_{2^n} \cup \{x_{p+1}\}, e_{2^n-1} \cup \{x_{p+1}\}, \dots, e_1 \cup \{x_{p+1}\}$$

convient pour $\{x_1, \dots, x_p, x_{p+1}\}$, ce qui prouve le résultat pour $p+1$ et achève la récurrence.

Exercice F

Remarquons tout d'abord que, pour toutes parties A et B de E, $f(A \cap B) \subset f(A) \cap f(B)$ (car $A \cap B \subset A$ et $A \cap B \subset B$). Il s'agit donc de montrer que f est injective si, et seulement si, pour toutes parties A et B de E, $f(A) \cap f(B) \subset f(A \cap B)$.
($\Rightarrow$) Soit A et B des parties de E. Soit $y \in f(A) \cap f(B)$. Il existe $a \in A$, $b \in B$ tels que $y = f(a) = f(b)$. Comme f est injective, $a = b$ donc $a \in A \cap B$ et, par suite, $y \in f(A \cap B)$ ce qui correspond bien à l'inclusion désirée.
($\Leftarrow$) Soit x, $x' \in E$ tels que $f(x) = f(x')$. Appliquons la propriété $A = \{x\}$ et $B = \{x'\}$:

$$f(\{x\} \cap \{x'\}) = \{f(x)\}.$$

Ainsi, $\{x\} \cap \{x'\}$ est non vide donc $x = x'$. La fonction f est donc injective.

Exercice G

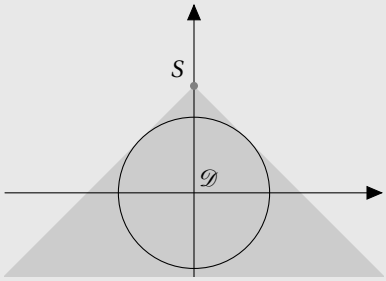
1. ▹ Étudions chacune des propriétés.
- La relation est bien réflexive, puisque, pour tout (x,y), $|x-x|=0\leq y-y$.
- Soit (x,y) et (x',y') tels que $|x-x'|\leq y-y'$ et $|x-x'|\leq y'-y$. Alors, $y-y'\in\mathbb{R}_+\cap\mathbb{R}_-$ donc $y=y'$ et par suite $x=x'$: la relation $\prec$ est bien antisymétrique.
- Soit (x,y), (x',y') et (x'',y'') tels que $|x-x'|\leq y-y'$ et $|x'-x''|\leq y'-y''$. Alors,

$$|x-x''|\leq|x-x'|+|x'-x''|\leq y-y'+y'-y''=y-y''.$$

Ainsi, $(x,y)\prec(x'',y'')$: la relation $\prec$ est transitive.
▹ La relation $\prec$ n'est pas totale : les points $(0,0)$ et $(1,0)$ ne sont pas comparables (de manière plus générale, deux points à la même ordonnée et à abscisses différentes ne peuvent satisfaire l'inégalité définissant la comparaison).
2. ▹ Commençons par montrer que S est un majorant de $\mathscr{D}$. Pour tout $(x,y)\in\mathscr{D}$, l'inégalité de Cauchy-Schwarz (le produit scalaire est inférieur au produit des normes) donne

$$|x|+y=|x|\times 1+y\times 1\leq\sqrt{x^2+y^2}\sqrt{1^2+1^2}\leq\sqrt{2}.$$

Par conséquent, $|x|\leq\sqrt{2}-y$ donc $(x,y)\prec(0,\sqrt{2})$.
Dans le schéma ci-dessous, le quart de plan grisé correspond aux points plus petits que S.



▹ Soit (x,y) un majorant de $\mathscr{D}$. Alors, en particulier, $(\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}})\prec(x,y)$ et $(-\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}})\prec(x,y)$, c'est-à-dire

$$\left|x-\tfrac{1}{\sqrt{2}}\right|\leq y-\tfrac{1}{\sqrt{2}}$$
$$\left|x+\tfrac{1}{\sqrt{2}}\right|\leq y-\tfrac{1}{\sqrt{2}}$$

Alors,

$$2|x|\leq\left|x-\tfrac{1}{\sqrt{2}}\right|+\left|x+\tfrac{1}{\sqrt{2}}\right|\leq 2y-\sqrt{2}.$$

Cette inégalité indique $(0,\sqrt{2})\prec(x,y)$ donc $(0,\sqrt{2})$ est bien le plus petit majorant de $\mathscr{D}$.

Corrigés des problèmes du chapitre 1

Problème 1

Partie A – Exemples et premières propriétés

1. a. L'ensemble $\mathscr{P}(E)$ contient $\varnothing$ et est clairement stable par passage au complémentaire et réunion (puisqu'il contient toutes les parties de E) : c'est une tribu.

b. L'ensemble $\{\varnothing, E\}$ contient $\varnothing$ et les complémentaires de ces éléments sont E et $\varnothing$: ils appartiennent à cet ensemble. On vérifie de même que toutes les réunions (au nombre de 2) sont également dedans : c'est une tribu.
c. Soit $A \subset E$. L'ensemble $\{\varnothing, A, \overline{A}, E\}$ contient $\varnothing$, les 4 complémentaires et les 4 réunions possibles appartiennent bien à cet ensemble : c'est une tribu.
2. Soit $\mathscr{A}$ une tribu, A, $B \in \mathscr{A}$.
▹ Considérons la suite (A_n) d'éléments de $\mathscr{A}$ définie par $A_0 = A$ et $A_n = B$ pour $n \geq 1$. Alors (d'après le troisième point de la définition)

$$A \cup B = \bigcup_{n \in \mathbb{N}} A_n \in \mathscr{A}.$$

▹ Comme $\overline{A}$ et $\overline{B}$ appartiennent à $\mathscr{A}$ (deuxième point de la définition), le résultat précédent indique que $\overline{A} \cup \overline{B} \in \mathscr{A}$. En utilisant une nouvelle fois la stabilité par passage au complémentaire et la loi de De Morgan, on obtient que $A \cap B \in \mathscr{A}$.
En conclusion, une tribu est stable par réunion et par intersection.
3. Soit $\mathscr{A}$ une tribu.
▹ Soit $A \in \mathscr{A}$ (qui existe d'après le premier point). Alors $\overline{A} \in \mathscr{A}$ (stabilité par passage au complémentaire) donc $E = A \cup \overline{A} \in \mathscr{A}$ (stabilité par réunion). Le premier point de la définition de classe monotone est établi.
▹ Soit A, $B \in \mathscr{A}$ tels que $B \subset A$. Alors $\overline{B} \in \mathscr{A}$ donc $A \cap \overline{B} \in \mathscr{A}$ d'après la question précédente : le deuxième point de la définition de classe monotone est établi.
▹ Le troisième point découle directement du point correspondant de la définition de tribu (on se restreint aux suites croissantes).
Ainsi, une tribu est une classe monotone.
4. a. Il s'agit d'une simple récurrence.
b. ▹ La suite (B_n) est, par définition, croissante donc $\bigcup_{k=0}^{n} B_k = B_n$ pour tout entier n.
▹ Remarquons que tout $x \in \bigcup_{k \in \mathbb{N}} B_k$ appartient à un ensemble B_n donc à

$$\bigcup_{k=0}^{n} A_k \subset \bigcup_{k \in \mathbb{N}} A_k.$$

Réciproquement, tout élément $x \in \bigcup_{k \in \mathbb{N}} A_k$ appartient à un ensemble

$$A_n \subset B_n \subset \bigcup_{k \in \mathbb{N}} B_k.$$

Par double inclusion,

$$\bigcup_{k \in \mathbb{N}} B_k = \bigcup_{k \in \mathbb{N}} A_k.$$

c. Montrons que la classe monotone $\mathscr{A}$ est une tribu.
▹ Le caractère non vide est établi d'après le premier point de la définition.
▹ Si A est un élément de $\mathscr{A}$, alors d'après le second point $\overline{A} = E \cap \overline{A}$ appartient à $\mathscr{A}$.
▹ Soit (A_n) une suite de parties de $\mathscr{A}$. Avec les notations de la question précédente,

$$\bigcup_{k \in \mathbb{N}} A_k = \bigcup_{k \in \mathbb{N}} B_k \in \mathscr{A}$$

d'après le troisième point de la définition de classe monotone.
5. Une classe monotone est toujours stable par passage au complémentaire. Par conséquent, avec la loi de De Morgan, une classe monotone stable par intersection est stable par réunion donc une tribu d'après la question précédente.

Partie B – Classes et tribus engendrées

1. Soit $(\mathscr{A}_i)_{i\in I}$ une famille de classes monotones de E.
▹ Comme chaque $\mathscr{A}_i$ contient E, l'intersection contient E.
▹ Soit A, B des éléments de l'intersection des classes tels que $B \subset A$. Pour tout i, $A \cap \overline{B} \in \mathscr{A}_i$ donc $A \cap \overline{B}$ appartient à l'intersection des classes.
▹ Soit $(A_n)_n$ une suite croissante de l'intersection. Alors, pour tout i, $(A_n)_n$ est une suite croissante de $\mathscr{A}_i$ donc sa réunion appartient à l'intersection des classes.
En conclusion, $\cup A_i$ est une classe monotone.
2. Une intersection d'une famille quelconque de tribus de E est encore une tribu (car la définition est une succession de propriétés d'appartenance).
3. L'intersection de toutes les classes monotones contenant $\mathscr{A}$ est encore une classe monotone contenant $\mathscr{A}$ (d'après la question précédente) ; or elle est contenue dans toutes les classes monotones contenant $\mathscr{A}$ donc c'est la plus petite.
Un raisonnement analogue permet de définir la tribu engendrée comme intersection de toutes les tribus contenant $\mathscr{A}$.

Partie C – Lemme de classe monotone

1. a. ▹ Il est clair que $E \in \lambda(\mathscr{A})$ et $A \cap E \in \lambda(\mathscr{A})$ donc $E \in \mathscr{C}_A$.
▹ Soit B_1, $B_2 \in \mathscr{C}_A$ tels que $B_2 \subset B_1$. Alors,

$$(B_1 \cap \overline{B_2}) \cap A = (B_1 \cap A) \cap \overline{(B_2 \cap A)} \in \lambda(\mathscr{A})$$

car $B_1 \cap A$ et $B_2 \cap A$ sont des éléments de $\lambda(\mathscr{A})$ qui vérifient $B_2 \cap A \subset B_1 \cap A$.
▹ Soit $(A_n)_n$ une suite croissante d'éléments de $\mathscr{C}_A$. Alors, pour tout n, $A_n \cap A \in \lambda(\mathscr{A})$. La suite $(A_n \cap A)_n$ est croissante donc

$$\Big(\bigcup_n A_n\Big) \cap A = \bigcup_n (A_n \cap A) \in \lambda(\mathscr{A}).$$

En conclusion, $\mathscr{C}_A$ est une classe monotone.
b. Soit $A \in \mathscr{A}$. Pour tout $B \in \mathscr{A}$, $A \cap B \in \mathscr{A} \subset \lambda(\mathscr{A})$ donc $B \in \mathscr{C}_A$. Ainsi, $\mathscr{C}_A$ est une classe monotone contenant $\mathscr{A}$ donc elle contient la plus petite classe monotone contenant $\mathscr{A}$: $\lambda(\mathscr{A}) \subset \mathscr{C}_A$.
c. Soit $A \in \mathscr{A}$, $B \in \lambda(\mathscr{A})$. Alors, $B \in \mathscr{C}_A$ d'après la question précédente. Ceci indique que $A \cap B \in \lambda(\mathscr{A})$ donc $A \in \mathscr{C}_B$.
d. Soit $B \in \lambda(\mathscr{A})$. La classe monotone $\mathscr{C}_B$ contient $\mathscr{A}$ donc contient la plus petite classe monotone contenant $\mathscr{A}$: $\lambda(\mathscr{A}) \subset \mathscr{C}_B$.
e. Soit A, $B \in \lambda(\mathscr{A})$. Comme $A \in \mathscr{C}_B$, $A \cap B \in \lambda(\mathscr{A})$. Ainsi, $\lambda(A)$ est stable par intersection.
2. On a établi le théorème 1.44.

Problème 2

Partie A – Partition en classes d'équivalence

1. • Raisonnons par l'absurde. Soit $x, y \in E$ tels que $\overline{x} \neq \overline{y}$ et $\overline{x} \cap \overline{y} \neq \varnothing$. Considérons $z \in \overline{x} \cap \overline{y}$. Par définition, $x \sim z$ et $y \sim z$ donc, par transitivité, $x \sim y$. Par conséquent, $\overline{x} = \overline{y}$: contradiction.
• Considérons les classes d'équivalence distinctes. Par définition, elles sont non vides et de réunion tout l'ensemble ; or, nous venons d'établir qu'elles étaient deux à deux disjointes, elles forment donc une partition de E.
2. a. • Pour tout $x \in E$, $f(x) = f(x)$ donc $x \sim x$: la relation $\sim$ est réflexive.
• Pour tous $x, y \in E$, $x \sim y$ si, et seulement si, $f(x) = f(y)$, si, et seulement si, $f(y) = f(x)$ c'est-à-dire $y \sim x$: la relation $\sim$ est symétrique.
• Pour tous $x, y, z \in E$ tels que $x \sim y$ et $y \sim z$, on a $f(x) = f(y)$ et $f(y) = f(z)$ donc $f(x) = f(z)$, c'est-à-dire $x \sim z$: la relation $\sim$ est transitive.

En conclusion, $\sim$ est une relation d'équivalence.
b. Soit $x \in E$. Alors

$$\begin{aligned} y \in \overline{x} &\Leftrightarrow f(y) = f(x) \\ &\Leftrightarrow y \in f^{-1}(\{f(x)\}). \end{aligned}$$

En conclusion, $\overline{x} = f^{-1}(\{f(x)\})$.
c. Remarquons tout d'abord que pour tout $n, k \in \mathbb{Z}$, $f(n+5k) = f(n)$, autrement dit $n+5k \in \overline{n}$ pour tout $k \in \mathbb{Z}$. Il suffit donc d'étudier les classes d'équivalence des entiers de $[\![0,4]\!]$. Or, $f(0) = 0[5]$, $f(1) = f(4) = 1[5]$ et $f(2) = f(3) = 4[5]$: il y a donc trois classes d'équivalence :

- E_0, l'ensemble des entiers congrus à 0 modulo 5 ;
- E_1, l'ensemble des entiers congrus à 1 ou 4 modulo 5 ;
- E_2, l'ensemble des entiers congrus à 2 ou 3 modulo 5.

3. a. On peut montrer les trois propriétés mais il est plus facile de se ramener au cas précédent avec l'application

$$f : \left\{ \begin{array}{ccc} \mathscr{P}(E) & \mapsto & \mathscr{P}(E) \\ X & \to & X \cup A \end{array} \right.$$

b. Remarquons que $X \sim Y$ si, et seulement si, $X \cup A = Y \cup A$ donc si, et seulement si, $X \cap A^{\complement} = Y \cap A^{\complement}$ (en partitionnant $X \cup A$ en A et $X \cap A^{\complement}$) : ainsi, deux parties sont dans la même classe d'équivalence si leurs intersections avec $A^{\complement}$ sont égales. Les classes d'équivalence sont donc de la forme

$$\{B \cup C,\ B \subset A\}$$

avec C une partie de $A^{\complement}$. Il y a donc autant de classes d'équivalence que de parties de $A^{\complement}$, soit $2^{\operatorname{card} A^{\complement}}$ classes d'équivalence.
c. D'après la description de la question précédente, toutes les classes d'équivalence ont pour cardinal $2^{\operatorname{card} A}$.

Partie B – Saturation d'une partie

1. Pour tout $x \in E$, $x \in \overline{x}$ (c'est la réflexivité de la relation d'équivalence $\sim$) c'est-à-dire $\{x\} \subset \overline{x}$. Ainsi, pour tout $A \subset E$,

$$A = \bigcup_{x \in A} \{x\} \subset \bigcup_{x \in A} \overline{x} = \operatorname{sat}(A).$$

2. Soit $A \subset E$. D'après la question précédente, $\operatorname{sat}(A) \subset \operatorname{sat}(\operatorname{sat}(A))$. Montrons qu'il y a en fait égalité. Soit $y \in \operatorname{sat}(\operatorname{sat}(A))$; par définition, il existe $z \in \operatorname{sat}(A)$ tel que $y \in \overline{z}$, c'est-à-dire $y \sim z$; or, comme $z \in \operatorname{sat}(A)$, il existe $x \in A$ tel que $z \in \overline{x}$ c'est-à-dire $z \sim x$. Par transitivité de $\sim$, $z \sim x$ donc $z \in \overline{x} \subset \operatorname{sat}(A)$. En conclusion, $\operatorname{sat}(\operatorname{sat}(A)) = \operatorname{sat}(A)$.
3. Soit $A \subset E$ et $x \in E$.

$$\begin{aligned} x \in \operatorname{sat}(A) &\Leftrightarrow \exists y \in A, \quad x \in \overline{y} \\ &\Leftrightarrow \exists y \in A, \quad x \sim y \\ &\Leftrightarrow \exists y \in A, \quad y \sim x \\ &\Leftrightarrow \exists y \in A, \quad y \in \overline{x} \\ &\Leftrightarrow \overline{x} \cap \operatorname{sat}(A) \neq \varnothing. \end{aligned}$$

Avec cette caractérisation, calculons $\operatorname{sat}(\operatorname{sat}(A)^{\complement})$:

$$\begin{aligned} x \in \operatorname{sat}(\operatorname{sat}(A)^{\complement}) &\Leftrightarrow \exists y \in \operatorname{sat}(A)^{\complement}, \quad x \sim y \\ &\Leftrightarrow \exists y \in \operatorname{sat}(A)^{\complement}, \quad \overline{x} = \overline{y} \\ &\Leftrightarrow \exists y \notin \operatorname{sat}(A), \quad \overline{x} = \overline{y} \\ &\Leftrightarrow \exists y \in E, \quad \overline{y} \cap \operatorname{sat}(A) = \varnothing, \quad \overline{x} = \overline{y} \\ &\Leftrightarrow \overline{x} \cap \operatorname{sat}(A) = \varnothing \\ &\Leftrightarrow x \notin \operatorname{sat}(A). \end{aligned}$$

Ainsi, $\operatorname{sat}(\operatorname{sat}(A)^{\complement}) = \operatorname{sat}(A)^{\complement}$.

4. Soit $(A_i)_{i\in I}$ une famille de parties de E.

$$\operatorname{sat}\left(\bigcup_{i\in I} A_i\right) = \bigcup_{x\in \bigcup_{i\in I} A_i} \overline{x} = \bigcup_{i\in I} \bigcup_{x\in A_i} \overline{x} = \bigcup_{i\in I} \operatorname{sat}(A_i).$$

Ainsi, $\operatorname{sat}\left(\bigcup_{i\in I} A_i\right) = \bigcup_{i\in I} \operatorname{sat}(A_i)$.

5. Soit $(A_i)_{i\in I}$ une famille de parties de E.

$$\operatorname{sat}\left(\bigcap_{i\in I} A_i\right) = \bigcup_{x\in \bigcap_{i\in I} A_i} \overline{x} \subset \bigcup_{x\in A_i} \overline{x} = \operatorname{sat}(A_i).$$

Ainsi, $\operatorname{sat}\left(\bigcap_{i\in I} A_i\right) \subset \bigcap_{i\in I} \operatorname{sat}(A_i)$.

L'autre inclusion est en général fausse comme on peut le voir avec $A_1 = \{x\}$ et $A_2 = \{y\}$ pour x, y distincts tels que $x \sim y$:

$$\operatorname{sat}(A_1 \cap A_2) = \operatorname{sat}(\varnothing) = \varnothing$$
$$\operatorname{sat}(A_1) \cap \operatorname{sat}(A_2) = \overline{x} \cap \overline{y} = \overline{x}.$$

Partie C – Parties saturées

1. a. Si une partie est saturée, alors elle est égale à sa saturation qui est par définition une réunion de classes d'équivalence.

Réciproquement, $A = \bigcup_{i\in I} \overline{x_i}$, alors, pour tout $x \in A$, il existe $i \in I$, tel que $x \in \overline{x_i}$ donc $\overline{x} \subset \overline{x_i}$. Par conséquent,

$$\operatorname{sat}(A) = \bigcup_{x\in A} \overline{x} \subset \bigcup_{i\in I} \overline{x_i} = A.$$

Comme on a toujours $A \subset \operatorname{sat}(A)$, on obtient $A = \operatorname{sat}(A)$, c'est-à-dire que A est saturée.

b. Pour cette relation d'équivalence, il y avait 3 classes d'équivalence E_0, E_1 et E_2. Par conséquent, il y a, d'après la question précédente, $2^3 = 8$ parties saturées

$$\varnothing,\ E_0,\ E_1,\ E_2,\ E_0 \cup E_1,\ E_0 \cup E_2,\ E_1 \cup E_2,\ E_0 \cup E_1 \cup E_2 = \mathbb{Z}.$$

2. Ces trois propriétés sont évidentes en utilisant la caractérisation comme réunion de classes d'équivalence (par exemple, l'intersection de classe d'équivalence est soit vide soit une réunion de classes d'équivalence).

3. a. Soit A une partie saturée et $B \subset E$ tels que $A \cap B = \varnothing$. Supposons, par l'absurde, qu'il existe $x \in A \cap \operatorname{sat}(B)$. Par définition de la saturation de B, il existe $y \in B$ tel que $x \sim y$ donc $y \in \overline{x} \subset A$: ainsi, $y \in A \cap B = \varnothing$: contradiction. En conclusion, $A \cap \operatorname{sat}(B) = \varnothing$.

b. L'hypothèse A saturée est nécessaire pour avoir le résultat précédent comme on peut le voir avec l'exemple suivant $A = \{x\}$ et $B = \{y\}$ pour x, y distincts tels que $x \sim y$:

$$A \cap B = \varnothing, \qquad A \cap \operatorname{sat}(B) = \{x\}.$$

4. Soit $B \subset f(E)$; notons $A = f^{-1}(B)$ (donc $B = f(A)$ car $B \subset f(E)$) et montrons que A est saturée :

$$\begin{aligned}\operatorname{sat}(A) &= \bigcup_{x\in A} \overline{x} = \bigcup_{x\in A} f^{-1}(\{f(x)\}) \\ &= \bigcup_{y\in f(A)} f^{-1}(\{y\}) = \bigcup_{y\in B} f^{-1}(\{y\}) = f^{-1}(B) = A.\end{aligned}$$

Ainsi, on peut définir les applications réciproques l'une de l'autre (car pour tout $B \subset f(E)$, $B = f(f^{-1}(B))$ et pour tout $A \in \mathscr{S}$, $A = f^{-1}(f(A))$) :

$$\left\{\begin{array}{ccc} \mathscr{S} & \to & \mathscr{P}(f(E)) \\ A & \mapsto & f(A) \end{array}\right. \qquad \left\{\begin{array}{ccc} \mathscr{P}(f(E)) & \to & \mathscr{S} \\ B & \mapsto & f^{-1}(B) \end{array}\right. .$$

En conclusion, $\mathscr{S}$ est en bijection avec $\mathscr{P}(f(E))$.

Problème 3

Partie A – Fonction de Grundy

1. Raisonnons par l'absurde et supposons qu'il n'existe pas de configuration x telle que $V(x)=\varnothing$.
▹ Soit x_0 une configuration quelconque.
▹ Soit $n \in \mathbb{N}$ tel qu'il existe une suite $x_0, \dots, x_n$ de configurations distinctes vérifiant

$$\forall k \in \{1, \dots, n\}, \qquad x_k \in V(x_{k-1}).$$

Comme $V(x_n) \neq \varnothing$, il existe $x_{n+1} \in S$ tel que $x_{n+1} \in V(x_n)$. Mais, comme le jeu est fini donc sans cycle, $x_{n+1} \notin \{x_0, \dots, x_n\}$.
Par récurrence, on construit une suite infinie de configurations distinctes : contradiction avec le caractère fini de S.
2. Reprenons le raisonnement par l'absurde de la première question en supposant qu'il existe $x_0 \in S$ tel que la valeur $f(x_0)$ n'est pas définie. Alors, $V(x_0) \neq \varnothing$ (sinon, la valeur $f(x_0)=0$ est correctement définie). Il existe donc $x_1 \in S$ tel que $x_1 \in V(x_0)$ et $f(x_1)$ n'est pas définie. On construit ainsi une suite infinie de configurations distinctes : contradiction, l'application f est correctement définie.
3. ▹ Remarquons que, par définition,
a. une configuration x vérifie $f(x)>0$ si, et seulement si, il existe une configuration y telle que $y \in V(x)$ et $f(y)=0$,
b. une configuration x vérifie $f(x)=0$ si, et seulement si, $x \in F$ ou, pour toute configuration y telle que $y \in V(x)$, $f(y)>0$.
▹ Le joueur ayant à jouer depuis une configuration x telle que $f(x)>0$ dispose d'une stratégie gagnante : se déplacer en une configuration $y \in V(x)$ telle que $f(y)=0$ (qui existe d'après le premier point ci-dessus). L'autre joueur, s'il n'a pas déjà perdu, le ramènera nécessairement en une configuration z telle que $f(z)>0$ (second point ci-dessus). Et l'on recommence... Seul le second joueur peut se retrouver en un sommet de $V^{-1}(\{\varnothing\})$, donc la stratégie est gagnante pour le premier joueur.
4. a. Les différentes configurations sont les positions de la dame. En notant ces positions par leurs coordonnées (entières), on obtient $S=[\![0,7]\!]^2$ et, pour tout $(x,y) \in S$, $V((x,y))$ est la réunion des ensembles

$$\{(x',y),\ x' \in [\![0,x]\!]\}, \qquad \{(x,y'),\ y \in [\![0,y]\!]\},$$
$$\{(x-k,y-k),\ k \in [\![1,\min(x,y)]\!]\}.$$

b. Un simple calcul donne la fonction dont les valeurs sont sur l'échiquier suivant :

7	8	6	9	0	1	4	5
6	7	8	1	9	10	3	4
5	3	4	0	6	8	10	1
4	5	3	2	7	6	9	0
3	4	5	6	2	0	1	9
2	0	1	5	3	4	8	6
1	2	0	4	5	3	7	8
0	1	2	3	4	5	6	7

c. Je refuse car la position étant évaluée à 0 par f, mon advsersaire dispose d'une stratégie gagnante.

Partie B – Jeux composés

1. Raisonnons par l'absurde comme dans la partie précédente. S'il existe une configuration qui n'est pas dans la réunion des parties $\mathscr{V}_n$, alors elle admet une configuration voisine qui n'est pas dans la réunion des parties $\mathscr{V}_n$. Par récurrence, on construit ainsi une suite de configurations distinctes donc une partie infinie : contradiction.

2. a. Soit $(x,y)\in\mathscr{V}_1$. Par définition de $\mathscr{V}_1$, $V(x,y)\subset\mathscr{V}_0=\varnothing$ donc $f(x,y)=0$, mais aussi

$$\{(s,y),\ s\in V_1(x)\}\cup\{(x,s),\ s\in V_2(y)\}=\varnothing$$

ce qui entraîne $V_1(x)=\varnothing$ et $V_2(y)=\varnothing$ donc $f_1(x)=0$ et $f_2(y)=0$. Par conséquent,

$$f(x,y)=0=0\oplus 0=f_1(x)\oplus f_2(y).$$

b. Supposons par l'absurde $f(x,y)>f_1(x)\oplus f_2(y)$. Comme $f(x,y)$ est le minimum exclu des valeurs de f sur $V(x,y)$, il existe $z\in V(x,y)$ tel que $f(z)=f_1(x)\oplus f_2(y)$.
Supposons sans perte de généralité que $z\in V_1(x)\times\{y\}$, c'est-à-dire qu'il existe $x'\in V_1(x)$ tel que $z=(x',y)$.
Comme $z\in\mathscr{V}_n$,

$$f_1(x)\oplus f_2(y)=f(z)=f_1(x')\oplus f_2(y)$$

ce qui entraîne $f_1(x)=f_1(x')$ et contredit $x'\in V_1(x)$.
Ainsi, $f(x,y)\leq f_1(x)\oplus f_2(y)$
c. Supposons par l'absurde $f(x,y)<f_1(x)\oplus f_2(y)$ et utilisons le lemme admis : $f(x,y)\oplus f_1(x)<f_2(y)$ ou $f(x,y)\oplus f_2(y)<f_1(x)$. Supposons sans perte de généralité que $f(x,y)\oplus f_1(x)<f_2(y)$.
Comme $f_2(y)$ est le minimum exclu des valeurs de f_2 sur $V_2(y)$, il existe $y'\in V_2(y)$ tel que

$$f(x,y)\oplus f_1(x)=f_2(y').$$

Par associativité et commutativité de $\oplus$, $f(x,y)=f_1(x)\oplus f_2(y')$. Or $(x,y')\in V(x,y)\subset\mathscr{V}_n$ donc $f(x,y)=f(x,y')$ (par hypothèse de récurrence) ce qui contredit la définition de $f(x,y)$.
En conclusion, $f(x,y)=f_1(x)\oplus f_2(y)$.
3. D'après le théorème de Sprague-Grundy, la valeur en $(x,x,\dots,x)$ de la fonction de Grundy du jeu itéré n fois est 0 si n est pair et $f(x)$ si n est impair. La position $(x,x,\dots,x)$ est donc gagnante si, et seulement si, n est impair et x est gagnant dans le jeu J.

Partie C – Jeu de Nim

1. Une configuration du jeu est un k-uplet $(x_1,\dots,x_k)$ d'entiers naturels tels que, pour tout i, $x_i\leq n_i$. Les configurations de $V((x_1,\dots,x_k))$sont les configurations $(y_1,\dots,y_k)$ telles qu'il existe un indice i vérifiant $y_i<x_i$ et, pour tout $j\neq i$, $y_j=x_j$.
2. Dans le cas d'un jeu de Nim avec un seul tas, la seule configuration sans voisin est 0 et une configuration quelconque k pointe vers toutes les configurations de $\{0,\dots,k-1\}$. Un récurrence immédiate permet de montrer que la fonction de Grundy est alors $f:k\mapsto k$.
3. D'après le théorème de Sprague-Grundy, la fonction de Grundy associée au jeu de Nim à k tas est

$$f:x_1,\dots,x_k\mapsto x_1\oplus\dots\oplus x_k.$$

4. Je refuse car la valeur de la fonction de Grundy pour cette configuration est

$$31\oplus 17\oplus 11\oplus 5=(2^4+2^3+2^2+2^1+2^0)\oplus(2^4+2^0)\oplus(2^3+2^1+2^0)\oplus(2^2+2^0)=0.$$

Problème 4

Partie A – Préliminaire

1. Une relation $\mathscr{R}$ sur E est transitive si, pour tous x, y et $z\in E$ tels que $x\mathscr{R}y$ et $y\mathscr{R}z$, on a $x\mathscr{R}z$.
2. ▹ La relation $\xrightarrow{*}$ est réflexive (il suffit de prendre $n=0$ dans la définition).
▹ Montrons que la relation $\xrightarrow{*}$ est transitive. Soit x, y et $z\in E$ tels que $x\xrightarrow{*}y$ et $y\xrightarrow{*}z$. Par définition, il existe des entiers n et m et des éléments $x_0,x_1,\dots,x_n\in E$ tels que $x=x_0$, $y=x_n$ et, pour tout $k\in[\![0,n-1]\!]$, $x_k\to x_{k+1}$ d'une part et des éléments $y_0,y_1,\dots,y_m\in E$ tels que $y=y_0$, $z=y_m$ et, pour tout $k\in[\![0,m-1]\!]$, $y_k\to y_{k+1}$ d'autre part.
Avec l'entier $n+m$ et la suite d'éléments $x_0,x_1,\dots,x_n=y_0,y_1,\dots,y_m\in E$, on observe de $x\xrightarrow{*}z$.

▹ La clôture transitive n'est en général pas une relation d'équivalence à cause de la propriété de symétrie : pour construire un contre-exemple, il suffit de considérer $E = \{0,1\}$ et la relation $\to$ telle que $0 \to 1$ et $1 \to 0$.
3. ▹ La relation sur $\mathbb{N}$ définie par $n \to n-1$ pour tout $n \geq 1$ termine.
▹ La relation sur $\mathbb{N}$ définie par $n \to n+1$ pour tout $n \in \mathbb{N}$ termine.

Partie B – Réécriture

1. a. Les éléments irréductibles de E sont ceux sans facteur ab donc les mots de la forme $b^k a^l$ pour des entiers naturels k et l.
b. La longueur des mots lors d'une succession de réécritures est une suite strictement décroissante d'entiers naturels ; or, il n'existe pas de telle suite infinie donc la relation termine.
c. ▹ On vérifie que, pour tous mots m, $m' \in \Sigma^*$ tels que $m \to m'$, on a $\delta(m) = \delta(m')$: la quantité est conservée par réécriture.
▹ Vérifions que la quantité Δ est aussi conservée par réécriture. Soit m, m' deux mots.

$$\begin{aligned}\Delta(mabm') &= \max\{\Delta(m), \delta(m)-1, \delta(m)+\Delta(m')\}\\ &= \max\{\Delta(m), \delta(m)+\Delta(m')\} \qquad (\text{car } \delta(m) \leq \Delta(m))\\ &= \Delta(mm').\end{aligned}$$

d. D'après la terminaison de $\to$ et la forme des mots irréductibles, il existe un couple d'entiers (p,q) tel que $m \xrightarrow{*} b^p a^q$.
Forts des deux remarques de la question précédente, on déduit de la réduction $m \xrightarrow{*} b^p a^q$ que $\delta(m) = p-q$ et $\Delta(m) = p$ donc $p = \Delta(m)$ et $q = \Delta(m) - \delta(m)$.
2. a. Les éléments irréductibles de E sont ceux sans facteur ab donc les mots de la forme $b^k a^l$ pour des entiers k, l éventuellement nuls.
b. Au cours d'une étape de réécriture pour cette relation, seul un terme dans cette quantité est modifié (celui qui correspond au b du motif réécrit ; pour toutes les autres occurrences de b, le nombre de a à gauche reste inchangé). Notons k le nombre de a à gauche de cette occurrence. Dans l'expression de la quantité, on passe donc de 3^k (avant réécriture) à $3^{k-1} + 3^{k-1} < 3^k$ (après réécriture).
c. La quantité calculée par f est un entier naturel et strictement décroissante au cours des réécritures : la relation termine.
3. On remarque que l'on a successivement $a^2 b \to ab^2a^2 \to b^2a^2ba^2$ et le mot initial est facteur du mot obtenu : on obtient donc une suite infinie de réécritures en recommençant les mêmes étapes sur ce facteur : la relation ne termine pas.

Partie C – Propriétés de confluence

1. Raisonnons par l'absurde en supposant qu'il existe $x \in E$ et deux éléments irréductibles distincts y et z tels que $x \xrightarrow{*} y$ et $x \xrightarrow{*} z$. D'après la propriété de confluence, il existe $t \in E$ tel que $y \xrightarrow{*} t$ et $z \xrightarrow{*} t$ ce qui n'est possible que si $y = t$ (irréductibilité de y) et $z = t$ (irréductibilité de z) d'où $y = z$: contradiction.
2. Pour montrer qu'une relation confluente est localement confluente, il suffit de remarquer que, pour tous x, y et $z \in E$, $x \to y$ et $x \to z$ entraîne $x \xrightarrow{*} y$ et $x \xrightarrow{*} z$.
3. ▹ La relation n'est pas confluente car $1 \xrightarrow{*} 0$ et $1 \xrightarrow{*} 3$ (en passant par 2) et pourtant, il n'existe pas de $x \in [\![0,3]\!]$ tel que $0 \xrightarrow{*} x$ et $3 \xrightarrow{*} x$ car 0 et 3 sont irréductibles.
▹ En revanche, la relation est localement confluente. En effet, par symétrie, un seul cas est à traiter, par exemple $1 \to 0$ et $1 \to 2$: mais on a $0 \xrightarrow{*} 0$ et $2 \xrightarrow{*} 0$.
4. ▹ La relation n'est pas confluente car $0 \xrightarrow{*} \circ$ et $0 \xrightarrow{*} \bullet$ (en passant par 1) et $\circ$ et $\bullet$ sont des éléments irréductibles.
▹ Pour étudier la confluence locale, il suffit d'étudier deux cas :
- $2k \to \circ$ et $2k \to 2k+1$ mais alors $\circ \xrightarrow{*} \circ$ et $2k+1 \xrightarrow{*} \circ$,
- $2k+1 \to \bullet$ et $2k+1 \to 2k+2$ mais alors $\bullet \xrightarrow{*} \bullet$ et $2k+2 \xrightarrow{*} \bullet$.

La relation est donc localement confluente.

Partie D – Lemme de Newman

1. Raisonnons par l'absurde en supposant qu'il n'existe pas d'élément irréductible et construisons par récurrence une suite $(x_n)_n$ telle que, pour tout $n \in \mathbb{N}$, $x_n \to x_{n+1}$.
• Posons $x_0 \in E$ un élément quelconque.
• Soit $n \in \mathbb{N}$. Supposons la suite construite jusqu'au rang n. L'élément x_n n'est pas irréductible donc il existe x_{n+1} tel que $x_n \to x_{n+1}$.
La suite ainsi construite contredit l'hypothèse donc il existe au moins un élément irréductible.
2. Raisonnons par l'absurde en supposant qu'il existe $x \in E$ tel que $\mathscr{P}(x) =$ Faux et construisons une suite $(x_n)_n$ telle que, pour tout $n \in \mathbb{N}$, $x_n \to x_{n+1}$ et $\mathscr{P}(x_n) =$ Faux.
• Posons $x_0 = x$.
• Soit $n \in \mathbb{N}$. Supposons la suite construite jusqu'au rang n. L'élément x_n n'est pas irréductible car $\mathscr{P}(x_n) =$ Faux et on a la première hypothèse. D'après la seconde hypothèse, il existe x_{n+1} tel que $x_n \to x_{n+1}$ et $\mathscr{P}(x_{n+1}) =$ Faux.
Ceci contredit l'hypothèse sur la relation $\to$ donc $\mathscr{P}(x) =$ Vrai pour tout $x \in E$.
3. Utilisons le principe d'induction de la question précédente avec, pour tout $x \in E$, $\mathscr{P}(x)$ la proposition : pour tous y et $z \in E$ tels que $x \xrightarrow{*} y$ et $x \xrightarrow{*} z$, il existe $t \in E$ tel que $y \xrightarrow{*} t$ et $z \xrightarrow{*} t$.
▹ La propriété est évidente pour les éléments irréductibles.
▹ Soit $y \in E$ tel que $\mathscr{P}(x) =$ Vrai pour tout $x \in E$ tel que $y \to x$. Considérons y_1 et $y_2 \in E$ tels que $y \xrightarrow{*} y_1$ et $y \xrightarrow{*} y_2$.
▹▹ Si $y = y_1$, il suffit de remarquer que $y_1 \xrightarrow{*} y_2$ et $y_2 \xrightarrow{*} y_2$. On procède de même si $y = y_2$.
▹▹ Si $y \neq y_1$ et $y \neq y_2$, il existe y_1' et y_2' tels que $y \to y_1'$, $y_1' \xrightarrow{*} y_1$, $y \to y_2'$ et $y_2' \xrightarrow{*} y_2$. Par hypothèse de confluence locale (en x), il existe z tel que $y_1' \xrightarrow{*} z$ et $y_2' \xrightarrow{*} z$. D'après la propriété $\mathscr{P}(y_1')$, il existe z_1 tel que $y_1 \xrightarrow{*} z_1$ et $z \xrightarrow{*} z_1$. Par transitivité de $\xrightarrow{*}$, $y_2' \xrightarrow{*} z_1$. On peut donc appliquer la propriété $\mathscr{P}(y_2')$: il existe z_2 tel que $z_1 \xrightarrow{*} z_2$ et $y_2 \xrightarrow{*} z_2$. En combinant tous les résultats obtenus, on a

$$y_1 \xrightarrow{*} z_1 \xrightarrow{*} z_2, \qquad y_2 \xrightarrow{*} z_2$$

ce qui établit donc $\mathscr{P}(x)$.
D'après la question précédente, la relation $\to$ est confluente.

Chapitre 2

Nombres complexes

1. Corps des nombres complexes — **2.** Exponentielle complexe — **3.** Équations algébriques

Objectifs et compétences du programme

- Savoir manipuler les nombres complexes.
- Interpréter une situation géométrique dans le plan avec les affixes complexes des points.
- Résoudre les équations algébriques de degré 2.
- Savoir déterminer des racines $n^{\text{ièmes}}$.

Jean-Robert Argand, mathématicien suisse (1768-1822)
Alors que les nombres complexes introduits par les algébristes italiens de la Renaissance semblent toujours aussi mystérieux et artificiels, le mathématicien amateur Argand a l'idée de les interpréter comme les points du « plan ». Il publie en 1806 son *Essai sur une manière de représenter les quantités imaginaires dans les constructions géométriques*, dans lequel il aborde :

- la représentation géométrique des nombres complexes ;
- l'identification d'un nombre complexe à un point du plan
- les opérations correspondant à l'addition et à la multiplication.

On lui doit également une autre pépite mathématique : une démonstration du théorème de D'Alembert et Gauss. La « preuve » analytique de D'Alembert est incomplète ; le premier travail de Gauss ajoute un peu de rigueur dans cette approche mais c'est bien à Argand que l'on doit la première preuve solide.

COURS

Dans ce chapitre, on explore une partie de la richesse des nombres complexes. Dans un premier temps, on les construit et on vérifie les différentes règles de manipulation. Dans un deuxième temps, on rejoint les notions d'angles *via* les arguments. Après ces considérations plus géométriques, on revient aux équations polynomiales (algébriques) qui étaient à l'origine des nombres complexes pour les mathématiciens de la Renaissance italienne.

1. Corps des nombres complexes

Ce premier paragraphe sert à comprendre une construction de $\mathbb{C}$.

1.1. Construction du corps $\mathbb{C}$

Commençons par donner une définition abstraite de la structure algébrique de corps. Intuitivement, il s'agit d'un ensemble avec deux opérations pour lesquelles toutes les manipulations simples se passent bien. On étudiera dans un autre chapitre les structures algébriques pour elles-mêmes.

Définition 2.1.

Un corps est un ensemble K muni de deux lois de composition internes $+$ et $\times$ telles que
- $+$ est associative, commutative, admet un élément neutre noté 0 et tout élément de K admet un symétrique pour $+$.
- $\times$ est associative, commutative, admet un élément neutre distinct de 0 et noté 1 et tout élément de K différent de 0 admet un symétrique pour $\times$.
- $\times$ est distributive sur $+$.

Exemples

Les ensembles $\mathbb{R}$, $\mathbb{Q}$ et $\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2},\ a, b \in \mathbb{Q}\}$ sont des corps pour l'addition et la multiplication usuelles.

L'idée que nous allons utiliser pour construire $\mathbb{C}$ consiste à partir de $\mathbb{R}^2$, l'ensemble des couples de réels, et d'en faire un corps en le munissant de deux lois de composition internes bien choisies. Pour cela posons, pour tous (x, y) et $(x', y') \in \mathbb{R}^2$,

$$(x, y) + (x', y') = (x + x', y + y'),$$
$$(x, y) \times (x', y') = (x x' - y y', x' y + x y').$$

Remarque

Comme on cherche à retrouver les complexes dont on a l'habitude, on définit le produit pour obtenir la formule suivante

$$(x + i y)(x' + i y') = (x x' - y y') + i(x y' + x' y).$$

On vérifie rapidement que $\mathbb{R}^2$ avec ces lois est bien un corps que l'on notera $\mathbb{C}$.

Notation

À partir de la construction précédente, on note le couple (x, y) comme $x + i y$ ce qui revient à noter $(1, 0) = 1$ et $(0, 1) = i$.
▹ On vérifie aisément que $i^2 = -1$ d'après la définition de la loi $\times$.
▹ Les réels x et y sont appelés respectivement partie réelle et partie imaginaire du complexe $x + i y$. Les parties réelle et imaginaire d'un complexe z sont notées $\mathfrak{Re}(z)$ et $\mathfrak{Im}(z)$ respectivement.

Proposition 2.2.

Les applications partie réelle et partie imaginaire définies de $\mathbb{C}$ dans $\mathbb{R}$ sont $\mathbb{R}$-linéaires, c'est-à-dire pour tous z, $z' \in \mathbb{C}$ et $\lambda, \lambda' \in \mathbb{R}$,

$$\mathfrak{Re}(\lambda z + \lambda' z') = \lambda \mathfrak{Re}(z) + \lambda' \mathfrak{Re}(z'),$$
$$\mathfrak{Im}(\lambda z + \lambda' z') = \lambda \mathfrak{Im}(z) + \lambda' \mathfrak{Im}(z').$$

Pour exprimer ces relations, on dit aussi que, pour la partie réelle ou la partie imaginaire, l'image d'une combinaison linéaire à coefficients réels est la combinaison linéaire (avec les mêmes coefficients) des images.

1.2. Conjugaison, module

Définition 2.3.

Soit $z = x + iy$ un nombre complexe avec x, $y \in \mathbb{R}$.
▹ Le conjugué de z est le nombre complexe $\overline{z} = x - iy$.
▹ Le module de z est le réel positif $|z| = \sqrt{x^2 + y^2}$.

Remarque

On remarque que $|z| = 0$ si, et seulement si, $z = 0$.

Des calculs immédiats donnent la proposition suivante.

Proposition 2.4.

Pour tout $z \in \mathbb{C}$,

$$z + \overline{z} = 2\mathfrak{Re}(z), \qquad z - \overline{z} = 2i\mathfrak{Im}(z), \qquad z\overline{z} = |z|^2.$$

Conseils méthodologiques

La dernière propriété reliant module et conjugué est très utile en pratique car elle permet de « se débarrasser » des nombres complexes non réels au dénominateur en multipliant par le conjugué. Plus précisément, pour tout complexe z non nul,

$$\frac{1}{z} = \frac{\overline{z}}{|z|^2}.$$

Exemple

Déterminons les parties réelle et imaginaire du nombre complexe sous forme de fraction $\frac{3+i}{2+i}$. En multipliant par le conjugué du dénominateur les deux termes de cette fraction, on obtient

$$\frac{3+i}{2+i} = \frac{(3+i)(2-i)}{5} = \frac{1}{5}(7-i).$$

Ainsi, la partie réelle est $\frac{7}{5}$, la partie imaginaire $-\frac{1}{5}$.

Une simple vérification permet d'obtenir la proposition suivante.

Proposition 2.5.

Pour tous z_1, $z_2 \in \mathbb{C}$,
$$\overline{z_1+z_2}=\overline{z_1}+\overline{z_2}, \qquad \overline{z_1 z_2}=\overline{z_1}.\overline{z_2}.$$

Démonstration

Notons $z_1 = x_1 + i\,y_1$ et $z_2 = x_2 + i\,y_2$ avec x_1, y_1, x_2 et $y_2 \in \mathbb{R}$. Alors,
$$\begin{aligned}\overline{z_1+z_2} &= \overline{x_1+x_2+i(y_1+y_2)}\\ &= x_1+x_2-i(y_1+y_2)\\ &= (x_1-i\,y_1)+(x_2-i\,y_2)=\overline{z_1}+\overline{z_2}.\end{aligned}$$

Par ailleurs,
$$\begin{aligned}\overline{z_1 z_2} &= \overline{x_1x_2-y_1y_2+i(x_1y_2+x_2y_1)}\\ &= x_1x_2-y_1y_2-i(x_1y_2+x_2y_1)\\ &= (x_1-i\,y_1)(x_2-i\,y_2)=\overline{z_1}.\overline{z_2}.\end{aligned}$$

■

On en déduit immédiatement le corollaire suivant.

Corollaire 2.6.

L'application conjugaison $z \longmapsto \overline{z}$ est $\mathbb{R}$-linéaire au sens défini ci-dessus.

Proposition 2.7.

Pour tous z_1, $z_2 \in \mathbb{C}$, $|z_1 z_2| = |z_1||z_2|$.
Si, de plus, $z_2 \neq 0$, alors
$$\left|\frac{z_1}{z_2}\right| = \frac{|z_1|}{|z_2|}.$$

Démonstration

Ces identités découlent directement de l'écriture $\overline{z}.z = |z|^2$. En effet,
$$|z_1 z_2|^2 = z_1 z_2 \overline{z_1 z_2} = z_1\overline{z_1}z_2\overline{z_2} = |z_1|^2|z_2|^2.$$

Comme les modules sont positifs, on obtient le résultat en passant aux racines carrées.
On procède de même pour le quotient. ■

Proposition 2.8. Identité remarquable

Pour tous z_1, $z_2 \in \mathbb{C}$,
$$|z_1+z_2|^2 = |z_1|^2 + 2\mathfrak{Re}(z_1\overline{z_2}) + |z_2|^2.$$

Démonstration

Il suffit d'utiliser la formule exprimant le module au carré :
$$\begin{aligned}|z_1+z_2|^2 &= (z_1+z_2)(\overline{z_1}+\overline{z_2})\\ &= z_1\overline{z_1}+z_2\overline{z_1}+z_1\overline{z_2}+z_2\overline{z_2}.\end{aligned}$$

On conclut avec la formule donnant la partie réelle. ■

Conseils méthodologiques

D'après toutes les formules précédentes, on déduit qu'il est plus pratique de manipuler les modules au carré que les modules. On pensera donc à élever au carré les modules qui apparaissent dans les expressions.

Précisons maintenant quelques inégalités en module.

Proposition 2.9.

Pour tout $z \in \mathbb{C}$,

$$|\mathfrak{Re}(z)| \leq |z| \qquad \text{et} \qquad |\mathfrak{Im}(z)| \leq |z|.$$

Il y a égalité respectivement pour $z \in \mathbb{R}$ et $z \in i\mathbb{R}$.

Démonstration

Pour tout $z \in \mathbb{C}$,

$$\mathfrak{Re}(z)^2 \leq \mathfrak{Re}(z)^2 + \mathfrak{Im}(z)^2 = |z|^2,$$
$$\mathfrak{Im}(z)^2 \leq \mathfrak{Re}(z)^2 + \mathfrak{Im}(z)^2 = |z|^2.$$

Le cas d'égalité correspond à $\mathfrak{Im}(z) = 0$ et $\mathfrak{Re}(z) = 0$ respectivement. ■

Proposition 2.10. Inégalités triangulaires

▹ Pour tous z_1, $z_2 \in \mathbb{C}$,

$$|z_1 + z_2| \leq |z_1| + |z_2|.$$

Il y a égalité si, et seulement si, $z_1\overline{z_2} \in \mathbb{R}_+$, c'est-à-dire s'il existe $\lambda \geq 0$, $z_1 = \lambda z_2$ ou $z_2 = 0$.

▹ Pour tous z_1, $z_2 \in \mathbb{C}$,

$$||z_1| - |z_2|| \leq |z_1 - z_2|.$$

Démonstration

- Soit z_1, $z_2 \in \mathbb{C}$; l'identité remarquable donne

$$|z_1 + z_2|^2 = |z_1|^2 + 2\mathfrak{Re}(z_1\overline{z_2}) + |z_2|^2.$$

  En utilisant l'inégalité (établie au point précédent), $2\mathfrak{Re}(z_1\overline{z_2}) \leq 2|z_1||z_2|$, on en déduit la majoration $|z_1 + z_2| \leq |z_1| + |z_2|$.
  Le cas d'égalité correspond à $2\mathfrak{Re}(z_1\overline{z_2}) = 2|z_1 z_2|$ donc $z_1\overline{z_2} \in \mathbb{R}_+$.
- Déduisons cette inégalité de la précédente. Pour tous $z_1, z_2 \in \mathbb{C}$,

$$|z_1| = |z_2 + z_1 - z_2| \leq |z_2| + |z_1 - z_2|,$$
$$|z_2| = |z_1 + z_2 - z_1| \leq |z_1| + |z_1 - z_2|.$$

  Donc $||z_1| - |z_2|| \leq |z_1 - z_2|$.

■

1.3. Interprétation géométrique

La construction de $\mathbb{C}$ repose sur une identification ensembliste de $\mathbb{C}$ au plan ; on peut en déduire facilement une utilisation des complexes en géométrie.

Définition 2.11.

▹ L'affixe d'un point $M(x,y)$ du plan est le complexe $z = x + iy$.

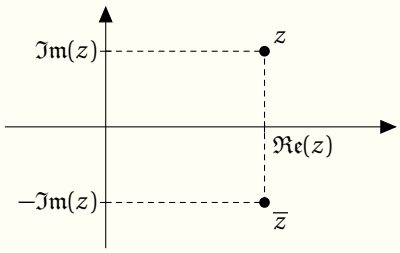
▹ L'affixe d'un vecteur $\overrightarrow{AB}$ est le complexe $z_B - z_A$ où les points A et B ont respectivement pour affixe z_A et z_B.

Proposition 2.12.

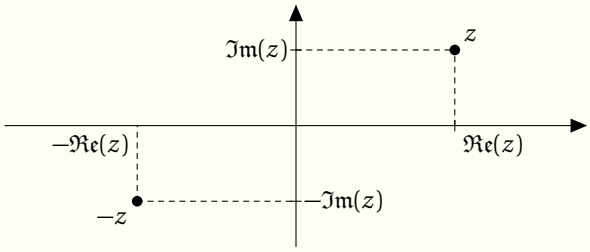
La partie réelle (respectivement imaginaire) d'un complexe z est la coordonnée du point d'affixe z sur l'axe des abscisses (respectivement ordonnées).

Remarque

On déduit de cette proposition que le conjugué d'un complexe z est l'affixe du symétrique du point d'affixe z par rapport à l'axe des abscisses.



On ne confondra pas $\overline{z}$ et $-z$:



On vérifie par un calcul rapide l'expression suivante de la distance entre deux points.

Proposition 2.13.

La distance entre deux points du plan A et B d'affixe z_A et z_B est le module $|z_A - z_B|$.

Remarque

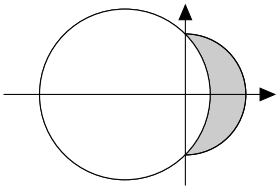
L'application $d : \left\{ \begin{array}{ccc} \mathbb{C}^2 & \to & \mathbb{R}_+ \\ (z,z') & \mapsto & |z-z'| \end{array} \right.$ est, en un sens plus général, une distance car elle vérifie les propriétés suivantes, déduites des propriétés du module détaillées précédemment.

- $\forall z, z' \in \mathbb{C}, \quad d(z,z') = 0 \quad \Rightarrow \quad z = z'$.
- $\forall z, z' \in \mathbb{C}, \quad d(z,z') = d(z',z)$.
- $\forall z, z', z'' \in \mathbb{C}, \quad d(z,z') \leq d(z,z'') + d(z'',z')$.

De cette expression de la distance usuelle en termes de complexes, on déduit que l'équation complexe d'un cercle de rayon $R > 0$ et de centre le point d'affixe $\omega \in \mathbb{C}$ est $|z-\omega| = R$ et que celle du disque correspondant est $|z-\omega| \leq R$.

Exemple Lunule d'Hippocrate

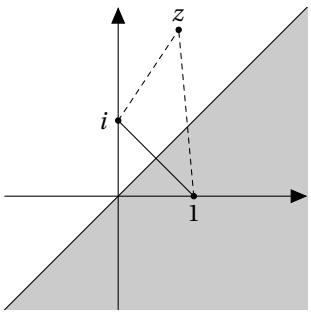
La lunule d'Hippocrate est la partie du plan intérieure au cercle de centre $(0,0)$ et de rayon 1 et extérieure au cercle de centre $(-1,0)$ et passant par $(0,1)$:



Un point M d'affixe z appartient à la lunule si, et seulement si, $|z| \leq 1$ (M est dans le petit disque) et $|z+1| \geq \sqrt{2}$ (mais pas dans le grand).

Exemple

▹ L'axe des ordonnées est l'ensemble des points d'affixe z vérifiant $\mathfrak{Re}(z) = 0$ ou, de manière équivalente $|z-1| = |z+1|$. Pour montrer ce deuxième point, on peut élever au carré puis utiliser l'identité remarquable.
Toutefois, il existe une interprétation géométrique plus intuitive. En appelant A, B, M les points d'affixe $1, -1$ et z, on trouve que la condition $|z-1| = |z+1|$ équivaut à $AM = BM$ soit, M appartient à la médiatrice du segment $[A, B]$.
▹ Avec un raisonnement analogue, on obtient que les points d'affixe z tels que $|z-i| = |z-1|$ sont les points de la médiatrice des points d'affixe 1 et i, c'est-à-dire à la première bissectrice. Par conséquent, les points d'affixe z tels que $|z-i| < |z-1|$ sont les points strictement au-dessus de la première bissectrice (zone grisée).



2. Exponentielle complexe

2.1. Compléments de trigonométrie : fonction tangente

On ne cherche pas à définir ici les fonctions trigonométriques mais on se limite à quelques rappels utiles pour la suite. Les fonctions sin et cos sont connues et on en déduit des propriétés de la fonction $\tan = \frac{\sin}{\cos}$.
Rappelons les formules d'addition des fonctions trigonométriques cos et sin.

Proposition 2.14. Formules d'addition et de duplication

Pour tous θ, $\varphi \in \mathbb{R}$,

$$\begin{aligned}\sin(\theta+\varphi) &= \sin(\theta)\cos(\varphi)+\cos(\theta)\sin(\varphi),\\ \cos(\theta+\varphi) &= \cos(\theta)\cos(\varphi)-\sin(\theta)\sin(\varphi).\end{aligned}$$

En particulier, pour tout $\theta \in \mathbb{R}$,

$$\begin{aligned}\sin(2\theta) &= 2\sin(\theta)\cos(\theta),\\ \cos(2\theta) &= \cos^2(\theta)-\sin^2(\theta) = 2\cos^2(\theta)-1 = 1-2\sin^2(\theta).\end{aligned}$$

Les formules d'addition sont particulièrement utiles lorsqu'il s'agit de linéariser des produits de deux termes sous la forme d'un cosinus ou d'un sinus.

Corollaire 2.15. Linéarisation de produit

Pour tous θ, $\varphi \in \mathbb{R}$,

$$\begin{aligned}2\sin(\theta)\cos(\varphi) &= \sin(\theta+\varphi)+\sin(\theta-\varphi)\\ 2\cos(\theta)\sin(\varphi) &= \sin(\theta+\varphi)-\sin(\theta-\varphi)\\ 2\cos(\theta)\cos(\varphi) &= \cos(\theta+\varphi)+\cos(\theta-\varphi)\\ -2\sin(\theta)\sin(\varphi) &= \cos(\theta+\varphi)-\cos(\theta-\varphi).\end{aligned}$$

Démonstration

Chacune de ses formules se démontre en utilisant les formules d'addition dans le membre de droite. ■

Remarque

Les formules ci-dessus permettent de passer d'une forme linéarisée comme somme de fonctions trigonométriques (utile pour calculer une primitive par exemple) à une forme factorisée comme produit (utile pour étudier le signe par exemple). Il convient de savoir les utiliser dans les deux sens.

Par exemple, on peut lire la troisième formule de droite à gauche de la façon suivante : pour tous θ, $\varphi \in \mathbb{R}$,

$$\cos(\theta)+\cos(\varphi) = 2\cos\left(\frac{\theta+\varphi}{2}\right)\cos\left(\frac{\theta-\varphi}{2}\right).$$

Exemple

Soit p, $q \in \mathbb{N}$. Calculons l'intégrale $I = \int_0^{2\pi} \cos(pt)\cos(qt)\,\mathrm{d}t$. Pour cela, linéarisons l'expression dans l'intégrale.

$$\begin{aligned}I &= \frac{1}{2}\int_0^{2\pi} \cos(p+q)t+\cos(p-q)t\,\mathrm{d}t\\ &= \begin{cases} 0 & \text{si } p \neq q,\\ \pi & \text{sinon.}\end{cases}\end{aligned}$$

Exemple

Déterminons le signe de la fonction $f : t \mapsto \cos(t)+\cos(3t)$ sur l'intervalle $[0,\pi]$. D'après la formule de linéarisation (utilisée de droite à gauche), pour tout $t \in [0,\pi]$,

$$\cos(t)+\cos(3t) = 2\cos(2t)\cos(-t).$$

Comme $\cos(2t)$ change de signe en $\frac{\pi}{4}$ et $\frac{3\pi}{4}$ et $\cos(-t)$ change de signe en $\frac{\pi}{2}$, on en déduit que la fonction f est positive sur $[0, \frac{\pi}{4}] \cup [\frac{\pi}{2}, \frac{3\pi}{4}]$.

Définition 2.16.

La fonction tangente, notée tan est la fonction définie sur $\mathbb{R} \setminus \{\frac{\pi}{2} + k\pi,\ k \in \mathbb{Z}\}$ comme le quotient du sinus par le cosinus.

Remarque

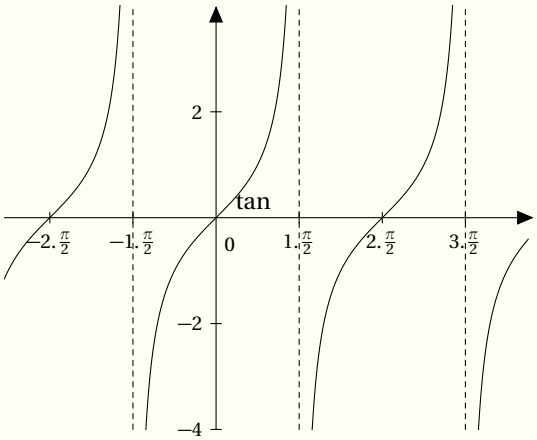
Cette définition nous permet d'obtenir la π-périodicité de tan puis quelques valeurs :

x	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$
$\tan(x)$	0	$\frac{1}{\sqrt{3}}$	1	$\sqrt{3}$

Le calcul avec les dérivées de cos et sin donne

$$\tan' = \frac{1}{\cos^2} = 1 + \tan^2.$$

Ces expressions permettent, entre autres, d'obtenir le tableau de variations puis l'allure suivante de la courbe représentative.



Corollaire 2.17.

Pour tous $\theta,\ \varphi \in \mathbb{R} \setminus \{\frac{\pi}{2} + k\pi,\ k \in \mathbb{Z}\}$ tels que $\theta + \varphi \in \mathbb{R} \setminus \{\frac{\pi}{2} + k\pi,\ k \in \mathbb{Z}\}$,

$$\tan(\theta + \varphi) = \frac{\tan(\theta) + \tan(\varphi)}{1 - \tan(\theta)\tan(\varphi)}.$$

En particulier, pour tout $\theta \in \mathbb{R} \setminus \{\frac{\pi}{2} + k\pi,\ k \in \mathbb{Z}\}$ tel que $2\theta \in \setminus \{\frac{\pi}{2} + k\pi,\ k \in \mathbb{Z}\}$,

$$\tan(2\theta) = \frac{2\tan(\theta)}{1 - \tan^2(\theta)}.$$

Démonstration
Écrivons $\tan = \frac{\sin}{\cos}$, utilisons les formules d'addition puis factorisons par le terme adapté

$$\begin{aligned}\tan(\theta+\varphi) &= \frac{\sin(\theta+\varphi)}{\cos(\theta+\varphi)} = \frac{\sin(\theta)\cos(\varphi)+\cos(\theta)\sin(\varphi)}{\cos(\theta)\cos(\varphi)-\sin(\theta)\sin(\varphi)}\\ &= \frac{\cos(\theta)\cos(\varphi)(\tan(\theta)+\tan(\varphi))}{\cos(\theta)\cos(\varphi)(1-\tan(\theta)\tan(\varphi))} = \frac{\tan(\theta)+\tan(\varphi)}{1-\tan(\theta)\tan(\varphi)}.\end{aligned}$$

■

Exemple

Calculons $t = \tan\frac{\pi}{8}$. Comme $\tan\frac{\pi}{4} = 1$, on obtient avec la formule donnant $\tan(\frac{\pi}{8}+\frac{\pi}{8})$ que

$$1 = \frac{2t}{1-t^2},$$

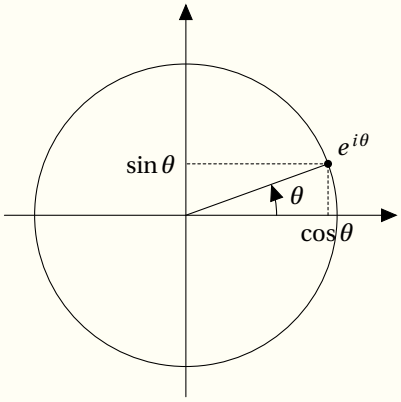
donc t est une solution de l'équation polynomiale $t^2+2t-1=0$. Comme les racines sont $-1\pm\sqrt{2}$. Or, $\frac{\pi}{8}\in[0,\frac{\pi}{2}[$, donc $t\geq 0$ et $\tan\frac{\pi}{8} = \sqrt{2}-1$.

2.2. Exponentielle d'un imaginaire pur

Définition 2.18.

Pour tout $\theta\in\mathbb{R}$, on définit le complexe $e^{i\theta}$ par $\cos\theta + i\sin\theta$.

Remarque
Cette définition permet de décrire les complexes de la forme $\cos\theta + i\sin\theta$ comme les affixes des points du cercle unité.

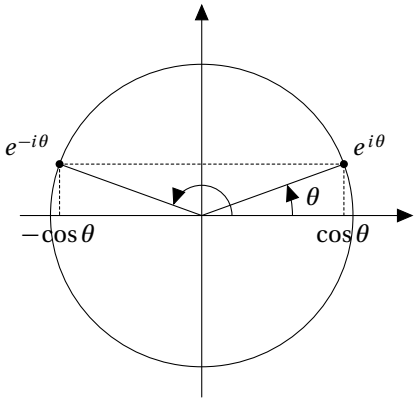


Conseils méthodologiques

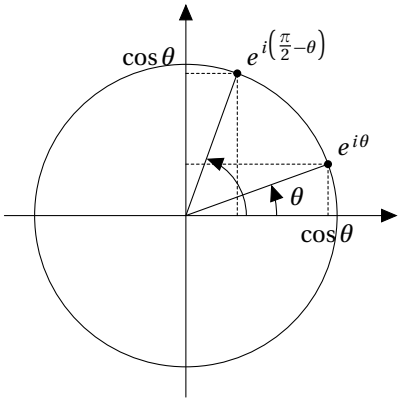
Cette représentation permet de retrouver rapidement des propriétés des fonctions trigonométriques.

Exemple

Voici l'illustration pour retrouver la relation trigonométrique $\cos(\pi-\theta)=-\cos(\theta)$ en considérant les abscisses (mais aussi $\sin(\pi-\theta)=\sin\theta$ en considérant les ordonnées)



et l'illustration pour retrouver la relation trigonométrique $\cos(\frac{\pi}{2}-\theta)=\sin(\theta)$



On en déduit directement que, pour tout $\theta\in\mathbb{R}$, $\overline{e^{i\theta}}=e^{-i\theta}$ et donc les relations suivantes comme expressions des parties réelles et parties imaginaires.

Proposition 2.19. Formules d'Euler

Pour tous $\theta\in\mathbb{R}$,

$$\cos\theta=\frac{e^{i\theta}+e^{-i\theta}}{2},\qquad \sin\theta=\frac{e^{i\theta}-e^{-i\theta}}{2i}.$$

On déduit des formules d'addition la propriété suivante.

Proposition 2.20. Propriété de morphisme

Pour tous $\theta,\varphi\in\mathbb{R}$, $e^{i\theta}e^{i\varphi}=e^{i(\theta+\varphi)}$.

Démonstration
Il suffit de revenir aux écritures cartésiennes

$$\begin{aligned} e^{i\theta}e^{i\varphi} &= (\cos\theta + i\sin\theta)(\cos\varphi + i\sin\varphi) \\ &= \cos(\theta)\cos(\varphi) - \sin(\theta)\sin(\varphi) + i(\sin(\theta)\cos(\varphi) + \cos(\theta)\sin(\varphi)) \\ &= \cos(\theta+\varphi) + i\sin(\theta+\varphi) = e^{i(\theta+\varphi)}. \end{aligned}$$

■

Conseils méthodologiques

On se sert souvent des relations d'Euler combinées à cette propriété pour factoriser une somme ou une différence de deux exponentielles complexes. À chaque fois, on procède de la même manière : on met en facteur l'exponentielle complexe avec pour argument la demi-somme des arguments

$$\begin{aligned} e^{i\theta} + e^{i\varphi} &= e^{i\frac{\theta+\varphi}{2}} 2\cos\frac{\theta-\varphi}{2}, \\ e^{i\theta} - e^{i\varphi} &= e^{i\frac{\theta+\varphi}{2}} 2i\sin\frac{\theta-\varphi}{2}. \end{aligned}$$

Remarque
On peut aussi factoriser une expression du type $a\cos(\theta) + b\sin(\theta)$ avec $(a,b) \neq (0,0)$. En effet, on choisit $\varphi \in \mathbb{R}$ tel que $a + ib = \sqrt{a^2+b^2}e^{i\varphi}$ et on remarque que

$$a\cos(\theta) + b\sin(\theta) = \mathfrak{Re}((a-ib)e^{i\theta}) = \sqrt{a^2+b^2}\mathfrak{Re}(e^{i(\theta-\varphi)}) = \sqrt{a^2+b^2}\cos(\theta-\varphi).$$

Cette factorisation s'interprète facilement pour un phénomène physique : les quantités $\sqrt{a^2+b^2}$ et φ sont respectivement l'amplitude et le déphasage du signal $a\cos(\theta) + b\sin(\theta)$.

Exemple

Calculons, pour tout $\theta \in \mathbb{R}$, le complexe $\frac{e^{i\theta}+1}{e^{i\theta}-1}$.
Pour cela, reconnaissons au numérateur comme au dénominateur la différence de deux exponentielles complexes (avec $1 = e^{i0}$) et factorisons l'exponentielle de la demi-somme des arguments.

$$\begin{aligned} \frac{e^{i\theta}+1}{e^{i\theta}-1} &= \frac{e^{i\frac{\theta}{2}}(e^{i\frac{\theta}{2}} + e^{-i\frac{\theta}{2}})}{e^{i\frac{\theta}{2}}(e^{i\frac{\theta}{2}} - e^{-i\frac{\theta}{2}})} \\ &= \frac{2\cos\frac{\theta}{2}}{2i\sin\frac{\theta}{2}} = -i\left(\tan\frac{\theta}{2}\right)^{-1}. \end{aligned}$$

Par récurrence sur l'entier n, on obtient la formule suivante.

Corollaire 2.21. Formule de Moivre

Pour tout $n \in \mathbb{N}$ et tout $\theta \in \mathbb{R}$, $(e^{i\theta})^n = e^{in\theta}$ et donc

$$(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta.$$

Démonstration
▹ Le résultat est évident pour $n = 0$ puisqu'elle s'écrit $1 = 1$.

▹ Soit $n \in \mathbb{N}$ tel que, pour tout $\theta \in \mathbb{R}$, $(e^{i\theta})^n = e^{in\theta}$. Alors, pour tout $\theta \in \mathbb{R}$, on applique successivement l'hypothèse de récurrence et la propriété de morphisme,

$$(e^{i\theta})^{n+1} = (e^{i\theta})^n e^{i\theta} = e^{in\theta} e^{i\theta} = e^{in\theta + i\theta} = e^{i(n+1)\theta}.$$

■

2.3. Groupe $\mathbb{U}$ des complexes de module 1

L'ensemble des complexes de module 1 est noté $\mathbb{U}$.

Proposition 2.22.

L'application $\left\{ \begin{array}{ccc} \mathbb{R} & \to & \mathbb{U} \\ \theta & \mapsto & e^{i\theta} \end{array} \right.$ est surjective mais pas injective.

En particulier, $\mathbb{U} = \left\{ e^{i\theta},\ \theta \in \mathbb{R} \right\}$.

Démonstration

▹ La non-injectivité est évidente d'après la 2π-périodicité des fonctions trigonométriques.

▹ Soit $z = x + iy \in \mathbb{U}$ avec $x, y \in \mathbb{R}$. Comme $x^2 \leq x^2 + y^2 = 1$, il existe θ tel que $x = \cos\theta$ et donc $y = \pm\sin\theta$ car $y^2 = 1 - x^2$. Ainsi, $z = e^{i\theta}$ ou $z = e^{-i\theta}$. ■

Définition 2.23. Groupe

Un ensemble G est un groupe s'il est muni d'une loi de composition interne $\times$ associative, qui admet un élément neutre et telle que tout élément de G admet un symétrique pour cette loi.
Si, de plus, $\times$ est commutative, alors G est abélien (ou commutatif).

Exemple

Les ensembles $(\mathbb{R}^*, \times)$, $(\mathbb{C}^*, \times)$ sont des groupes abéliens.

Proposition 2.24.

L'ensemble $\mathbb{U}$ est un groupe abélien pour la multiplication.

Démonstration

La loi $\times$ est associative et commutative sur $\mathbb{U}$ car elle l'est sur $\mathbb{C}$; l'élément neutre 1 appartient à $\mathbb{U}$. De plus, pour tout $z, z' \in \mathbb{U}$, $|zz'| = |z||z'| = 1$ et $\left|\frac{1}{z}\right| = \frac{1}{|z|} = 1$ donc $zz' \in \mathbb{U}$ et $\frac{1}{z} \in \mathbb{U}$. ■

Remarque

On aurait également pu utiliser la surjectivité de l'application exponentielle et que, pour tous θ et $\varphi \in \mathbb{R}$,

$$e^{i\theta}.e^{i\varphi} = e^{i(\theta+\varphi)} \in \mathbb{U}$$
$$\left(e^{i\theta}\right)^{-1} = e^{-i\theta} \in \mathbb{U}$$

Exemple

Soit a et b deux complexes de modules distincts et $z \in \mathbb{U}$. Montrons que $\frac{az+b}{\overline{b}z+\overline{a}} \in \mathbb{U}$. En effet, en écrivant que $\overline{a} = \overline{a}|z|^2$ puis $|z|^2 = z\overline{z}$,

$$\begin{aligned}|\overline{b}z+\overline{a}| &= |\overline{b}z+\overline{a}|z|^2| = |z|.|\overline{b}+\overline{a}z| \\ &= |\overline{b+az}| = |az+b|.\end{aligned}$$

2.4. Arguments d'un nombre complexe

Nous avons vu que l'application exponentielle complexe qui, à un réel, associe un complexe de module 1 est surjective. Cela autorise la définition suivante.

Définition 2.25.

Soit $z \in \mathbb{C}$ non nul. L'ensemble des arguments de z est l'ensemble des antécédents de $\frac{z}{|z|}$ par l'application exponentielle complexe qui, à un réel, associe un complexe de module 1.

Autrement dit, un réel θ est un argument de z si $z = |z|e^{i\theta}$ (cette écriture est appelée trigonométrique ou géométrique de z par opposition à l'écriture cartésienne en termes de parties réelle et imaginaire).

Définition 2.26.

L'exponentielle d'un complexe $z = x + iy$ avec $x, y \in \mathbb{R}$ est le complexe $e^x e^{iy}$.

Cette définition étend évidemment la propriété de morphisme.

Proposition 2.27.

L'application $\left\{\begin{array}{ccc} \mathbb{C} & \to & \mathbb{C}^* \\ z & \mapsto & e^z \end{array}\right.$ est surjective mais pas injective.

Démonstration

Pour obtenir un antécédent $x + iy$ de $a \in \mathbb{C}^*$, il suffit de poser $x = \ln|a|$ et y un argument de a. ■

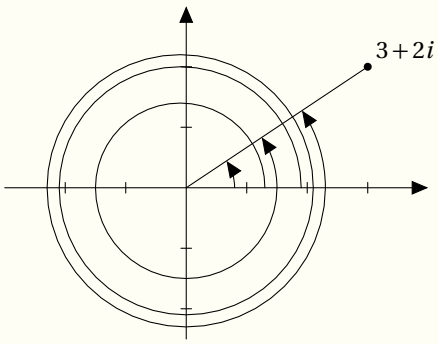
Proposition 2.28.

L'ensemble des arguments d'un complexe non nul est de la forme $\theta + 2\pi\mathbb{Z} = \{\theta + 2k\pi,\ k \in \mathbb{Z}\}$.

Démonstration

Deux réels θ et φ sont des arguments d'un même complexe si, et seulement si, $e^{i\theta} = e^{i\varphi}$ donc, par propriété de morphisme, $e^{i(\theta-\varphi)} = 1$. Cette condition équivaut à $\cos(\theta-\varphi) = 1$ et $\sin(\theta-\varphi) = 0$, c'est-à-dire à $\theta - \varphi \in 2\pi\mathbb{Z}$. ■

Voici une représentation de différents arguments du complexe $3+2i$ séparés par 2π et 4π.



Définition 2.29.

L'argument principal d'un nombre complexe z non nul est le plus petit argument positif de z. Il est souvent noté $\arg(z)$.

Remarque

D'après la proposition précédente, l'argument principal d'un complexe appartient à $[0, 2\pi[$. Un réel est un complexe d'argument principal 0 ou π.

Proposition 2.30.

Soit z_1, z_2 deux complexes non nuls.
▹ Un argument de $z_1 z_2$ est obtenu en additionnant des arguments de z_1 et de z_2.
▹ Un argument de $\frac{z_1}{z_2}$ est obtenu en soustrayant un argument de z_2 à un argument de z_1.
▹ Un argument de $\overline{z_1}$ est l'opposé d'un argument de z_1.

Démonstration

Notons $z_1 = |z_1|e^{i\theta_1}$ et $z_2 = |z_2|e^{i\theta_2}$. Alors,

$$z_1 z_2 = |z_1|.|z_2|e^{i(\theta_1+\theta_2)}, \qquad \frac{z_1}{z_2} = \frac{|z_1|}{|z_2|}e^{i(\theta_1-\theta_2)},$$

d'où l'on déduit les deux premiers points. La dernière affirmation provient de l'égalité $\overline{z_1} = \frac{|z_1|^2}{z_1}$ et de la propriété pour le quotient. ■

2.5. Interprétation géométrique

Comme un argument d'un nombre complexe z s'interprète comme l'angle entre l'axe des abscisses et la demi-droite issue de l'origine et dirigée par le vecteur d'affixe z, on obtient directement

Proposition 2.31.

Soit a, b, c des complexes deux à deux distincts. L'argument du complexe $\frac{c-a}{b-a}$ est une mesure de l'angle orienté $(\overrightarrow{AB}, \overrightarrow{AC})$ où A, B, C sont les points d'affixe a, b et c respectivement.

Exemple Angle au centre

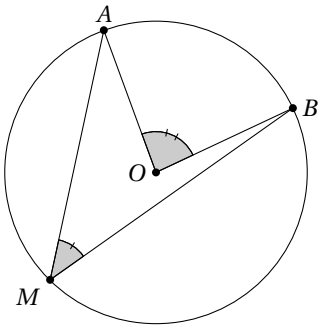
Soit α, β, $\theta \in \mathbb{R}$. Notons A, B, M, O les points d'affixes $e^{i\alpha}$, $e^{i\beta}$, $e^{i\theta}$ et 0 et montrons que $2(\overrightarrow{MB}, \overrightarrow{MA}) = (\overrightarrow{OB}, \overrightarrow{OA})[2\pi]$. En effet, en factorisant par l'exponentielle de la demi-somme des arguments,

$$\frac{e^{i\alpha}-e^{i\theta}}{e^{i\beta}-e^{i\theta}} = \frac{e^{i\frac{\alpha+\theta}{2}}(e^{i\frac{\alpha-\theta}{2}}-e^{-i\frac{\alpha-\theta}{2}})}{e^{i\frac{\beta+\theta}{2}}(e^{i\frac{\beta-\theta}{2}}-e^{-i\frac{\beta-\theta}{2}})} = \frac{e^{i\frac{\alpha+\theta}{2}}\sin\frac{\alpha-\theta}{2}}{e^{i\frac{\beta+\theta}{2}}\sin\frac{\beta-\theta}{2}} = e^{i\frac{\alpha-\beta}{2}}\frac{\sin\frac{\alpha-\theta}{2}}{\sin\frac{\beta-\theta}{2}}.$$

Par conséquent, un argument de $\frac{e^{i\alpha}-e^{i\theta}}{e^{i\beta}-e^{i\theta}}$, ou de manière équivalente une mesure de l'angle $(\overrightarrow{MA}, \overrightarrow{MB})$, est égale à $\frac{\alpha-\beta}{2}[\pi]$ (l'angle est exprimé modulo π et non modulo 2π puisque l'on ne connaît pas le signe du réel $\frac{\sin\frac{\alpha-\theta}{2}}{\sin\frac{\beta-\theta}{2}}$). Par conséquent,

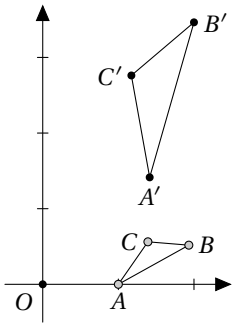
$$2(\overrightarrow{MB}, \overrightarrow{MA}) = \alpha-\beta[2\pi] = (\overrightarrow{OB}, \overrightarrow{OA})[2\pi].$$

Voici une illustration de cette propriété d'angle au centre.



Exemple Similitudes directes

Soit $a \neq 0$ et b deux complexes.
▹ L'application $z \mapsto az$ est la composition (commutative) de l'homothétie de centre l'origine et de rapport $|a|$ et de la rotation de centre l'origine et d'angle $\arg(a)$.
Voici, par exemple, l'effet de la transformation $z \mapsto 2e^{i\frac{\pi}{4}}$ sur un triangle ABC : une mesure de l'angle $(\overrightarrow{OA}, \overrightarrow{OA'})$ vaut $\frac{\pi}{4}$ et le rapport de longueur $\frac{OA'}{OA}$ vaut 2.



On note, en particulier, que la rotation de centre l'origine et d'angle $\frac{\pi}{2}$ est simplement la multiplication par i.
▹ Pour étudier l'application $z \mapsto az+b$ (pour $a \neq 1$ car dans le cas contraire c'est une simple translation), on se ramène au cas précédent en changeant de repère. Considérons $z_0 = \frac{b}{1-a}$ l'unique point fixe de l'application. Alors, $z' = az+b$ si, et seulement si, $z'-z_0 = a(z-z_0)$ donc le point d'affixe z' se déduit du point d'affixe z par l'homothétie de centre z_0 et de rapport $|a|$ et de la rotation de centre z_0 et d'angle $\arg(a)$.

Proposition 2.32.

Le produit scalaire de deux vecteurs d'affixes z et z' est $\mathfrak{Re}(\overline{z}z') = \frac{1}{2}(\overline{z}z' + z\overline{z'})$.

Démonstration

Notons $z = |z|e^{i\theta}$ et $z' = |z'|e^{i\theta'}$. Alors,

$$\overline{z}z' = |z||z'|e^{i(\theta'-\theta)} = |z||z'|\cos(\theta'-\theta) + i|z||z'|\sin(\theta'-\theta),$$

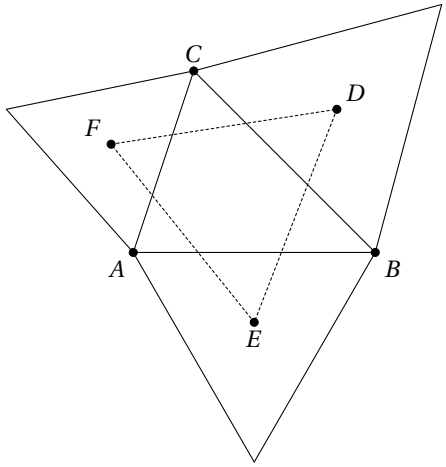
et l'on reconnaît l'expression du produit scalaire. ■

Remarque

La partie imaginaire de $\overline{z}z'$ est appelée déterminant des vecteurs d'affixes z et z'. La nullité de cette quantité sert à caractériser la condition de colinéarité. En effet, pour $z \neq 0$, $\mathfrak{Im}(\overline{z}z') = 0$ si, et seulement s'il existe $a \in \mathbb{R}$ tel que $\overline{z}z' = a$ soit $z' = \frac{a}{|z|^2}z$.

Exemple

Considérons un triangle ABC et D, E et F les centres des triangles équilatéraux construits (extérieurement) sur les côtés de ce triangle comme indiqué sur la figure suivante :



Notons a, b, c, d, e et f les affixes des points. D'après les propriétés des triangles équilatéraux,

$$(d-b) = j(d-c), \qquad (f-c) = j(f-a), \qquad (e-a) = j(e-b)$$

soit

$$(1-j)d = b - jc, \qquad (1-j)f = c - ja, \qquad (1-j)e = a - jb.$$

Ainsi,

$$(1-j)(e-d) = a - jb - b + jc = -j(1-j)(e-f)$$

donc $e-d = e^{-i\frac{\pi}{3}}(e-f)$. En conclusion, DEF est équilatéral. Ce résultat est improprement appelé *théorème de Napoléon*.

Exemple Symétrie orthogonale

Soit A, B deux points d'affixes a et b telles que $|a| = |b|$. Alors, le symétrique orthogonal du point M d'affixe z par rapport à la droite (AB) est le point M' d'affixe

$$z' = a + b - \frac{1}{|a|^2} ab\overline{z}.$$

En effet, vérifions que la droite (AB) est la médiatrice du segment $[MM']$ avec une condition d'orthogonalité (donc un produit scalaire nul) et une condition de distance :

$$\begin{aligned}\mathfrak{Re}((z'-z)\overline{(b-a)}) &= \mathfrak{Re}\left(\left(-\tfrac{1}{|a|^2} ab\overline{z} - z\right)\overline{(b-a)}\right)\\ &= -\mathfrak{Re}(a\overline{z} - \overline{a}z + \overline{b}z - b\overline{z}) = 0,\\ |z'-a| &= \frac{|b|}{|a|^2}\left||a|^2 - a\overline{z}\right| = |\overline{a} - \overline{z}| = |z-a|.\end{aligned}$$

Dans le premier calcul, on a utilisé que

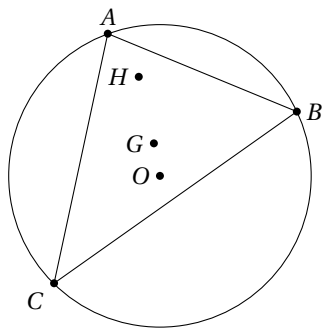
$$\mathfrak{Re}((a+b)\overline{(a-b)}) = \mathfrak{Re}(|a|^2 - |b|^2 + 2i\mathfrak{Im}(b\overline{a})).$$

Exemple Points remarquables dans un triangle

Soit A, B, C trois points non-alignés d'affixes respectives a, b et c tels que O, le centre du cercle circonscrit au triangle ABC, ait une affixe nulle (et donc $|a| = |b| = |c|$). Montrons que l'orthocentre H a pour affixe $a+b+c$; pour cela, il suffit de vérifier que le vecteur $\overrightarrow{AH}$ est orthogonal au vecteur $\overrightarrow{BC}$ (on a les autres propriétés de hauteur par symétrie)

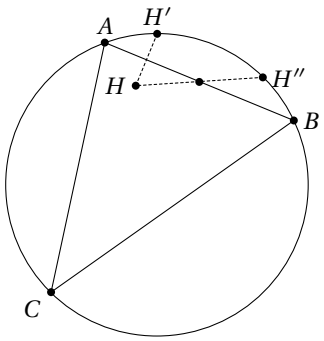
$$\mathfrak{Re}((a+b+c-a)\overline{(c-b)}) = \mathfrak{Re}((b+c)\overline{(c-b)}) = 0.$$

Comme le centre de gravité G admet pour affixe $\frac{1}{3}(a+b+c)$, on retrouve avec ce calcul que $\overrightarrow{OH} = 3\overrightarrow{OG}$.



▹ En utilisant les calculs de l'exemple précédent, on trouve que le symétrique orthogonal H' de H par rapport à un côté AB admet pour affixe $-\frac{ab\overline{c}}{|a|^2}$ donc appartient au cercle circonscrit (de centre O et de rayon $|a| = |b| = |c|$).

▹ Soit z'' l'affixe du symétrique H'' de H par rapport au milieu du segment $[AB]$. Alors, $\frac{z''+a+b+c}{2} = \frac{a+b}{2}$ donc $z'' = -c$ et le point d'affixe z appartient lui-aussi au cercle circonscrit.



Parmi les autres applications, il peut être utile de savoir reconnaître l'équation complexe d'une droite.

Proposition 2.33.

Une équation complexe d'une droite du plan est de la forme $\overline{a}z + a\overline{z} + b = 0$ où $a \in \mathbb{C}^*$ et $b \in \mathbb{R}$.

Démonstration

Soit $\mathscr{D}$ une droite du plan complexe. Notons a l'affixe d'un vecteur normal à $\mathscr{D}$ et z_0 l'affixe d'un point de $\mathscr{D}$. Alors, un point d'affixe z appartient à $\mathscr{D}$ si, et seulement si, $\mathfrak{Re}(\overline{a}(z-z_0)) = 0$, c'est-à-dire si

$$\overline{a}(z-z_0) + a\overline{(z-z_0)} = 0$$

c'est-à-dire

$$\overline{a}z + a\overline{z} = \overline{a}z_0 + a\overline{z_0}.$$

On conclut en posant $b = -(\overline{a}z_0 + a\overline{z_0})$. ■

Remarque

Pour une droite verticale, un vecteur est normal si son affixe a est réelle (non nulle) ; l'équation de la droite devient alors $a\mathfrak{Re}(z) + b = 0$ c'est-à-dire $\mathfrak{Re}(z) = -\frac{b}{a}$.

Exemple

Soit M_0 un point d'affixe z_0 et $\overrightarrow{u}$ un vecteur non nul d'affixe b. Calculons une équation complexe de la droite passant par M_0 et dirigée par $\overrightarrow{u}$. Un point M d'affixe z appartient à cette droite si, et seulement si, les vecteurs $\overrightarrow{MM_0}$ et $\overrightarrow{u}$ sont colinéaires, c'est-à-dire $\mathfrak{Im}((z-z_0)\overline{b}) = 0$ que l'on peut réécrire sous la forme $z\overline{b} - b\overline{z} = z_0\overline{b} - b\overline{z_0}$.

3. Équations algébriques

3.1. Équations de degré 2

On commence par se limiter à l'équation de degré 2 pour laquelle les résultats et les méthodes ressemblent à ce qui est déjà connu sur $\mathbb{R}$.

Proposition 2.34.

Pour tout $a \in \mathbb{C}^*$, il existe deux solutions (opposées) à l'équation $z^2 = a$.

Donnons deux preuves de ce résultat.

Démonstration

Écrivons $a = Re^{i\varphi}$ avec $R \in \mathbb{R}_+^*$ et $\varphi \in \mathbb{R}$. Cherchons z sous la forme $re^{i\theta}$ avec $r \in \mathbb{R}_+$ et $\theta \in \mathbb{R}$. L'équation $z^2 = a$ se traduit $r^2 e^{2i\theta} = Re^{i\varphi}$ soit $r^2 = R$ et $2\theta = \varphi[2\pi]$. Les solutions sont donc

$$\sqrt{R}e^{i\frac{\varphi}{2}}, \qquad \sqrt{R}e^{i\left(\frac{\varphi}{2}+\pi\right)} = -\sqrt{R}e^{i\frac{\varphi}{2}}.$$

■

Malheureusement, en pratique, cette démonstration ne donne pas toujours le moyen d'obtenir les racines puisqu'il faut disposer de l'écriture géométrique de a. Voici avec la seconde preuve le moyen d'obtenir les racines lorsque l'on ne connaît que l'écriture cartésienne de a.

Démonstration

Identifions les parties imaginaires et réelles dans l'équation $(x+iy)^2 = \alpha + i\beta$; on obtient le système non-linéaire

$$\begin{cases} x^2 - y^2 & = \alpha \\ 2xy & = \beta \end{cases}$$

Pour simplifier la résolution de celui-ci, on ajoute la condition sur le module $|x+iy|^2 = |\alpha+i\beta|$ qui s'écrit $x^2+y^2 = \sqrt{\alpha^2+\beta^2}$. On obtient ainsi

$$\begin{cases} x^2 & = \frac{\sqrt{\alpha^2+\beta^2}+\alpha}{2} \\ y^2 & = \frac{\sqrt{\alpha^2+\beta^2}-\alpha}{2} \\ 2xy & = \beta \end{cases}$$

ce qui permet de trouver rapidement les deux valeurs possibles de (x, y)

$$\left(\sqrt{\frac{\sqrt{\alpha^2+\beta^2}+\alpha}{2}}, \frac{\beta}{\sqrt{2(\sqrt{\alpha^2+\beta^2}+\alpha)}}\right) \text{ et } \left(-\sqrt{\frac{\sqrt{\alpha^2+\beta^2}+\alpha}{2}}, -\frac{\beta}{\sqrt{2(\sqrt{\alpha^2+\beta^2}+\alpha)}}\right).$$

Les premières équations donnent $|x|$ et $|y|$, la dernière nous donne le signe de x en fonction de celui de y.

■

Remarque

L'écriture $\sqrt{z}$ n'a aucun sens pour $z \notin \mathbb{R}_+$, aucune des deux « racines carrées » de z n'est privilégiée.

Exemple

Cherchons les complexes $z = x+iy$ avec $x, y \in \mathbb{R}$ tels que $z^2 = 3-4i$. Avec la méthode précédente, on obtient $x = \pm 2$, $y = \pm 1$ et $2xy = -4$ donc les solutions sont $z = 2-i$ et $z = -2+i$.

Énonçons désormais la résolution complète de l'équation de degré 2 (qui généralise les résultats connus sur $\mathbb{R}$)

Proposition 2.35.

Soit a, b, c trois complexes tels que $a \neq 0$. Notons $\Delta = b^2 - 4ac$ le discriminant.

- Si $\Delta = 0$, l'équation $az^2 + bz + c = 0$ admet une unique solution $-\frac{b}{2a}$.
- Si $\Delta \neq 0$, l'équation $az^2 + bz + c = 0$ admet deux solutions distinctes $\frac{-b \pm \delta}{2a}$ où δ est l'une des solutions de $z^2 = \Delta$.

Démonstration

Comme dans le cas réel, on se ramène à une écriture canonique et on utilise une identité remarquable

$$\begin{aligned} az^2 + bz + c &= a\left(z^2 + \frac{b}{a}z + \frac{c}{a}\right) \\ &= a\left(\left(z + \frac{b}{2a}\right)^2 + \frac{c}{a} - \frac{b^2}{4a^2}\right) = a\left(\left(z + \frac{b}{2a}\right)^2 - \frac{\Delta}{4a^2}\right) \\ &= a\left(z - \frac{-b+\delta}{2a}\right)\left(z - \frac{-b-\delta}{2a}\right) \end{aligned}$$

■

Exemple

Cherchons les solutions de $z^2 - 2\cos\theta\, z + 1 = 0$ pour $\theta \in \mathbb{R}$.
Le discriminant est alors $4(\cos^2\theta - 1) = -4\sin^2\theta$ donc les solutions sont

$$\cos\theta \pm i\sin\theta = e^{\pm i\theta}.$$

Exemple

Cherchons les solutions de $z^2 - 2e^{i\theta}z + 1 = 0$ pour $\theta \in [0, \pi]$.
Le discriminant est $\Delta = 4(e^{2i\theta} - 1) = 8e^{i(\theta + \frac{\pi}{2})}\sin\theta$. Les solutions de l'équation $z^2 = \Delta$ sont $\pm 2\sqrt{2\sin\theta}\, e^{i(\frac{\theta}{2} + \frac{\pi}{4})}$ donc les solutions recherchées sont

$$e^{i\theta} \pm \sqrt{2\sin\theta}\, e^{i(\frac{\theta}{2} + \frac{\pi}{4})}.$$

Exemple

Cherchons les solutions de $z^2 + 2iz - 4 + 4i = 0$.
Le discriminant est $\Delta = 12 - 16i$ dont une racine est $\delta = 4 - 2i$. Les solutions sont donc $\frac{-2i-(4-2i)}{2} = -2$ et $\frac{-2i+(4-2i)}{2} = 2 - 2i$.

Remarque

Si l'on écrit $az^2 + bz + c = a(z - z_1)(z - z_2)$ et que l'on développe le second membre, on obtient

$$z_1 + z_2 = -\frac{b}{a}, \qquad z_1 z_2 = \frac{c}{a}.$$

Ceci est un cas particulier des relations entre coefficients et racines pour un polynôme : c'est particulièrement pratique lorsque que l'on connaît une racine pour en déduire la seconde. Ces relations se généralisent à des équations de plus haut degré.

Remarque

Si $z \in \mathbb{C}$ est solution d'une équation à coefficients réels, alors $\overline{z}$ est aussi solution de cette équation (il suffit de conjuguer l'équation).

3.2. Racines $n^{\text{ièmes}}$

Dans cette section, on se limite à un exemple important : celui de l'équation $z^n = a$ avec $a \in \mathbb{C}$.

Définition 2.36.

Soit $n \in \mathbb{N} \setminus \{0\}$. L'ensemble des racines $n^{\text{ièmes}}$ de l'unité, noté $\mathbb{U}_n$, est l'ensemble des complexes z tels que $z^n = 1$.

Proposition 2.37.

L'ensemble $\mathbb{U}_n$ est un groupe (à n éléments) abélien pour la multiplication.

Démonstration

La loi $\times$ est associative et commutative sur $\mathbb{U}_n$ car elle l'est sur $\mathbb{C}$; l'élément neutre 1 appartient à $\mathbb{U}_n$. Pour tout $z, z' \in \mathbb{U}_n$,

$$(zz')^n = z^n z'^n = 1, \qquad \left(\frac{1}{z}\right)^n = \frac{1}{z^n} = 1.$$

donc $zz' \in \mathbb{U}_n$ et $\frac{1}{z} \in \mathbb{U}_n$. Ainsi, $\mathbb{U}_n$ est un groupe. ■

Remarque

L'ensemble $\mathbb{U}_n$ est stable par passage au conjugué car $z^n = 1$ si, et seulement si, $\overline{z}^n = 1$. Ici, cette propriété peut se retrouver également par le fait que $\mathbb{U}_n \subset \mathbb{U}$ donc que, pour tout $z \in \mathbb{U}_n$,

$$\frac{1}{z} = \frac{\overline{z}}{|z|^2} = \overline{z}.$$

Il s'agit toutefois d'une propriété plus générale, si z est racine d'une équation polynomiale à coefficients réels, alors $\overline{z}$ l'est également (en conjugant).

Proposition 2.38.

Pour tout $n \in \mathbb{N} \setminus \{0\}$,

$$\mathbb{U}_n = \left\{e^{\frac{2ik\pi}{n}},\ k \in \mathbb{Z}\right\} = \left\{e^{\frac{2ik\pi}{n}},\ k \in [\![0, n-1]\!]\right\}.$$

Démonstration

▹ Un complexe $z = |z|e^{i\theta}$ appartient à $\mathbb{U}_n$ si, et seulement si, $|z|^n = 1$ et $n\theta \in 2\pi\mathbb{Z}$. Comme $|z| \in \mathbb{R}_+$, $|z| = 1$. D'où le résultat.

▹ Pour tout $k \in \mathbb{Z}$, il existe $(q, r) \in \mathbb{Z} \times [\![0, n-1]\!]$ (quotient et reste de la division euclidienne de k par n) tel que $k = qn + r$. Ainsi,

$$e^{\frac{2ik\pi}{n}} = e^{\frac{2ir\pi}{n}} \in \left\{e^{\frac{2ik\pi}{n}},\ k \in [\![0, n-1]\!]\right\}.$$

Donc

$$\left\{e^{\frac{2ik\pi}{n}},\ k \in \mathbb{Z}\right\} \subset \left\{e^{\frac{2ik\pi}{n}},\ k \in [\![0, n-1]\!]\right\}.$$

L'autre inclusion est évidente et l'on vérifie que ces n complexes sont deux à deux distincts (car les arguments principaux sont distincts). ■

Exemple

Commençons par quelques petites valeurs

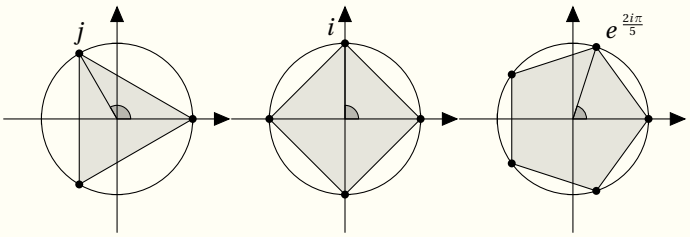
- $\mathbb{U}_2 = \{\pm 1\}$.
- $\mathbb{U}_3 = \{1, j, j^2\} = \{1, j, \overline{j}\}$ où $j = e^{\frac{2i\pi}{3}}$ (et donc $\overline{j} = j^2$).
- $\mathbb{U}_4 = \{\pm 1, \pm i\}$.

Remarque

Les points d'affixes les racines $n^{\text{ièmes}}$ de l'unité forment un polygone régulier à n sommets, inscrit dans le cercle unité et qui admet $(1,0)$ pour sommet. En effet, l'angle au centre défini par deux racines $n^{\text{ièmes}}$ de l'unité consécutives est un argument de

$$\frac{e^{\frac{2i(k+1)\pi}{n}} - 0}{e^{\frac{2ik\pi}{n}} - 0} = e^{\frac{2i\pi}{n}}$$

donc de mesure $\frac{2\pi}{n}$.
Voici une illustration pour $n = 3$, $n = 4$ et $n = 5$:



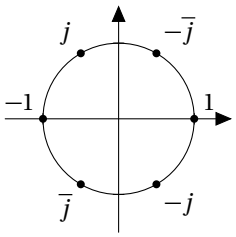
On étudie dans le chapitre suivant d'autres propriétés géométriques des racines $n^{\text{ièmes}}$ de l'unité.

Exemple

Pour $n = 6$, les racines $6^{\text{ièmes}}$ de l'unité sont successivement

$$e^{\frac{2i0\pi}{6}} = 1,\ e^{\frac{2i1\pi}{6}} = e^{\frac{i\pi}{3}} = -j^2,\ e^{\frac{2i2\pi}{6}} = e^{\frac{2i\pi}{3}}\ j,\ e^{\frac{2i3\pi}{6}} = -1,\ e^{\frac{2i4\pi}{6}} = e^{\frac{4i\pi}{3}} = j^2,\ e^{\frac{2i5\pi}{6}} = e^{\frac{5i\pi}{3}} = -j.$$

On obtient donc le dessin suivant des points correspondants à ces affixes



Proposition 2.39.

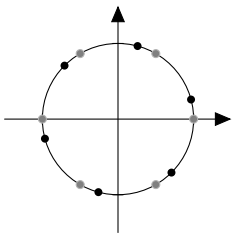
Soit $a \in \mathbb{C}^*$. Les solutions de $z^n = a$ sont les complexes $|a|^{\frac{1}{n}} e^{i\frac{\arg(a)+2k\pi}{n}}$ pour $k \in [\![0, n-1]\!]$.

En d'autres termes, pour obtenir les racines $n^{\text{ièmes}}$ de a, on commence par en trouver une z_0, à savoir $|a|^{\frac{1}{n}} e^{i\frac{\arg(a)}{n}}$ avec les notations de la proposition, puis on déduit toutes les autres en multipliant z_0 par une racine $n^{\text{ième}}$ de l'unité.

Démonstration
Notons $z' = \frac{z}{|a|^{\frac{1}{n}} e^{i\frac{\arg(a)}{n}}}$. Alors, $z^n = a(z')^n$ donc z est solution de $z^n = a$ si, et seulement si, z' est une racine $n^{\text{ième}}$ de l'unité. ■

Exemple

Les solutions de $z^n = i$ sont les complexes de la forme $e^{\frac{2ki\pi}{n}} \times e^{\frac{i\pi}{2n}}$ avec $k \in [\![1, n]\!]$.
Voici, en noir, le dessin des points correspondants à ces affixes pour $n = 6$ (les points grisés correspondent aux points d'affixes dans $\mathbb{U}_6$) :



On remarque que l'on a « tourné » la figure pour $\mathbb{U}_6$ d'un angle correspondant à l'une des racines $6^{\text{ièmes}}$ de i, à savoir $\frac{\pi}{12}$.

3.3. Cas général

Citons un théorème essentiel (certains l'appellent même *théorème fondamental de l'algèbre*).

Théorème 2.40. D'Alembert-Gauss

Tout polynôme non-constant à coefficients dans $\mathbb{C}$ admet une racine.

Il existe de nombreuses preuves de ce résultat mais aucune n'est accessible à ce stade du cours.

Corollaire 2.41.

Un polynôme non constant définit une application surjective de $\mathbb{C}$ dans $\mathbb{C}$.

Démonstration
Soit $P \in \mathbb{C}[X]$ un polynôme non constant et $z \in \mathbb{C}$. Le polynôme $P - z$ est non constant donc admet une racine $a \in \mathbb{C}$ d'après le théorème de D'Alembert-Gauss. Par définition, $P(a) = z$ donc le complexe z admet un antécédent. Ainsi, tout complexe appartient à l'image de P et $P : \mathbb{C} \to \mathbb{C}$ est surjectif. ■

FICHE

SYNTHÈSE

L'écriture naturelle pour un nombre complexe est l'écriture cartésienne définie avec deux réels.

Écriture cartésienne

▹ Un complexe z admet une unique écriture $x+iy$ avec $x,\ y\in\mathbb{R}$; x est la partie réelle et y est la partie imaginaire de z. On les note $\mathfrak{Re}(z)$ et $\mathfrak{Im}(z)$.
▹ Le conjugué de $z=x+iy$ est $\overline{z}=x-iy$.
▹ Les applications parties réelle, imaginaire et l'application conjugué sont $\mathbb{R}$-linéaires.

Dès qu'il s'agit d'effectuer des calculs avec des produits et des quotients, il convient de préférer l'écriture géométrique.

Écriture géométrique

Un complexe z non nul admet une unique écriture $re^{i\theta}$ avec $r>0$ et $\theta\in[0,2\pi[$; r est le module de z et θ son argument principal. On les note $|z|$ et $\arg(z)$.
▹ Pour tout $z\in\mathbb{C}$, $|z|^2=z\overline{z}$.
▹ Pour tous z et $z'\in\mathbb{C}$, $|zz'|=|z||z'|$ et $\arg(zz')=\arg(z)+\arg(z')[2\pi]$.
▹ Pour tout $z\neq 0$, $\left|\frac{1}{z}\right|=\frac{1}{|z|}$ et $\arg(\frac{1}{z})=-\arg(z)[2\pi]$.

Ces derniers résultats sur le module et l'argument sont des conséquences des propriétés de l'exponentielle complexe.

Exponentielle complexe

L'application $\theta\mapsto e^{i\theta}=\cos\theta+i\sin\theta$ est surjective de $\mathbb{R}$ dans $\mathbb{U}$ (non injective).
▹ (formule d'Euler) Pour tous $\theta\in\mathbb{R}$,

$$\cos\theta=\frac{e^{i\theta}+e^{-i\theta}}{2},\qquad \sin\theta=\frac{e^{i\theta}-e^{-i\theta}}{2i}.$$

▹ (propriété de morphisme) Pour tous $\theta,\ \varphi\in\mathbb{R}$, $e^{i\theta}e^{i\varphi}=e^{i(\theta+\varphi)}$.
▹ (formule de Moivre) Pour tout $n\in\mathbb{N}$ et tout $\theta\in\mathbb{R}$, $(e^{i\theta})^n=e^{in\theta}$.

Parmi les ensembles remarquables de nombres complexes, on remarque quelques groupes.

$\mathbb{U}$, $\mathbb{U}_n$

L'ensemble $\mathbb{U}$ des complexes de module 1 et l'ensemble $\mathbb{U}_n$ des complexes dont la puissance $n^{\text{ième}}$ est égale à 1 sont des groupes pour la multiplication.

$$\mathbb{U}_n = \{e^{\frac{2ik\pi}{n}},\ k \in \mathbb{Z}\} = \{e^{\frac{2ik\pi}{n}},\ 0 \leq k \leq n-1\}.$$

L'ensemble des solutions de $z^n = re^{i\theta}$ avec $r > 0$ et $\theta \in \mathbb{R}$ fixés est

$$\{r^{\frac{1}{n}} e^{i\frac{\theta+2k\pi}{n}},\ 0 \leq k \leq n-1\} = r^{\frac{1}{n}} e^{i\frac{\theta}{n}} \mathbb{U}_n.$$

Les nombres complexes peuvent aussi être utiles aux calculs en géométrie plane. Pour cela, on identifie un point M de coordonnées (x, y) et son affixe, le complexe égal à $x + iy$.

Application à la géométrie

Soit A, B, C, D des points du plan d'affixes a, b, c et $d \in \mathbb{C}$ respectivement.

▹ La distance AB est $|b-a|$.

▹ Une détermination de l'angle $(\overrightarrow{AB}, \overrightarrow{AC})$ est $\arg\left(\frac{c-a}{b-a}\right)$.

▹ Le produit scalaire $\overrightarrow{AB}.\overrightarrow{CD}$ est $\mathfrak{Re}((\overline{b-a})(d-c))$.

De nombreuses formules trigonométriques permettent de mener à bien les calculs (en linéarisant ou en factorisant les expressions). On peut les retrouver à partir des formules d'addition suivantes.

Formules d'addition

Pour tous θ, $\varphi \in \mathbb{R}$,

$$\sin(\theta+\varphi) = \sin(\theta)\cos(\varphi) + \cos(\theta)\sin(\varphi),$$
$$\cos(\theta+\varphi) = \cos(\theta)\cos(\varphi) - \sin(\theta)\sin(\varphi),$$
$$\tan(\theta+\varphi) = \frac{\tan(\theta)+\tan(\varphi)}{1-\tan(\theta)\tan(\varphi)}.$$

avec les restrictions θ, φ, $\theta+\varphi \in \mathbb{R} \setminus \{\frac{\pi}{2} + k\pi,\ k \in \mathbb{Z}\}$ pour la dernière égalité.

EXERCICES

Vrai ou faux ?

		Vrai	Faux
a)	La partie réelle de $\left(\frac{1}{2+i}\right)^2$ est $\frac{1}{4}$.	☐	☐
b)	Un argument de $e^{i\frac{\pi}{6}}+e^{i\frac{9\pi}{6}}$ est $\frac{5\pi}{6}$.	☐	☐
c)	L'exponentielle est surjective de $\mathbb{C}$ dans $\mathbb{C}^*$.	☐	☐
d)	Soit $\theta \notin 2\pi\mathbb{Z}$, $z \mapsto e^{i\theta}z+1$ est une écriture complexe de la rotation de centre d'affixe 1 et d'angle θ.	☐	☐
e)	Les deux racines non réelles d'un polynôme de degré 2 à coefficients complexes sont conjuguées.	☐	☐
f)	Le complexe $-j$ est racine de X^2-X+1.	☐	☐
g)	Pour tout $k \in \mathbb{N}\setminus\{0\}$, j^k est racine de X^2+X+1.	☐	☐
h)	Si j est racine d'un polynôme réel P, alors j^2 est aussi racine de P.	☐	☐
i)	Si m divise n, alors $\mathbb{U}_m \subset \mathbb{U}_n$.	☐	☐
j)	Les ensembles $\mathbb{U}_n$ et $\{\omega,\ \omega^n=-1\}$ sont en bijection.	☐	☐
k)	$\{\overline{\omega},\ \omega\in\mathbb{U}_n\}=\mathbb{U}_n$.	☐	☐
l)	$\{\omega^2,\ \omega\in\mathbb{U}_n\}=\mathbb{U}_n$ pour n impair.	☐	☐
m)	Un polynôme complexe qui prend deux valeurs distinctes réalise une surjection de $\mathbb{C}$ dans $\mathbb{C}$.	☐	☐
n)	Une fonction polynomiale injective sur $\mathbb{C}$ est de degré 1.	☐	☐

Exercices d'application du chapitre 2

●○○ **Exercice 1**

Soit $n\in\mathbb{N}$ et $\theta\in\mathbb{R}\setminus\{\frac{\pi}{2}+k\pi,\ k\in\mathbb{Z}\}$. Calculer les complexes suivants $(i-\sqrt{3})^n$ et $(1+i\tan\theta)^n$.

●○○ **Exercice 2**

Déterminer la partie réelle de $\left(\frac{\sqrt{3}-i}{1+i}\right)^{15}$.

●○○ **Exercice 3**

Déterminer les entiers $n \in \mathbb{N}$ tels que $(1+i)^n \in i\mathbb{R}$.

●○○ **Exercice 4**

Résoudre l'équation $(1+i)z^2 - z - 1 + i = 0$.

●○○ **Exercice 5**

Soit $\theta \in \mathbb{R}$. Résoudre l'équation $z^4 - 2\cos(\theta)z^2 + 1 = 0$.

●○○ **Exercice 6**

Résoudre l'équation $z^2 - 2z + i = 0$.

●●○ **Exercice 7**

Soit $\alpha \in \mathbb{C} \setminus \mathbb{R}$ et $n \in \mathbb{N} \setminus \{0\}$. Montrer que les complexes z vérifiant l'équation $\left(\frac{1+iz}{1-iz}\right)^n = \alpha$ sont tous réels si, et seulement si, $|\alpha| = 1$.

●●○ **Exercice 8**

Soit a, b, c et d des complexes de module 1. Montrer que $|ab - cd| \leq |a-c| + |b-d|$.

●●○ **Exercice 9**

Soit a, $b \in \mathbb{C}$. Montrer que

$$|a| + |b| \leq |a+b| + |a-b|.$$

Étudier le cas d'égalité puis interpréter cette inégalité pour les longueurs remarquables d'un parallélogramme.

●○○ **Exercice 10**

Soit $z \in \mathbb{U}$. Calculer $|z-1|^2 + |z+1|^2$ puis interpréter géométriquement le résultat obtenu.

●●○ **Exercice 11**

Soit $a, b \in \mathbb{U}$ distincts, $z \in \mathbb{C}$ et $u = \frac{z + ab\overline{z} - a - b}{a-b}$. Montrer que $u^2 \in \mathbb{R}$.

●○○ **Exercice 12**

Soit $a \in i\mathbb{R}$, $n \in \mathbb{N} \setminus \{0\}$, M_n, M_{n+1} et M_{n+2} les points d'affixe a^n, a^{n+1} et a^{n+2}. Montrer que le triangle $M_n M_{n+1} M_{n+2}$ est rectangle.

●○○ **Exercice 13**

Dessiner les trois zones du plan dont les points sont d'affixe z vérifiant

1. $1 \leq |z+1| \leq 2$.
2. $\mathfrak{Re}(z+1) \geq 0$.
3. $|iz+1| \leq 1$.

●●○ **Exercice 14**

Soit $a, b, c \in \mathbb{C}$ deux à deux distincts. Montrer que les points d'affixes a, b, c forment un triangle équilatéral si, et seulement si,

$$|2a - b - c| = |2b - a - c| = |2c - a - b|.$$

●●● **Exercice 15**

Soit a, b, $c \in \mathbb{C}$ deux à deux distincts. Montrer que les propositions suivantes sont équivalentes

i. les points d'affixes a, b, c forment un triangle équilatéral ;

ii. j ou j^2 est solution de $az^2 + bz + c = 0$;

iii. $a^2 + b^2 + c^2 = ab + ac + bc$;

iv. $\frac{1}{a-b} + \frac{1}{b-c} + \frac{1}{c-a} = 0$.

●●○ **Exercice 16**

Soit u, v, $w \in \mathbb{U}$ tels que $u + v + w = 0$. Montrer que $u = jv = j^2 w$ ou $u = jw = j^2 v$.

●●● **Exercice 17**

Déterminer les complexes $z_1, z_2, z_3 \in \mathbb{U}$ tels que $z_1 + z_2 + z_3 = z_1 z_2 z_3 = 1$.
On pourra remarquer que z_1, z_2, z_3 sont les solutions de $(z - z_1)(z - z_2)(z - z_3) = 0$ que l'on développera.

●●● **Exercice 18**

1. Soit a, $b \in \mathbb{C}$. Montrer que
$$|a - b|^2 \leq (1 + |a|^2)(1 + |b|^2).$$
2. Montrer qu'il y a égalité si, et seulement si, $a = re^{i\theta}$ et $b = -\frac{1}{r}e^{i\theta}$ avec $r \in \mathbb{R}^*$.

●●○ **Exercice 19**

Considérons l'application
$$f : \left\{ \begin{array}{ccc} \mathbb{C} \setminus \{-1\} & \to & \mathbb{C} \setminus \{0\} \\ z & \mapsto & \frac{2}{z+1} \end{array} \right.$$
Déterminer $f(\mathbb{U})$.

●●○ **Exercice 20**

Soit $P = \{z \in \mathbb{C},\ \mathfrak{Im}(z) > 0\}$ et $D = \{z \in \mathbb{C},\ |z| < 1\}$.

1. Soit $z \in P$. Montrer que $\frac{z-i}{z+i} \in D$.
2. Soit $z \neq -i$ tel que $\frac{z-i}{z+i} \in D$. Montrer que $z \in P$.
3. Montrer que l'application
$$f : \left\{ \begin{array}{ccc} P & \to & D \\ z & \mapsto & \frac{z-i}{z+i} \end{array} \right.$$
est une bijection puis calculer sa réciproque.

Exercices d'approfondissement du chapitre 2

●○○ **Exercice A**

Soit a, b et $c \in \mathbb{Z}$ tels que $ab + bc + ca = 1$. Montrer que $(1 + a^2)(1 + b^2)(1 + c^2)$ est un carré d'entier.

EXERCICES

●○○ **Exercice B**

Considérons, pour tout $\alpha \in \mathbb{R}$, l'équation (E_α) suivante

$$z^3 + 2z^2 - e^{i\alpha} z - 2e^{i\alpha} = 0.$$

1. Déterminer les solutions réelles de $(E_{\pi/2})$.
2. Montrer qu'une de ces solutions est solution de (E_α) pour tout $\alpha \in \mathbb{R}$.
3. En déduire l'ensemble des solutions de (E_α).
4. Déterminer les valeurs de $\frac{\alpha}{2}$ telles que les points dont les affixes sont les solutions de (E_α) forment un triangle rectangle.

●○○ **Exercice C**

Résoudre l'équation $z^3 + (i-2)z^2 + (3-3i)z + 2i - 2 = 0$ sachant qu'elle admet une solution réelle.

●●○ **Exercice D**

1. Soit $a, b \in \mathbb{C}$. Développer $(a + ib)^2$.
2. Soit $a, b \in \mathbb{C}$. Résoudre l'équation $z^2 = 4ab + 2i(a^2 - b^2)$.

●○○ **Exercice E**

Résoudre l'équation $1 - \overline{z} = |z|$.

●●○ **Exercice F**

Résoudre l'équation $z^n = \overline{z}$.

●○○ **Exercice G**

Soit $\omega \in \mathbb{U}_7$. Calculer

$$\frac{\omega}{1+\omega^2} + \frac{\omega^2}{1+\omega^4} + \frac{\omega^3}{1+\omega^6}.$$

On pourra admettre que $1 + \omega + \omega^2 + \omega^3 + \omega^4 + \omega^5 + \omega^6 = 0$ *(ce résultat sera établi dans le chapitre suivant).*

●●○ **Exercice H**

Soit $n \in \mathbb{N} \setminus \{0\}$. Résoudre l'équation $\mathfrak{Re}(z^n) = \mathfrak{Im}(z^n)$.

●●○ **Exercice I**

Montrer que, pour tout $z \in \mathbb{C}$,

$$|z-1| \leq |z-j| + |z-j^2|.$$

●○○ **Exercice J**

Soit z et z' deux complexes de module au plus 1. Montrer que

$$\min(|z-z'|, |z+z'|) \leq \sqrt{2}.$$

●●● **Exercice K**

Soit $u \in \mathbb{C} \setminus \{1\}$ et $z \in \mathbb{C} \setminus \mathbb{R}$. Montrer que

$$\frac{z - u\overline{z}}{1-u} \in \mathbb{R} \qquad \Leftrightarrow \qquad |u| = 1.$$

Problèmes du chapitre 2

●○○ **Problème 1**

Considérons la loi de composition interne $\boxtimes$ sur $\mathbb{C}$ définie, pour tous z, $z' \in \mathbb{C}$, par

$$z \boxtimes z' = \frac{1}{2}(z\overline{z'} + \overline{z}z' + zz' - \overline{zz'}).$$

On considère l'application $f : z \mapsto \mathfrak{Re}(z)^2 - \alpha\mathfrak{Im}(z)^2$ avec $\alpha \in \mathbb{R}$.

1. Déterminer $\alpha \in \mathbb{R}$ tel que
$$\forall (z, z') \in \mathbb{C}^2, \qquad f(z \boxtimes z') = f(z)f(z').$$
On fixe dorénavant $\alpha = 1$ et on pose $G = \{z \in \mathbb{C},\ f(z) \neq 0\}$.
2. Montrer que $\boxtimes$ est une loi de composition interne sur G.
3. Montrer que $(G, \boxtimes)$ est un groupe abélien.
4. Soit $\tilde{f} : \left\{ \begin{array}{ccc} G & \to & \mathbb{R}^* \\ z & \mapsto & f(z) \end{array} \right.$.
   a. L'application $\tilde{f}$ est-elle surjective ? bijective ?
   b. Calculer et dessiner l'ensemble $\tilde{f}^{-1}(\mathbb{R}_+^*)$.
5. Déterminer parmi les ensembles suivants lesquels sont des groupes pour $\boxtimes$

   a. $\{z \in \mathbb{C},\ f(z) = 1\}$
   b. $\{z \in \mathbb{C},\ f(z) > 0\}$
   c. $\{z \in \mathbb{C},\ f(z) < 0\}$
   d. $\{z \in \mathbb{C},\ f(z) = 0\}$

●●● **Problème 2 Homographies stabilisant $\mathbb{U}_n$**

Dans ce sujet, on étudie des applications $f : \mathbb{U}_n \to \mathbb{C}$ stabilisant $\mathbb{U}_n$, c'est-à-dire telles que $f(\mathbb{U}_n) \subset \mathbb{U}_n$.

Partie A – Exemples préliminaires

1. Déterminer, pour chacune des applications f suivantes, les valeurs de l'entier $n \geq 3$ pour lesquelles $f(\mathbb{U}_n) \subset \mathbb{U}_n$:

   a. $f_1 : z \mapsto \overline{z}$
   b. $f_2 : z \mapsto 1 + z$
   c. $f_3 : z \mapsto jz$
   d. $f_4 : z \mapsto \frac{1}{z}$
   e. $f_5 : z \mapsto i$
   f. $f_6 : z \mapsto z^3$

2. Reprendre la question précédente avec la condition $f(\mathbb{U}_n) = \mathbb{U}_n$.

On fixe désormais $n \geq 3$ et on se limite à certaines fonctions « homographiques », c'est-à-dire aux fonctions de la forme

$$z \mapsto \frac{az + b}{cz + d}$$

avec a, b, c et d des complexes tels que $c\omega + d \neq 0$ pour tout $\omega \in \mathbb{U}_n$.

Partie B – Stabilisation de $\mathbb{U}$

Soit f une homographie stabilisant $\mathbb{U}_n$ et $a \neq 0$, b, $c \neq 0$ et d des complexes associés comme ci-dessus.

1. Justifier que, pour tout $\omega \in \mathbb{U}_n$, $|a\omega + b| = |c\omega + d|$.
2. a. Retrouver l'identité remarquable $|z + z'|^2$ pour z, $z' \in \mathbb{C}$.
   b. Montrer que $|a|^2 + |b|^2 = |c|^2 + |d|^2$ et $a\overline{b} = c\overline{d}$.
   *Indication : on pourra utiliser la question **1** et utiliser que la somme des éléments de $\mathbb{U}_n$ vaut 0.*
   c. En déduire que $|c|^2$ et $|d|^2$ sont solutions de l'équation $z^2 - (|a|^2 + |b|^2)z + |a|^2|b|^2 = 0$.

PROBLÈMES

d. Conclure que

$$\begin{cases} |a| = |c| \\ |b| = |d| \end{cases} \quad \text{ou} \quad \begin{cases} |a| = |d| \\ |b| = |c| \end{cases}$$

3. Supposons $|a| = |c|$ et considérons $\theta \in \mathbb{R}$ tel que $a = c\,e^{i\theta}$. Montrer que $b = d\,e^{i\theta}$ puis en déduire f.
4. Supposons $|a| = |d|$.
   a. Montrer qu'il existe $\theta \in \mathbb{R}$ tel que $a = \overline{d}\,e^{i\theta}$.
   Fixons ce réel θ.
   b. Montrer que $b = \overline{c}\,e^{i\theta}$ puis en déduire la forme de l'application f.
5. Vérifier que l'application f stabilise $\mathbb{U}$.

Partie C

On considère dorénavant une application

$$f : z \mapsto e^{i\theta}\,\frac{az+b}{\overline{b}z+\overline{a}}$$

avec $a,\ b \in \mathbb{C} \setminus \{0\}$ et on suppose qu'elle stabilise $\mathbb{U}_n$.

1. Montrer que, pour tout $\omega \in \mathbb{U}_n$, $e^{in\theta}(a\omega+b)^n = \left(\overline{b}\omega+\overline{a}\right)^n$.
   On admet que, pour tout $k \in \mathbb{Z}$,

$$\sum_{\omega \in \mathbb{U}_n} \omega^k = \begin{cases} n & \text{si } n \text{ divise } k, \\ 0 & \text{sinon.} \end{cases}$$

2. En déduire que $b\,a^{n-1}e^{in\theta} = \overline{a}\,\overline{b}^{n-1}$.
3. Conclure que $|a| = |b|$.
4. Déterminer f.

Partie D – Conclusion

Préciser toutes les fonctions homographiques (au sens de l'énoncé) qui stabilisent $\mathbb{U}_n$.

CORRIGÉS

Corrigés des Vrai/Faux

a) Faux, la partie réelle vaut $\frac{3}{25}$.

b) Faux, en effet, $e^{i\frac{\pi}{6}} + e^{i\frac{9\pi}{6}} = e^{i\frac{5\pi}{6}} 2\cos\frac{2\pi}{3} = -e^{i\frac{5\pi}{6}}$ dont un argument est $-\frac{\pi}{6}$.

c) Vrai.

d) Faux, c'est bien une rotation mais le centre est le point d'affixe z solution de $z = e^{i\theta} z + 1$ c'est-à-dire $z = \frac{1}{1-e^{i\theta}} \neq 1$.

e) Faux, il manque une condition « à coefficients réels ». Pour un contre-exemple, on peut considérer le polynôme $(X-i)(X-2i)$.

f) Vrai car j est racine du polynôme X^2+X+1.

g) Faux, si n est un multiple de 3, $j^n = 1$ n'est pas racine de X^2+X+1.

h) Vrai car $j^2 = \overline{j}$ et P est à coefficients réels.

i) Vrai, pour tout $z \in \mathbb{U}_m$, $z^m = 1$ donc $z^n = 1$ (puisque m divise n) donc $z \in \mathbb{U}_n$.

j) Vrai. En fait, $\mathbb{U}_n$ est en bijection avec tous les ensembles $\{z \in \mathbb{C},\ z^n = a\}$ pour $a \neq 0$ d'après la structure de l'ensemble des racines $n^{\text{ièmes}}$ de a.

k) Vrai car $z^n = 1$ si, et seulement si, $\overline{z}^n = 1$.

l) Vrai. Il est clair que $\{\omega^2,\ \omega \in \mathbb{U}_n\} \subset \mathbb{U}_n$. Réciproquement, écrivons $n = 2k+1$ et remarquons que, pour tout $\omega \in \mathbb{U}_n$,

$$\omega = \omega^{1-n} = \omega^{-2k} = (\omega^{-k})^2,$$

d'où l'autre inclusion.

m) Vrai. Si un polynôme prend deux valeurs distinctes, alors il est non constant donc surjectif d'après le théorème de D'Alembert-Gauss.

n) Vrai, il est clair que toute fonction constante n'est pas injective. De plus, si une fonction polynomiale est de degré n, alors chaque complexe admet n antécédents d'après le théorème de D'Alembert-Gauss : ainsi le degré d'une fonction polynomiale injective est 1.

Corrigés des exercices d'application du chapitre 2

Exercice 1

On passe par les écritures géométriques $i - \sqrt{3} = 2e^{\frac{5i\pi}{6}}$ et $1 + i\tan\theta = \frac{1}{\cos\theta} e^{i\theta}$ et on en déduit

$$(i-\sqrt{3})^n = 2^n e^{\frac{5ni\pi}{6}}, \qquad (1+i\tan\theta)^n = \frac{1}{(\cos\theta)^n} e^{ni\theta}.$$

Exercice 2
Commençons par écrire le numérateur et le dénominateur sous forme géométrique en mettant en facteur le module et en reconnaissant l'argument :

$$\sqrt{3}-i=2e^{-\frac{i\pi}{6}}, \qquad 1+i=\sqrt{2}e^{\frac{i\pi}{4}}.$$

Alors,

$$\left(\frac{\sqrt{3}-i}{1+i}\right)^{15}=\sqrt{2}^{-15}e^{15\left(-\frac{i\pi}{6}-\frac{i\pi}{4}\right)}=\sqrt{2}^{-15}e^{-\frac{25i\pi}{4}}.$$

La partie réelle est donc $\sqrt{2}^{-15}\cos(-\frac{25\pi}{4})=2^{-8}$.

Exercice 3
Commençons par écrire $1+i$ sous forme géométrique $\sqrt{2}e^{i\frac{\pi}{4}}$. La question se reformule en la recherche des entiers $n\in\mathbb{N}$ tels que $\sqrt{2}^n e^{in\frac{\pi}{4}}\in i\mathbb{R}$ soit en

$$n\frac{\pi}{4}=\frac{\pi}{2}[\pi].$$

On obtient ainsi la condition nécessaire et suffisante de congruence suivante : $n=2[4]$.

Exercice 4
On commence par calculer le discriminant $\Delta=1-4(-1+i)(1+i)=9$. Les solutions sont donc $\frac{1\pm3}{2(1+i)}$ soit $1-i$ et $-\frac{1}{2}(1-i)$.

Exercice 5
Un complexe z est solution de cette équation si, et seulement si, $Z=z^2$ est solution de l'équation $Z^2-2\cos(\theta)Z+1=0$, c'est-à-dire si $Z=e^{i\theta}$ ou $Z=e^{-i\theta}$. Les solutions sont donc $e^{i\frac{\theta}{2}}$, $-e^{i\frac{\theta}{2}}$, $e^{-i\frac{\theta}{2}}$ et $-e^{-i\frac{\theta}{2}}$.

Exercice 6
On pourrait calculer le discriminant mais il peut être plus malin ici de remarquer que l'équation se réécrit $(z-1)^2=1-i$ soit encore $(z-1)^2=2^{\frac{1}{2}}e^{-i\frac{\pi}{4}}$. On obtient alors les solutions, d'après le résultat sur les solutions d'une équation de la forme $Z^2=a$,

$$1\pm2^{\frac{1}{4}}e^{-i\frac{\pi}{8}}.$$

On remarque que dans cette solution, nous avons en fait choisi d'utiliser la mise « sous forme canonique » plutôt que la routine sur le discriminant.

Exercice 7
Résolvons l'équation. Notons $\alpha\in\mathbb{C}$ tel que $\alpha^n=\alpha$. Un complexe z vérifie $\left(\frac{1+iz}{1-iz}\right)^n=\alpha$ si, et seulement si,

$$\frac{1+iz}{1-iz}=\alpha\omega,$$

avec $\omega\in\mathbb{U}_n$. Ceci équivaut encore à

$$z=i\frac{1-\alpha\omega}{1+\alpha\omega}=i\frac{(1-\alpha\omega)(1+\overline{\alpha\omega})}{|1+\alpha\omega|^2}.$$

Pour que z soit réel, il faut et il suffit que

$$(1-\alpha\omega)(1+\overline{\alpha\omega})=1-|\alpha|^2-2i\mathfrak{Im}(\alpha\omega)$$

soit un imaginaire pur, c'est-à-dire que $|\alpha|=1$. Comme $|\alpha|=|\alpha|^n$, la condition recherchée est $|\alpha|=1$.

Exercice 8

$$\begin{aligned}|ab-cd| &= |ab-bc+bc-cd| \\ &\leq |ab-bc|+|bc-cd| \\ &\leq |a-c|+|b-d|,\end{aligned}$$

car $|b|=|c|=1$.

Exercice 9

▹ Écrivons $2a=(a+b)+(a-b)$ et $2b=(a+b)-(a-b)$ puis utilisons l'inégalité triangulaire.

$$\begin{aligned}|a|+|b| &\leq \frac{1}{2}(|a+b|+|a-b|)+\frac{1}{2}(|a+b|+|a-b|) \\ &\leq |a+b|+|a-b|.\end{aligned}$$

▹ Par ailleurs, on a égalité si, et seulement si, on a égalité dans l'inégalité triangulaire avec $a+b$ et $a-b$ et dans l'inégalité triangulaire entre $a+b$ et $b-a$. Ainsi, $a=b$ où il existe $\lambda \geq 0$ et $\mu \geq 0$ tels que $a+b=\lambda(a-b)$ et $a+b=\mu(b-a)$. Ces deux dernières égalités entraînent $a=-b$. En conclusion, le cas d'égalité correspond à $a=\pm b$.

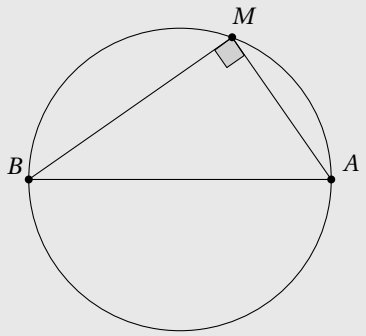
▹ En notant A et B les points d'afixes a et b, l'inégalité indique que, dans le parallélogramme construit sur les vecteurs $\overrightarrow{OA}$ et $\overrightarrow{OB}$, la somme des longueurs de deux côtés consécutifs est inférieure à la somme des longueurs des diagonales.

Exercice 10

Utilisons l'expression du module au carré comme le produit du complexe par son conjugué. Alors,

$$\begin{aligned}|z-1|^2+|z+1|^2 &= (z-1)(\overline{z}-1)+(z+1)(\overline{z}+1) \\ &= |z|^2-z-\overline{z}+1+|z|^2+z+\overline{z}+1=4.\end{aligned}$$

Pour interpréter géométriquement ce résultat, introduisons les points A, B et M d'affixes 1, -1 et z réciproquement. La quantité calculée est donc AM^2+BM^2 ; comme le triangle ABM est rectangle en M (ce sont trois points du cercle unité et AB en est un diamètre), $AM^2+BM^2=AB^2=4$ d'après le théorème de Pythagore.



Exercice 11

Vérifions que $\overline{u} = -u$; en effet,

$$\overline{u} = \frac{\overline{z} + \overline{a}\,\overline{b}\,z - \overline{a} - \overline{b}}{\overline{a} - \overline{b}} = \frac{\overline{z} + \frac{1}{a}\frac{1}{b}z - \frac{1}{a} - \frac{1}{b}}{\frac{1}{a} - \frac{1}{b}} = \frac{a\,b\,\overline{z} + z - b - a}{b - a} = -u.$$

Ainsi, $u^2 = -|u|^2 \in \mathbb{R}_-$.

Exercice 12

Pour obtenir une détermination de l'angle $(\overrightarrow{M_{n+1}M_n}, \overrightarrow{M_{n+1}M_{n+2}})$, calculons un argument de

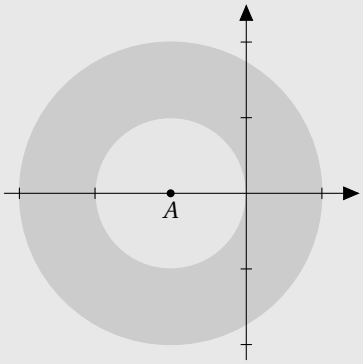
$$\frac{a^n - a^{n+1}}{a^{n+2} - a^{n+1}} = \frac{1-a}{a(a-1)} = -\frac{\overline{a}}{|a|^2} \in i\mathbb{R}.$$

Ainsi, $(\overrightarrow{M_{n+1}M_n}, \overrightarrow{M_{n+1}M_{n+2}}) = \frac{\pi}{2}[\pi]$.

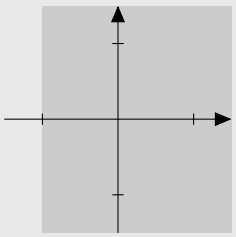
Exercice 13

Dans ce corrigé on note M le point d'affixe z.

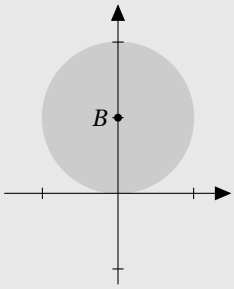
1. Soit A le point d'affixe -1. M appartient à la zone indiquée si, et seulement si, $AM \in [1,2]$. Il s'agit de la couronne centrée sur A, de rayon intérieur 1 et de rayon extérieur 2.



2. Avec la linéarité de la partie réelle, l'équation se réécrit $\mathfrak{Re}(z) \geq -1$. On obtient donc le demi-plan délimité par la droite verticale d'équation $x = -1$.



3. En factorisant par i, on a $|iz+1| = |i|.|z-i| = |z-i|$. L'équation indique donc que la distance entre M et le point d'affixe B d'affixe i est inférieure à ou égale à 1.



Exercice 14

Notons g l'affixe du centre de gravité. La condition se réécrit alors

$$|a-g| = |b-g| = |c-g|,$$

c'est-à-dire que le centre de gravité est le centre du cercle circonscrit : le triangle est équilatéral.

Exercice 15

▹ Montrons l'équivalence entre **i.** et **ii.**. Les points d'affixes a, b, c forment un triangle équilatéral si, et seulement si, l'une des rotations de centre le point d'affixe b et d'angle $\pm\frac{\pi}{3}$ envoie le point d'affixe a sur le point d'affixe c, c'est-à-dire :

$$c-b = e^{i\frac{\pi}{3}}(a-b) \qquad \text{ou} \qquad c-b = e^{-i\frac{\pi}{3}}(a-b).$$

Cette condition se réécrit alors $-xa+(x-1)b+c=0$ avec $x \in \{e^{i\frac{\pi}{3}}, e^{-i\frac{\pi}{3}}\}$, soit encore, après de rapides calculs,

- $a\,j^2+b\,j+c=0$ si $x=e^{i\frac{\pi}{3}}$;
- $a\,j+b\,j^2+c=0$ si $x=e^{-i\frac{\pi}{3}}$.

On retrouve bien **ii.**

▹ Montrons l'équivalence entre **iii.** et **ii.** La condition **iii.** signifie que c est solution de l'équation

$$z^2-(a+b)z+a^2+b^2-a\,b=0,$$

dont le discriminant est $-3(a-b)^2$ et dont les racines sont

$$\frac{a+b\pm i\sqrt{3}(a-b)}{2}$$

ou encore $-a\,j^2-b\,j$ et $-a\,j-b\,j^2=-a\,j^4-b\,j^2$. Le complexe c est donc égal à l'un de ces deux nombres si, et seulement si, j ou j^2 est solution de $a\,z^2+b\,z+c=0$.

▹ Pour l'équivalence entre **iii.** et **iv.**, il suffit de mettre au même dénominateur car

$$(b-c)(c-a)+(a-b)(c-a)+(a-b)(b-c)=a\,b+a\,c+b\,c-a^2-b^2-c^2.$$

Exercice 16

Soit α, β et γ les arguments principaux de u, v et w. Sans perte de généralité, on peut supposer que $\alpha \leq \beta \leq \gamma$. Or, $1=|w|=|u+v|=2\left|\cos\frac{\beta-\alpha}{2}\right|$. Ainsi, $\beta-\alpha \in \left\{-\frac{2\pi}{3}, \frac{2\pi}{3}\right\}$ c'est-à-dire $v=j\,u$ ou $v=j^2\,u$ et donc $w=j^2\,u$ ou $w=j\,u$ respectivement.

Exercice 17

Remarquons tout d'abord qu'en conjuguant, $\overline{z_1}+\overline{z_2}+\overline{z_3}=1$ et comme les complexes sont de module 1,

$$\frac{1}{z_1}+\frac{1}{z_2}+\frac{1}{z_3}=1.$$

Ainsi, $z_1z_2+z_2z_3+z_3z_1=1$. D'après les relations entre coefficients et racines, les trois complexes sont les racines du polynôme $X^3-X^2+X-1=(X^2+1)(X-1)$. En conclusion, $\{z_1,z_2,z_3\}=\{1,i,-i\}$.

Exercice 18

1. Par inégalité triangulaire, $|a-b|\leq|a|+|b|$ donc

$$|a-b|^2\leq|a|^2+|b|^2+2|ab|.$$

Il reste à voir que l'inégalité arithmético-géométrique donne $2|ab|\leq|ab|^2+1$ pour conclure.
2. Il suffit d'avoir l'égalité dans les deux majorations utilisées ci-dessus donc $|ab|=1$ d'une part et a et $-b$ de même argument modulo 2π.

Exercice 19

Un complexe z appartient à $f(\mathbb{U})$ si, et seulement si, il existe $\theta\in\mathbb{R}\setminus\{\pi+2k\pi,\ k\in\mathbb{Z}\}$ tel que

$$z=\frac{2}{e^{i\theta}+1}=e^{-i\frac{\theta}{2}}\frac{1}{\cos\frac{\theta}{2}}=1-i\tan\frac{\theta}{2}.$$

Par ailleurs, on remarque que l'image directe de $\mathbb{R}\setminus\{\pi+2k\pi,\ k\in\mathbb{Z}\}$ par la fonction tan est $\mathbb{R}$. Ainsi, tout complexe de partie réelle 1 est de la forme $1-i\tan\frac{\theta}{2}$ avec $\theta\in\mathbb{R}\setminus\{\pi+2k\pi,\ k\in\mathbb{Z}\}$.
Ainsi, $z\in f(\mathbb{U})$ si, et seulement si, z appartient à la droite verticale d'équation $\mathfrak{Re}(z)=1$.

Exercice 20

1. Soit $z=x+iy\in P$ avec $x\in\mathbb{R}$ et $y>0$. Calculons le module de $\frac{z-i}{z+i}$.

$$\begin{aligned}\left|\frac{z-i}{z+i}\right|^2&=\frac{x^2+(y-1)^2}{x^2+(y+1)^2}\\&=1-\frac{4y}{x^2+(y+1)^2}<1.\end{aligned}$$

2. Soit $z=x+iy\neq-i$ avec $x,\ y\in\mathbb{R}$ tel que $\frac{z-i}{z+i}\in D$. Le calcul de la question précédente donne

$$\left|\frac{z-i}{z+i}\right|^2=1-\frac{4y}{x^2+(y+1)^2}.$$

Par conséquent, $y>0$.
3. Soit $z'\in D$. Résolvons l'équation $z'=f(z)$.

$$\begin{aligned}z'=f(z)&\Leftrightarrow z'(z+i)=z-i\\&\Leftrightarrow z(z'-1)=-i(1+z')\\&\Leftrightarrow z=\frac{-i(1+z')}{z'-1}.\end{aligned}$$

D'après la question précédente, l'unique antécédent z de z' appartient à P. Ainsi, f est bijective de réciproque

$$\left\{\begin{array}{ccc}D&\to&P\\z&\mapsto&\frac{-i(1+z)}{z-1}\end{array}\right.$$

Corrigés des exercices d'approfondissement du chapitre 2

Exercice A

Le terme considéré est le module au carré du complexe $z=(a+i)(b+i)(c+i)$ dont la partie imaginaire est $ab+bc+ca-1=0$ et la partie réelle est un entier.

Exercice B

1. Un réel x est solution de $(E_{\pi/2})$ si $x^3+2x^2-ix-2i=0$ soit en prenant partie réelle et imaginaire $x^3+2x^2=0$ et $x+2=0$: la seule solution réelle est donc $x=-2$.
2. On vérifie immédiatement que -2 est solution des équations (E_α) pour tout $\alpha\in\mathbb{R}$.
3. Factorisons $z+2$ dans l'équation (E_α) :

$$(z+2)(z^2-e^{i\alpha})=0.$$

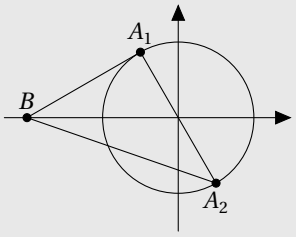
Les solutions de (E_α) sont donc -2, $e^{i\alpha/2}$ et $-e^{i\alpha/2}$.

4. Notons A_1, A_2 et B les points d'affixes $e^{i\alpha/2}$, $-e^{i\alpha/2}$ et -2 respectivement.

Le triangle A_1A_2B ne peut être rectangle en B car B n'appartient pas au cercle unité. En revanche, ce triangle est rectangle en A_1 si $A_1B^2+A_1A_2^2=BA_2^2$ (théorème de Pythagore) soit en complexes

$$\left|-2-e^{i\alpha/2}\right|^2+\left|-e^{i\alpha/2}-e^{i\alpha/2}\right|^2=\left|-2+e^{i\alpha/2}\right|^2.$$

En développant, on obtient $\cos\frac{\alpha}{2}=-\frac{1}{2}$ soit $\frac{\alpha}{2}=\pm\frac{2\pi}{3}[2\pi]$.



De même, il est rectangle en A_2 si

$$\left|-2+e^{i\alpha/2}\right|^2+\left|e^{i\alpha/2}+e^{i\alpha/2}\right|^2=\left|-2-e^{i\alpha/2}\right|^2,$$

c'est-à-dire $\cos\frac{\alpha}{2}=\frac{1}{2}$ soit $\frac{\alpha}{2}=\pm\frac{\pi}{3}[2\pi]$.

Exercice C

Cherchons tout d'abord une solution réelle x : elle vérifie $\mathfrak{Re}(x^3+(i-2)x^2+(3-3i)x+2i-2)=0$ soit $x^3-2x^2+3x-2=0$ donc $x=1$ est une solution possible. On vérifie immédiatement qu'elle est effectivement solution de l'équation initiale.

Factorisons l'équation par $z-1$:

$$z^3+(i-2)z^2+(3-3i)z+2i-2=(z-1)(z^2+(i-1)z+2-2i).$$

Les autres solutions de l'équation initiale sont donc les solutions de l'équation $z^2+(i-1)z+2-2i=0$. Résolvons cette équation avec la méthode habituelle : le discriminant est $\Delta=-8+6i$ dont les racines carrées sont $\pm(1+3i)$. Ainsi, les solutions sont $1+i$ et $-2i$.

En conclusion, les solutions recherchées sont -1, $1+i$ et $-2i$.

CORRIGÉS

Exercice D

1. D'après l'identité remarquable, $(a+ib)^2 = a^2 - b^2 + 2iab$.
2. D'après la première question, cette équation se réécrit $z^2 = 2i(a-ib)^2$ soit encore
$$z^2 = \left(\sqrt{2}e^{i\frac{\pi}{2}}(a-ib)\right)^2.$$

On en déduit que les deux solutions sont $\pm\sqrt{2}e^{i\frac{\pi}{2}}(a-ib)$.

Exercice E

Remarquons tout d'abord que, pour toute solution z, $|z| \in \mathbb{R}$ donc $\overline{z} \in \mathbb{R}$ soit encore $z \in \mathbb{R}$.
Il suffit alors de distinguer les cas selon le signe du réel z. Si $z \geq 0$, l'équation devient $1-z = z$ donc $z = \frac{1}{2}$.
Si $z < 0$, alors l'équation devient $1-z = -z$ ce qui est impossible.
En conclusion, la seule solution est $\frac{1}{2}$.

Exercice F

Remarquons tout d'abord que si z est solution, alors $|z|^n = |z|$ donc $|z| = 0$ ou $|z| = 1$.
Cherchons désormais les solutions z non nulles (donc de module 1) :
$$\begin{aligned} z^n = \overline{z} &\Leftrightarrow z^{n+1} = \overline{z}z \\ &\Leftrightarrow z^{n+1} = 1. \end{aligned}$$
En conclusion, les solutions sont 0 et les racines $n+1^{\text{ièmes}}$ de l'unité.

Exercice G

Si $\omega = 1$, cette quantité vaut $\frac{3}{2}$. Sinon, en mettant au même dénominateur, on obtient
$$\begin{aligned} \frac{\omega}{1+\omega^2} + \frac{\omega^2}{1+\omega^4} + \frac{\omega^3}{1+\omega^6} &= \frac{2(1+\omega+\omega^2+\omega^3+\omega^4+\omega^5)}{1+\omega+\omega^2+\omega^3+\omega^4+\omega^5+2\omega^6} \\ &= \frac{-2\omega^6}{\omega^6} = -2, \end{aligned}$$
en utilisant que $1+\omega+\omega^2+\omega^3+\omega^4+\omega^5+\omega^6 = 0$.

Exercice H

Cherchons les solutions sous la forme $z = re^{i\theta}$ avec $r \in \mathbb{R}_+$ et $\theta \in \mathbb{R}$. L'équation se réécrit alors $r^n\cos(n\theta) = r^n\sin(n\theta)$ soit $r = 0$ ou $n\theta = \frac{\pi}{4}[\pi]$.
En conclusion, les solutions de l'équation sont donc 0 et, pour tout $r > 0$ et tout entier k,
$$r\exp\left(i\frac{(4k+1)\pi}{4n}\right).$$

Exercice I

Faisons apparaître $j + j^2$ dans le membre de gauche en utilisant que cette quantité vaut -1. Ainsi, on obtient en appliquant ensuite l'inégalité triangulaire,
$$\begin{aligned} |z-1| &= |(z-1)(j+j^2)| \\ &= |jz - j^2 + j^2z - j| \\ &\leq |jz - j^2| + |j^2z - j| \\ &= |j||z-j| + |j^2||z-j^2| \\ &= |z-j| + |z-j^2|. \end{aligned}$$

Exercice J

Remarquons que

$$|z-z'|\times|z+z'|=|z^2-z'^2|\leq|z^2|+|z'^2|\leq 2$$

par inégalité triangulaire. Or, $|z-z'|\times|z+z'|\geq\min(|z-z'|,|z+z'|)^2$ d'où le résultat annoncé.

On remarque que cette majoration est optimale avec $z=1$, $z'=i$ par exemple puisque, dans ce cas, $|z+z'|=|z-z'|=\sqrt{2}$.

Exercice K

Le complexe $\frac{z-u\overline{z}}{1-u}$ est réel si, et seulement s'il est égal à son conjugué, c'est-à-dire

$$\begin{aligned}\frac{z-u\overline{z}}{1-u}=\frac{\overline{z}-\overline{u}z}{1-\overline{u}} &\Leftrightarrow (z-u\overline{z})(1-\overline{u})=(\overline{z}-\overline{u}z)(1-u)\\ &\Leftrightarrow z-\overline{u}z-u\overline{z}+|u|^2\overline{z}=\overline{z}-u\overline{z}-\overline{u}z+|u|^2z\\ &\Leftrightarrow (z-\overline{z})(1-|u|^2)=0\\ &\Leftrightarrow |u|^2=1,\end{aligned}$$

car z n'est pas réel donc $z\neq\overline{z}$.

Corrigés des problèmes du chapitre 2

Problème 1

1. On remarque que $f(i\boxtimes i)=1$ et que $f(i)=-\alpha$, d'où $\alpha^2=1$. Réciproquement,

- Pour $\alpha=-1$, $f:z\mapsto|z|^2$. Alors,

$$f(1+i)^2=4,$$
$$f((1+i)\boxtimes(1+i))=f(2+2i)=8.$$

Ainsi, la valeur -1 ne convient pas.

- Pour $\alpha=1$, on a, pour tout x, x', y, $y'\in\mathbb{R}$,

$$\begin{aligned}f((x+iy)\boxtimes(x'+iy'))&=f((xx'+yy')+i(xy'+x'y))\\ &=(xx'+yy')^2-(xy'+x'y)^2\\ &=x^2(x'^2-y'^2)+y^2(y'^2-x'^2)\\ &=(x^2-y^2)(x'^2-y'^2)\\ &=f(x+iy)f(x'+iy').\end{aligned}$$

Ainsi, la valeur $\alpha=1$ convient.

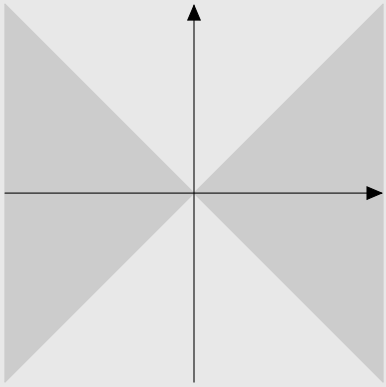
2. Soit z, $z'\in G$. Alors, $f(z)\neq 0$ et $f(z')\neq 0$ d'où $f(z\boxtimes z')=f(z)f(z')\neq 0$ donc $z\boxtimes z'\in G$.

3. L'associativité et la commutativité sont faciles à vérifier ; $1\in G$ est l'élément neutre ; le symétrique de $z\in G$ est $\frac{\overline{z}}{f(z)}\in G$: en effet,

$$z\boxtimes f(z)=\frac{1}{2}\left(\frac{z^2}{f(z)}-\frac{\overline{z}^2}{f(z)}\right)=1.$$

4. a. L'application $\tilde{f}$ est surjective : en effet, si $x\geq 0$, $x=\tilde{f}(\sqrt{x})$ et si $x<0$, $x=\tilde{f}(\sqrt{-x})$.
En revanche, $\tilde{f}$ n'est pas bijective car $\tilde{f}(-1)=\tilde{f}(1)$.

b. L'ensemble $\tilde{f}^{-1}(\mathbb{R}_+^*)$ est l'ensemble des complexes dont la partie réelle est en valeur absolue strictement supérieure à la partie imaginaire, soit la zone colorée ci-dessous



5. On vérifie que les deux premiers sont des groupes ; les deux derniers ne sont pas des groupes (par exemple, ils ne contiennent pas l'élément neutre 1).

Problème 2

Partie A – Exemples préliminaires

1. a. D'après le cours, tout groupe $\mathbb{U}_n$ est stable par f_1 (car pour $z \in \mathbb{C}$, $\overline{z^n} = \overline{z}^n$).
b. Aucun groupe $\mathbb{U}_n$ n'est stable par f_2 car $1 \in \mathbb{U}_n$ et pourtant $2 = f(1) \notin \mathbb{U}_n$.
c. Un groupe $\mathbb{U}_n$ est stable par f_3 si, et seulement si, pour tout $z \in \mathbb{U}_n$, $(jz)^n = 1$ soit $j^n = 1$. Ainsi, les seuls groupes $\mathbb{U}_n$ stables par f_3 sont ceux pour lesquels 3 divise n.
d. D'après le cours, tout groupe $\mathbb{U}_n$ est stable par f_4 : en effet, pour tout $z \in \mathbb{U}_n$, $\frac{1}{z} = z^{n-1} \in \mathbb{U}_n$.
e. Un groupe $\mathbb{U}_n$ est stable par f_5 si, et seulement si, $i \in \mathbb{U}_n$ c'est-à-dire si 4 divise n.
f. Tout groupe $\mathbb{U}_n$ est stable par f_6 puisque le produit est une loi de composition interne à $\mathbb{U}_n$.
2. Il n'y a rien à étudier pour f_2. Comme les applications f_1, f_3, f_4 sont injectives, une inclusion entraîne l'égalité (les n éléments de $\mathbb{U}_n$ ont n images distinctes) donc la réponse est identique à celle de la question précédente pour ces applications.
L'application f_5 n'atteint pas 1 donc il n'existe pas de n tel que $f_5(\mathbb{U}_n) = \mathbb{U}_n$.
Pour f_6, on remarque que si 3 divise n, alors plusieurs éléments de $\mathbb{U}_n$ admettent la même image : on n'atteint pas les n éléments de $\mathbb{U}_n$ à l'arrivée. En revanche, sinon, 3 est premier avec n donc, d'après le théorème de Bezout, il existe des entiers u et v tels que $3u + nv = 1$: pour tout $z \in \mathbb{U}_n$,

$$f_6(z^u) = z^{3u} = z^{3u+nv} = z.$$

Ainsi, tout $\mathbb{U}_n$ est atteint par f_6. La réponse est donc tous les groupes $\mathbb{U}_n$ avec n non divisible par 3.

Partie B – Stabilisation de $\mathbb{U}$

1. Pour tout $\omega \in \mathbb{U}_n$, $f(\omega) \in \mathbb{U}_n \subset \mathbb{U}$ donc $|f(\omega)| = 1$ ce qui se traduit par $|a\omega + b| = |c\omega + d|$.
2. a. Développons en utilisant le conjugué. Pour tous z, $z' \in \mathbb{C}$,

$$|z + z'|^2 = (z + z')(\overline{z} + \overline{z'}) = |z|^2 + |z'|^2 + z\overline{z'} + z'\overline{z}.$$

On peut encore écrire cette identité sous la forme

$$|z + z'|^2 = |z|^2 + |z'|^2 + 2\mathfrak{Re}(z\overline{z'}).$$

b. D'après la question **1**, pour tout $\omega \in \mathbb{U}_n$, $|a\omega + b|^2 = |c\omega + d|^2$ soit, en développant

$$|a|^2 + |b|^2 - |c|^2 - |d|^2 = 2\mathfrak{Re}((c\overline{d} - a\overline{b})\omega).$$

En additionnant toutes ces relations et en utilisant la linéarité de $\mathfrak{Re}$, on obtient que

$$n\left(|a|^2+|b|^2-|c|^2-|d|^2\right)=2\mathfrak{Re}\left((c\overline{d}-a\overline{b})\sum_{\omega\in\mathbb{U}_n}\omega\right)=0.$$

Ainsi, $|a|^2+|b|^2=|c|^2+|d|^2$ et par conséquent, $\mathfrak{Re}((c\overline{d}-a\overline{b})\omega)=0$ pour tout $\omega\in\mathbb{U}_n$. En utilisant tour à tour les valeurs $\omega=1$ et $\omega=e^{\frac{2i\pi}{n}}$, on obtient que $\mathfrak{Re}(c\overline{d}-a\overline{b})=0$ puis $\mathfrak{Im}(c\overline{d}-a\overline{b})=0$. En conclusion, $a\overline{b}=c\overline{d}$.

c. D'après la question précédente, $|c|^2+|d|^2=|a|^2+|b|^2$ et $|c|^2|d|^2=|c\overline{d}|^2=|a\overline{b}|^2=|a|^2|b|^2$. Connaissant somme et produit, on en déduit que $|c|^2$ et $|d|^2$ sont solutions de l'équation

$$z^2-(|a|^2+|b|^2)z+|a|^2|b|^2=0.$$

d. Les solutions de l'équation précédente sont $|a|^2$ et $|b|^2$ donc

$$\{|a|^2,|b|^2\}=\{|c|^2,|d|^2\}.$$

Comme les modules sont positifs, on en déduit $\{|a|,|b|\}=\{|c|,|d|\}$ ce qui est la condition recherchée dans l'énoncé.

3. En reportant dans la relation $a\overline{b}=c\overline{d}$ et en simplifiant par $c\neq 0$, on obtient $e^{i\theta}\overline{b}=\overline{d}$ donc $b=de^{i\theta}$. Par conséquent, $f:z\mapsto e^{i\theta}$.

4. a. Par hypothèse $|a|=|\overline{d}|$ donc $d\neq 0$ et $|a/d|=1$. Ainsi, il existe $\theta\in\mathbb{R}$ tel que $a=\overline{d}e^{i\theta}$.

b. Une nouvelle fois, on reporte dans la relation $a\overline{b}=c\overline{d}$ et on simplifie pour obtenir $b=\overline{c}e^{i\theta}$. Alors f est de la forme

$$z\mapsto e^{i\theta}\frac{\overline{d}z+\overline{c}}{cz+d}.$$

5. Il est clair que $fz\mapsto e^{i\theta}$ stabilise $\mathbb{U}$. Supposons dorénavant que f est de la forme obtenue à la question **4b**. Alors, pour tout $z\in\mathbb{U}$ dans le domaine de définition de f :

$$\left|e^{i\theta}(\overline{d}z+\overline{c})\right|=\left|\overline{d}z+\overline{c}\right|=\left|d\overline{z}+c\right|=\frac{1}{|z|}|d+cz|=|d+cz|.$$

En conclusion, $f(\mathbb{U})=\mathbb{U}$.

Partie C

1. Par définition de stabilisation, pour tout $\omega\in\mathbb{U}_n$, $f(\omega)^n=1$ donc $e^{in\theta}(a\omega+b)^n=(\overline{b}\omega+\overline{a})^n$.
2. Calculons avec la formule du binôme, le théorème de Fubini et la question précédente :

$$\begin{aligned}\sum_{\omega\in\mathbb{U}_n}\omega(a\omega+b)^n&=\sum_{\omega\in\mathbb{U}_n}\omega\sum_{k=0}^{n}\binom{n}{k}a^kb^{n-k}\omega^k\\&=\sum_{k=0}^{n}\binom{n}{k}a^kb^{n-k}\sum_{\omega\in\mathbb{U}_n}\omega^{k+1}\\&=a^{n-1}b.\end{aligned}$$

De même,

$$\sum_{\omega\in\mathbb{U}_n}\omega(\overline{b}\omega+\overline{a})^n=\overline{a}\overline{b}^{n-1}.$$

Avec l'identité de la question **1**, on obtient $ba^{n-1}e^{in\theta}=\overline{a}\overline{b}^{n-1}$.

3. En passant au module dans la question précédente, on obtient $|a|^{n-1}|b|=|a||b|^{n-1}$ donc $|a|^{n-2}=|b|^{n-2}$. Comme les modules sont positifs, $|a|=|b|$.
4. Notons $a=e^{i\varphi}b$ avec $\varphi\in\mathbb{R}$. Alors f est de la forme

$$z\mapsto e^{i\theta}\frac{e^{i\varphi}z+1}{z+e^{-i\varphi}}=e^{i(\theta+\varphi)}.$$

Partie D – Conclusion

Discutons si coefficients a et b sont nuls.

- Si a et $b \neq 0$, alors d'après les parties B et C, f est constante ; elle stabilise $\mathbb{U}_n$ si, et seulement si, cette constante appartient à $\mathbb{U}_n$.
- Si $a = 0$ et $b \neq 0$, alors f est de la forme $z \mapsto \frac{\kappa}{z}$; elle stabilise $\mathbb{U}_n$ si, et seulement si, $\kappa \in \mathbb{U}_n$ (condition nécessaire car $f(1) = \kappa \in \mathbb{U}_n$).
- Si $a \neq 0$ et $b = 0$, alors f est de la forme $z \mapsto \kappa z$; elle stabilise $\mathbb{U}_n$ si, et seulement si, $\kappa \in \mathbb{U}_n$ (condition nécessaire car $f(1) = \kappa \in \mathbb{U}_n$).

Les fonctions homographiques (au sens de cet énoncé) qui stabilisent $\mathbb{U}_n$ sont les $z \mapsto \omega$, $z \mapsto \omega z$ et $z \mapsto \frac{\omega}{z}$ pour $\omega \in \mathbb{U}_n$.

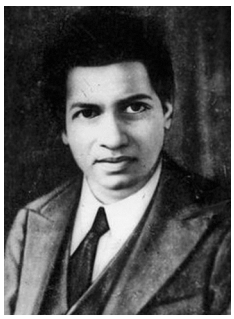
Chapitre 3

Sommes et produits (finis)

1. Règles de manipulation — **2.** Quelques sommes classiques — **3.** Formule du binôme

Objectifs et compétences du programme

- Apprendre à manipuler les quantités exprimées avec $\sum$ et $\prod$.
- Découvrir les aspects calculatoires (par opposition aux aspects combinatoires) des coefficients binomiaux.
- Reconnaître des sommes classiques (géométriques, binomiales...).



Srinivasa Ramanujan, mathématicien indien (1887-1920)
Il est l'un des rares mathématiciens à avoir vu son histoire inspirer un film de grande diffusion (*The Man Who Knew Infinity* réalisé par Matthew Brown, en 2016). Deux livres dévorés pendant son adolescence lui serviront de formation initiale en mathématiques. En autodidacte, il va développer une singulière vision des mathématiques, de géniales intuitions calculatoires (mais aussi un style de rédaction laconique) qui entraîneront une riche collaboration avec de grands mathématiciens britanniques comme Geoffrey Hardy.
Les carnets de Srinivasa Ramanujan débordent littéralement de formules toutes plus exotiques les unes que les autres et il est toujours difficile même pour un mathématicien actuel de comprendre le cheminement intellectuel qui a amené Ramanujan à intuiter les valeurs de telles ou telles sommes.

COURS

Ce chapitre, à vocation essentiellement technique, permet de se familiariser avec une notation compacte et précise pour les sommes et produits (et par conséquent, d'éviter les disgracieux « ... »). Il convient de prêter attention aux différentes règles de manipulation et de recalculer les sommes classiques présentées en exemples pour les intégrer comme naturelles.

1. Règles de manipulation

Cette section rassemble des exemples importants de manipulations des sommes et produits finis. Commençons par rappeler les notations $\sum$ et $\prod$.

Notation

Soit une famille finie non vide $(u_k)_{k\in I}$ de complexes,

$$\sum_{k\in I} u_k, \qquad \prod_{k\in I} u_k$$

désignent les somme et produit de tous les complexes de cette famille.
En particulier, si $u_1, u_2, \ldots, u_n$ sont des complexes, on note

$$\sum_{k=1}^{n} u_k = u_1 + u_2 + \ldots + u_n,$$
$$\prod_{k=1}^{n} u_k = u_1 \times u_2 \times \ldots \times u_n.$$

Par convention, la somme des éléments de la famille vide vaut 0 et son produit vaut 1.

On indique quelquefois une condition supplémentaire sur les valeurs de l'indice. Par exemple, si on ne considère que les indices entre p et q, qui sont diviseurs de n, c'est-à-dire l'ensemble d'indices

$$I = \{k \in [\![p, q]\!],\ k|n\},$$

on note la somme et le produit par

$$\sum_{\substack{k=p\\ k|n}}^{q} u_k, \qquad \prod_{\substack{k=p\\ k|n}}^{q} u_k.$$

De même, les quantités

$$\sum_{\substack{k=1\\ k \text{ pair}}}^{n} u_k, \qquad \prod_{\substack{k=1\\ k \text{ pair}}}^{n} u_k$$

désigne les sommes sur l'ensemble d'indices $I = \{k \in [\![1, n]\!],\ k = 0[2]\}$.

Proposition 3.1.

Soit I un ensemble fini et $a \in \mathbb{C}$. Alors,

$$\sum_{i\in I} a = a|I|$$

où $|I|$ désigne le nombre d'éléments de I.

Exemple

Soit $m < n$ deux entiers naturels, $a \in \mathbb{C}$. Alors, $\sum_{k=m}^{n} a = (n-m+1)a$. En effet, $n-m+1$ est le nombre d'entiers entre m et n (au sens large).

Voici un exemple de produit de référence.

Définition 3.2.

Soit $n \in \mathbb{N}$. La factorielle de n, prononcée « factorielle n », est l'entier noté $n!$ défini par

$$n! = \prod_{k=1}^{n} k.$$

En particulier, d'après la convention sur le produit vide, $0! = 1$.

Exemple

Soit $m < n$ deux entiers naturels. Alors, $\prod_{k=m}^{n} k = \frac{\prod_{k=1}^{n} k}{\prod_{k=1}^{m-1} k} = \frac{n!}{(m-1)!}$.

On peut résumer la philosophie générale des règles de manipulation ainsi : *dans le terme de droite d'une égalité, on doit retrouver chaque terme du membre de gauche une et une seule fois et aucun autre terme.* La première « règle » est que l'indice est muet donc on peut changer sa notation (à condition bien sûr de la changer à toutes ses apparitions) et qu'il n'est pas défini en dehors de la somme/du produit.

Proposition 3.3.

Soit $u_1, \ldots, u_n \in \mathbb{C}$. Alors, $\sum_{k=0}^{n} u_k = \sum_{l=0}^{n} u_l$, et $\prod_{k=0}^{n} u_k = \prod_{l=0}^{n} u_l$.

Remarque

On peut aussi choisir des noms de variables plus exotiques mais il convient de rester raisonnable. On évitera par exemple, le pourtant très juste,

$$\sum_{k=0}^{n} u_k = \sum_{\text{Bernardo}=0}^{n} u_{\text{Bernardo}}.$$

Proposition 3.4. Linéarité de la somme

Soit $u_1, v_1 \ldots, u_n, v_n, \lambda \in \mathbb{C}$. Alors,

$$\sum_{k=0}^{n} (u_k + v_k) = \sum_{k=0}^{n} u_k + \sum_{k=0}^{n} v_k.$$

$$\sum_{k=0}^{n} \lambda u_k = \lambda \sum_{k=0}^{n} u_k.$$

La première propriété est une conséquence de l'associativité et de la commutativité de l'addition. La seconde propriété signifie simplement que, par distributivité de la multiplication sur l'addition, on peut mettre en facteur un terme qui ne dépend pas de l'indice de sommation.

Conseils méthodologiques

On sort d'une somme (c'est-à-dire, on met en facteur) ce qui ne dépend pas de l'indice de sommation.

Exemple

Calculons la somme des écarts à la moyenne dans une suite finie de complexes $u_1, \dots, u_n$.

$$\begin{aligned}\sum_{k=1}^{n}\Big(u_k - \frac{1}{n}\sum_{i=1}^{n} u_i\Big) &= \sum_{k=1}^{n} u_k - \sum_{k=1}^{n}\Big(\frac{1}{n}\sum_{i=1}^{n} u_i\Big)\\ &= \sum_{k=1}^{n} u_k - n\frac{1}{n}\sum_{i=1}^{n} u_i = 0.\end{aligned}$$

On a utilisé la mise en facteur du terme $\frac{1}{n}\sum_{i=1}^{n} u_i$ qui ne dépend pas de k. Ainsi, la somme des écarts à la moyenne est nulle.

Proposition 3.5. Associativité du produit

Soit $u_0, v_0 \dots, u_n, v_n \in \mathbb{C}$. Alors,

$$\prod_{k=0}^{n}(u_k v_k) = \Big(\prod_{k=0}^{n} u_k\Big)\Big(\prod_{k=0}^{n} v_k\Big).$$

On notera le cas particulier suivant, pour $u_0, \dots, u_n, \lambda \in \mathbb{C}$,

$$\prod_{k=0}^{n}(\lambda u_k) = \lambda^{n+1}\prod_{k=0}^{n} u_k.$$

Il y a $n+1$ termes dans le produit donc λ sort avec l'exposant $n+1$.

Démonstration

▹ Si l'un des complexes considérés est nul, les deux membres sont nuls et il y a bien égalité.

▹ Sinon, posons, pour tout $k \in [\![1, n]\!]$, a_k et $b_k \in \mathbb{C}$ tels que $u_k = e^{a_k}$ et $v_k = e^{b_k}$ (ce qui est possible car $\exp : \mathbb{C} \to \mathbb{C}^*$ est surjective). Ainsi,

$$\begin{aligned}\prod_{k=0}^{n}(u_k v_k) &= \prod_{k=0}^{n}\exp(a_k + b_k) = \exp\Big(\sum_{k=0}^{n}(a_k + b_k)\Big)\\ &= \exp\Big(\sum_{k=0}^{n} a_k\Big)\exp\Big(\sum_{k=0}^{n} b_k\Big)\\ &= \Big(\prod_{k=0}^{n} u_k\Big)\Big(\prod_{k=0}^{n} v_k\Big).\end{aligned}$$

■

Dans la démonstration précédente, on s'est ramené à la proposition analogue pour les sommes. On pourra le faire pour la plupart des propositions suivantes donc on se contentera d'énoncer les résultats pour les sommes.

Exemple

▹ Le produit des nombres pairs entre 1 et $2n$ vaut

$$\prod_{\substack{k=1\\ k \text{ pair}}}^{2n} k = \prod_{l=1}^{n} (2l) = 2^n n!.$$

▹ Par ailleurs,

$$(2n)! = \prod_{k=1}^{2n} k = \Big(\prod_{\substack{k=1\\ k \text{ pair}}}^{2n} k\Big)\Big(\prod_{\substack{k=1\\ k \text{ impair}}}^{2n} k\Big).$$

Ainsi, le produit des nombres impairs de 1 à $2n$ vaut

$$\prod_{\substack{k=1\\ k \text{ impair}}}^{2n} k = \frac{(2n)!}{\displaystyle\prod_{\substack{k=1\\ k \text{ pair}}}^{2n} k} = \frac{(2n)!}{2^n n!}.$$

1.1. Associativité

On rappelle que l'associativité de l'addition permet de parenthéser à sa guise une somme de plusieurs termes. Combinée avec la commutativité, cela permet de regrouper notre somme « par paquets ». Voici un exemple d'une telle manipulation.

Proposition 3.6.

Soit $u_0, \dots, u_n \in \mathbb{C}$. Alors,

$$\sum_{k=0}^{n} u_k = \sum_{\substack{k=0\\ k \text{ pair}}}^{n} u_k + \sum_{\substack{k=0\\ k \text{ impair}}}^{n} u_k.$$

Cette manipulation correspond à partitionner l'ensemble $[\![0, n]\!]$ des indices de la somme de gauche en deux parties :

- celle des indices pairs $[\![0, n]\!] \cap 2\mathbb{N}$,
- celle des indices impairs $[\![0, n]\!] \cap (2\mathbb{N}+1)$.

Exemple

L'entier 111111 n'est pas premier : en effet, $111111 = 111 \times 10^3 + 111 = 111$. Généralisons cet exemple et montrons que l'entier n dont l'écriture décimale compte ab chiffres 1 avec $a, b > 1$ n'est pas premier. Par définition,

$$n = \sum_{k=0}^{ab-1} 10^k.$$

Séparons cette somme selon la valeur du quotient de l'indice par b, c'est-à-dire que l'on commence par calculer la somme sur les indices k dont le quotient par b est 0, puis sur les indices

dont le quotient par b est 1 et ainsi de suite.

$$\begin{aligned} n &= \sum_{q=0}^{a-1}\Big(\sum_{r=0}^{b-1} 10^{qb+r}\Big) \\ &= \sum_{q=0}^{a-1}\Big(10^{qb}\sum_{r=0}^{b-1} 10^{r}\Big) \\ &= \Big(\sum_{r=0}^{b-1} 10^{r}\Big)\Big(\sum_{q=0}^{a-1} 10^{qb}\Big). \end{aligned}$$

Ainsi, n est le produit de deux entiers différents de 1 ou -1 donc n'est pas premier.

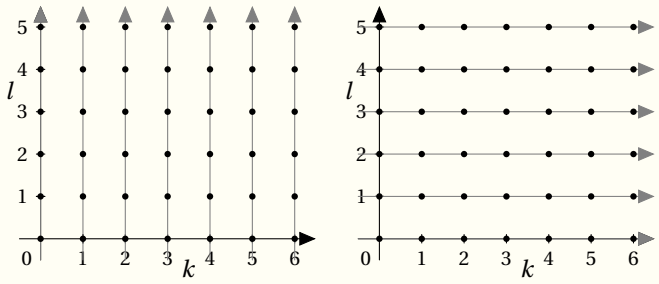
1.2. Sommes doubles

Proposition 3.7. Théorème de Fubini

Soit $(u_{k,l})_{k\leq p, l\leq q}$ une famille de complexes. Alors,

$$\sum_{k=0}^{p}\sum_{l=0}^{q} u_{k,l} = \sum_{l=0}^{q}\sum_{k=0}^{p} u_{k,l}.$$

Ceci revient à partitionner différemment l'ensemble $[\![0,p]\!]\times[\![0,q]\!]$ des couples d'indices (k,l) : d'un côté, selon la valeur du premier indice ; de l'autre, selon la valeur du second indice.



Corollaire 3.8.

Soit $u_0, \ldots, u_n, v_0, \ldots, v_m \in \mathbb{C}$. Alors,

$$\left(\sum_{k=0}^{n} u_k\right)\left(\sum_{k=0}^{m} v_k\right) = \sum_{k=0}^{n}\sum_{l=0}^{m} u_k v_l.$$

En particulier,

$$\left(\sum_{k=0}^{n} u_k\right)^2 = \sum_{k=0}^{n}\sum_{l=0}^{n} u_k u_l.$$

Remarque
Écrivons la formule du carré d'une somme dans les cas particuliers de 2 ou 3 termes :

$$\begin{aligned}(u_0+u_1)^2 &= u_0^2+u_0u_1+u_1u_0+u_1^2 = u_0^2+2u_0u_1+u_1^2,\\ (u_0+u_1+u_2)^2 &= u_0^2+u_0u_1+u_0u_2+u_1u_0+u_1^2+u_1u_2+u_2u_0+u_2u_1+u_2^2\\ &= u_0^2+u_1^2+u_2^2+2u_0u_1+2u_1u_2+2u_2u_0.\end{aligned}$$

1.3. Changement d'indice(s)

Il s'agit de faire un changement de variables entre deux ensembles d'entiers.

Proposition 3.9.

Soit I et J deux ensembles finis, $\varphi : J \to I$ une bijection et $(x_i)_{i\in I}$ une famille de complexes. Alors

$$\sum_{i\in I} u_i = \sum_{j\in J} u_{\varphi(j)}.$$

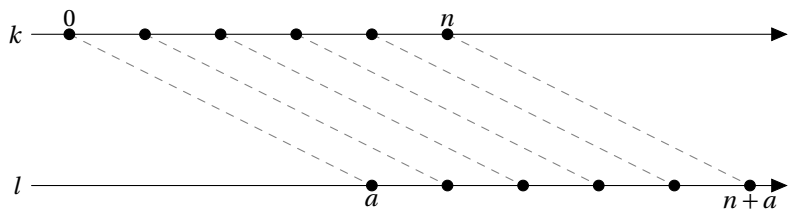
Le cas le plus fréquent est le cas d'un changement par « translation ».

Proposition 3.10.

Soit $u_0, \dots, u_n \in \mathbb{C}$. Alors,

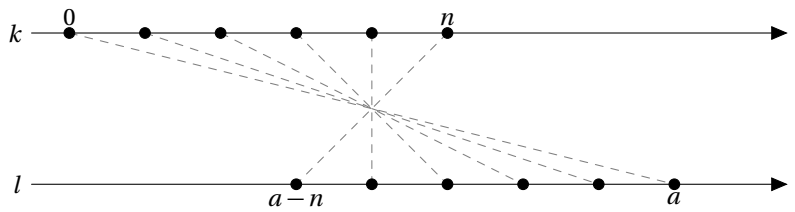
- En posant $l = k+a$ ou de manière équivalente $k = l-a$,

$$\sum_{k=0}^{n} u_k = \sum_{l=a}^{n+a} u_{l-a}.$$



- En posant $l = -k+a$ ou de manière équivalente $k = a-l$,

$$\sum_{k=0}^{n} u_k = \sum_{l=a-n}^{a} u_{a-l}.$$



Regardons comme une conséquence des propositions précédentes la propriété de télescopage.

Proposition 3.11. Somme télescopique

Soit $u_0, \dots, u_{n+1} \in \mathbb{C}$. Alors,

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

Démonstration

On utilise la linéarité puis on effectue un changement de variables par translation

$$\begin{aligned}\sum_{k=0}^{n}(u_{k+1}-u_k) &= \sum_{k=0}^{n}u_{k+1}-\sum_{k=0}^{n}u_k\\ &= \sum_{l=1}^{n+1}u_l-\sum_{k=0}^{n}u_k\\ &= \sum_{k=1}^{n+1}u_k-\sum_{k=0}^{n}u_k\\ &= \Big(u_{n+1}+\sum_{k=1}^{n}u_k\Big)-\Big(\sum_{k=1}^{n}u_k+u_0\Big) = u_{n+1}-u_0\end{aligned}$$

■

Remarques

Notons que $\frac{1}{k(k+1)} = \frac{1}{k} - \frac{1}{k+1}$. Par conséquent, la somme suivante est télescopique

$$\sum_{k=1}^{n}\frac{1}{k(k+1)} = \sum_{k=1}^{n}\Big(\frac{1}{k}-\frac{1}{k+1}\Big) = 1-\frac{1}{n+1}.$$

Remarquons également que $1+\frac{1}{k} = \frac{k+1}{k}$. Par conséquent, le produit suivant est télescopique

$$\prod_{k=1}^{n}\Big(1+\frac{1}{k}\Big) = \prod_{k=1}^{n}\frac{k+1}{k} = \frac{n+1}{1} = n+1.$$

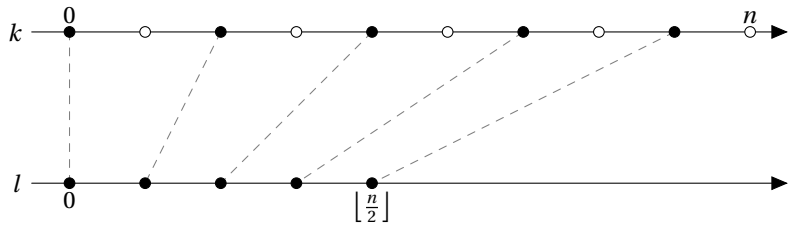
Voici un autre exemple utile de changement d'indices qu'il faut comprendre et savoir manipuler.

Exemple

Soit $u_0, \dots, u_n \in \mathbb{C}$. Alors,

$$\sum_{\substack{k=0\\ k \text{ pair}}}^{n} u_k = \sum_{l=0}^{\lfloor \frac{n}{2} \rfloor} u_{2l}, \qquad \sum_{\substack{k=0\\ k \text{ impair}}}^{n} u_k = \sum_{l=0}^{\lfloor \frac{n-1}{2} \rfloor} u_{2l+1},$$

où $\lfloor x \rfloor$ désigne la partie entière du réel x, c'est-à-dire le plus grand entier inférieur ou égal à x.



La véritable difficulté dans cette proposition consiste à trouver les bornes de la nouvelle somme. Par exemple, dans la seconde somme, on pose $k = 2l + 1$ et on effectue le petit raisonnement suivant :

$$\begin{aligned} 0 \leq k \leq n &\Leftrightarrow 0 \leq 2l + 1 \leq n \\ &\Leftrightarrow -\frac{1}{2} \leq l \leq \frac{n-1}{2}. \end{aligned}$$

Or, l est un entier donc la condition $0 \leq k \leq n$ est équivalente à $0 \leq l \leq \lfloor \frac{n-1}{2} \rfloor$.

Pour une somme double, le paradigme est le même pour le changement avec deux indices : dans le terme de droite, on doit retrouver chaque terme du membre de gauche une et une seule fois. Le cas le plus simple reste le théorème de Fubini déjà énoncé.

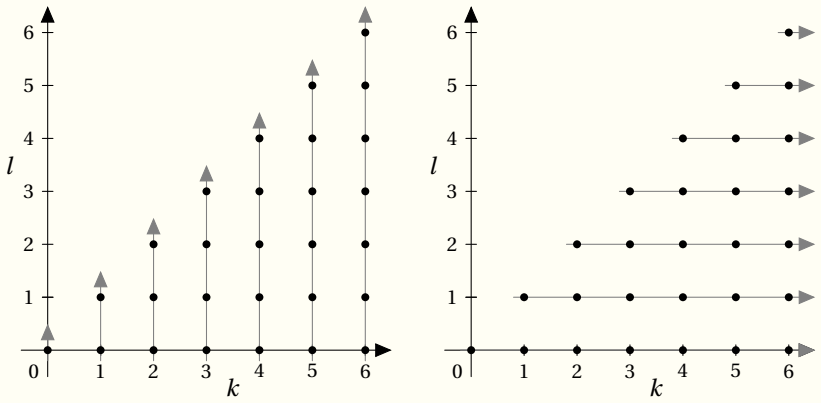
En choisissant une famille nulle en dehors du triangle d'indices (k, l) avec $0 \leq l \leq k \leq n$, on obtient le résultat suivant sur les sommes triangulaires.

Corollaire 3.12.

Soit $(u_{k,l})_{0 \leq l \leq k \leq n}$ une famille de complexes. Alors,

$$\sum_{k=0}^{n} \sum_{l=0}^{k} u_{k,l} = \sum_{l=0}^{n} \sum_{k=l}^{n} u_{k,l}.$$

D'un côté, on commence par « fixer » k puis l vérifie $0 \leq l \leq k$; de l'autre, on « fixe » l puis k vérifie $l \leq k \leq n$.



Exemple

Soit $u_0, u_1, \dots, u_n \in \mathbb{C}$. Reprenons l'exemple du carré d'une somme

$$\begin{aligned}\Big(\sum_{k=0}^{n} u_k\Big)^2 &= \sum_{k=0}^{n} u_k\Big(\sum_{l=0}^{k-1} u_l + u_k + \sum_{l=k+1}^{n} u_l\Big)\\ &= \sum_{k=0}^{n} u_k^2 + \sum_{k=0}^{n} u_k \sum_{l=0}^{k-1} u_l + \sum_{k=0}^{n}\sum_{l=k+1}^{n} u_k u_l\\ &= \sum_{k=0}^{n} u_k^2 + \sum_{k=0}^{n} u_k \sum_{l=0}^{k-1} u_l + \sum_{l=0}^{n}\sum_{k=0}^{l-1} u_k u_l\\ &= \sum_{k=0}^{n} u_k^2 + 2\sum_{k=0}^{n}\sum_{l=0}^{k-1} u_k u_l.\end{aligned}$$

On retrouve bien la construction de l'identité remarquable avec d'une part la somme des carrés et d'autre part les doubles produits.

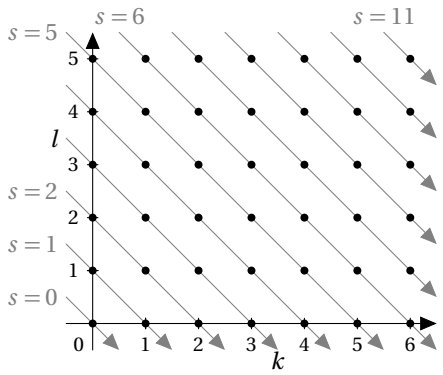
On peut toutefois effectuer d'autres manipulations.

Exemple

Soit $(u_{k,l})_{k\le p, l\le q}$ une famille de complexes. Pour calculer

$$\sum_{k=0}^{p}\sum_{l=0}^{q} u_{k,l}$$

on peut bien entendu partitionner l'ensemble des indices selon la valeur de $s = k+l$. Cette fois-ci, on partitionne l'ensemble des couples d'indices selon les « diagonales ».



Ainsi,

$$\sum_{k=0}^{p}\sum_{l=0}^{q} u_{k,l} = \sum_{k=0}^{p}\sum_{s=k}^{k+q} u_{k,s-k} = \sum_{s=0}^{p+q}\sum_{k=\max(0,s-q)}^{\min(p,s)} u_{k,s-k}.$$

Les bornes de la dernière somme se retrouvent en remarquant que

$$\begin{cases} 0\le k\le p\\ 0\le l\le q\end{cases} \iff \begin{cases} 0\le k\le p\\ 0\le s-k\le q\end{cases} \iff \begin{cases} 0\le k\le p\\ s-q\le k\le s\\ 0\le s\le p+q\end{cases}$$

2. Quelques sommes classiques

2.1. Sommes géométriques

Proposition 3.13. Somme géométrique

Soit $a \in \mathbb{C}$. Alors, pour tous entiers relatifs $p \leq q$

$$\sum_{k=p}^{q} a^k = \begin{cases} \dfrac{a^p - a^{q+1}}{1-a} & \text{si } a \neq 1; \\ q-p+1 & \text{sinon.} \end{cases}$$

On mémorise qu'une somme géométrique de raison différente de 1 est la différence du premier terme écrit et du premier terme non-écrit divisée par un moins la raison.

Démonstration

▹ Le résultat est évident pour $a = 1$ car il y a $q-p+1$ éléments dans $[\![p, q]\!]$ et le terme général de la somme est constant égal à 1.

▹ Si $a \neq 1$, on a, par télescopage,

$$\begin{aligned}(1-a)\sum_{k=p}^{q} a^k &= \sum_{k=p}^{q}(a^k - a^{k+1}) \\ &= a^p - a^{q+1}.\end{aligned}$$

Le résultat annoncé s'obtient en divisant par $1-a$. ■

Exemple

Effectuons quelques calculs sur les racines $n^{\text{ièmes}}$ de l'unité. Notons $\omega = e^{\frac{2i\pi}{n}}$.

▹ La somme des racines $n^{\text{ièmes}}$ de l'unité est nulle. En effet,

$$\sum_{z \in \mathbb{U}_n} z = \sum_{k=0}^{n-1} \omega^k = \frac{1-\omega^n}{1-\omega} = 0.$$

Ainsi, la moyenne des racines $n^{\text{ièmes}}$ de l'unité est nulle. En termes géométriques, l'isobarycentre ou centre de gravité des points d'affixes dans $\mathbb{U}_n$ (donc des sommets d'un polygone régulier inscrit dans le cercle unité) est l'origine.

▹ Reprenons ce calcul en ajoutant un exposant $m \in \mathbb{Z}$. Cette fois-ci, il s'agit d'une somme géométrique de raison ω^m :

$$\sum_{z \in \mathbb{U}_n} z^m = \sum_{k=0}^{n-1} (\omega^k)^m = \sum_{k=0}^{n-1} (\omega^m)^k = \begin{cases} n & \text{si } m = 0[n] \\ 0 & \text{sinon.} \end{cases}$$

Exemple

Calculons, pour tout $\theta \in \mathbb{R}$ et tout $n \in \mathbb{N} \setminus \{0\}$, la somme $\sum_{k=0}^{n} \cos(k\theta)$. Pour cela, remarquons que le cosinus est la partie réelle de l'exponentielle complexe puis utilisons que la partie réelle

est linéaire.

$$\sum_{k=0}^{n}\cos(k\theta)=\mathfrak{Re}\Big(\sum_{k=0}^{n}e^{ik\theta}\Big)=\mathfrak{Re}\Big(\sum_{k=0}^{n}(e^{i\theta})^k\Big).$$

On est ainsi ramené à calculer la somme géométrique de raison $e^{i\theta}$.
▹ Si $e^{i\theta}=1$, c'est-à-dire si $\theta\in 2\pi\mathbb{Z}$, la somme recherchée vaut $n+1$ (car chacun des $n+1$ cosinus de la somme initiale vaut 1) ;
▹ sinon,

$$\begin{aligned}\sum_{k=0}^{n}\cos(k\theta)&=\mathfrak{Re}\Big(\frac{1-e^{i(n+1)\theta}}{1-e^{i\theta}}\Big)\\&=\mathfrak{Re}\Big(\frac{-2i\sin\frac{(n+1)\theta}{2}}{-2i\sin\frac{\theta}{2}}e^{i\frac{n\theta}{2}}\Big)\\&=\frac{\sin\frac{(n+1)\theta}{2}}{\sin\frac{\theta}{2}}\mathfrak{Re}\Big(e^{i\frac{n\theta}{2}}\Big)\\&=\frac{\sin\frac{(n+1)\theta}{2}}{\sin\frac{\theta}{2}}\cos\frac{n\theta}{2}.\end{aligned}$$

Exemple Noyau de Dirichlet

Calculons, pour tout $n\in\mathbb{N}\setminus\{0\}$ et tout $\theta\in\mathbb{R}\setminus 2\pi\mathbb{Z}$, $D_n(\theta)=\sum\limits_{k=-n}^{n}e^{ik\theta}$.

$$\begin{aligned}D_n(\theta)&=\frac{e^{-in\theta}-e^{i(n+1)\theta}}{1-e^{i\theta}}\\&=\frac{e^{i\frac{\theta}{2}}(e^{-i(n+\frac{1}{2})\theta}-e^{i(n+\frac{1}{2})\theta})}{e^{i\frac{\theta}{2}}(e^{-i\frac{\theta}{2}}-e^{i\frac{\theta}{2}})}\\&=\frac{-2i\sin\big((n+\frac{1}{2})\theta\big)}{-2i\sin\big(\frac{\theta}{2}\big)}=\frac{\sin\big((n+\frac{1}{2})\theta\big)}{\sin\big(\frac{\theta}{2}\big)}.\end{aligned}$$

On remarque que l'on a utilisé la méthode canonique pour simplifier la somme/différence de deux complexes de module 1, qui consiste à factoriser l'exponentielle de la demi-somme des arguments.

2.2. Formule de Bernoulli

Généralisons le résultat des sommes géométriques (qui se retrouve pour $b=1$).

Proposition 3.14. Formule de Bernoulli

Pour tous a, $b\in\mathbb{C}$ et tout $n\in\mathbb{N}$,

$$a^{n+1}-b^{n+1}=(a-b)\sum_{k=0}^{n}a^k b^{n-k}.$$

Démonstration
Partons du second membre et développons

$$\begin{aligned}(a-b)\sum_{k=0}^{n} a^k b^{n-k} &= \sum_{k=0}^{n} a^{k+1} b^{n-k} - \sum_{k=0}^{n} a^k b^{n+1-k} \\ &= \sum_{l=1}^{n+1} a^l b^{n+1-l} - \sum_{k=0}^{n} a^k b^{n+1-k} \\ &= \Big(a^{n+1} + \sum_{l=1}^{n} a^l b^{n+1-l}\Big) - \Big(\sum_{k=1}^{n} a^k b^{n+1-k} + b^{n+1}\Big) \\ &= a^{n+1} - b^{n+1}.\end{aligned}$$

■

Exemple

Montrons que si l'entier n est composé, alors $2^n - 1$ l'est également. Pour cela, notons $n = ab$ avec $a, b > 1$ et d'après la formule de Bernoulli,

$$\begin{aligned}2^n - 1 &= (2^b)^a - 1^a \\ &= (2^b - 1)\sum_{k=0}^{a-1} 2^{kb}.\end{aligned}$$

Ainsi, $2^n - 1$ est composé.

2.3. Sommes géométriques dérivées

Proposition 3.15.

Pour tout $a \in \mathbb{C} \setminus \{1\}$ et tout $n \in \mathbb{N}$,

$$\sum_{k=0}^{n} k a^{k-1} = \frac{1-(n+1)a^n + n a^{n+1}}{(1-a)^2}.$$

Démonstration
Considérons l'application

$$f : \begin{cases} \mathbb{R} \setminus \{1\} & \to & \mathbb{R} \\ x & \mapsto & \sum_{k=0}^{n} x^k \end{cases}$$

D'après la somme géométrique, pour tout $x \neq 1$, $f(x) = \frac{1-x^{n+1}}{1-x}$. En dérivant cette relation, on obtient l'égalité pour $x \in \mathbb{R} \setminus \{1\}$

$$\sum_{k=0}^{n} k x^{k-1} = \frac{1-(n+1)x^n + n x^{n+1}}{(1-x)^2}.$$

Par conséquent, l'équation polynomiale

$$(1-x)^2 \sum_{k=0}^{n} k x^{k-1} - (1-(n+1)x^n + n x^{n+1}) = 0$$

admet une infinité de solutions (tous les réels) donc est satisfaite sur tout $\mathbb{C}$ (une équation polynomiale non nulle de degré au plus $n+1$ admet au plus $n+1$ solutions). Par conséquent, l'égalité est vérifiée pour tout $a \in \mathbb{C} \setminus \{1\}$ en divisant par $a-1$. ■

Exemple

Notons encore $\omega = e^{\frac{2i\pi}{n}}$. Alors, la formule de somme géométrique dérivée donne

$$\begin{aligned}\sum_{k=0}^{n-1} k\omega^{k-1} &= \frac{1-n\omega^{n-1}+(n-1)\omega^n}{(1-\omega)^2}\\ &= \frac{n-n\omega^{-1}}{(1-\omega)^2} = \frac{n\omega^{-1}(\omega-1)}{(1-\omega)^2} = \frac{n}{\omega(\omega-1)}.\end{aligned}$$

Il peut être tentant de faire tendre a vers 1 dans la formule des sommes géométriques dérivées. C'est tout à fait possible et on obtient alors la formule suivante par un argument de continuité.

Proposition 3.16.

Pour tout $n \in \mathbb{N}$,

$$\sum_{k=0}^{n} k = \frac{n(n+1)}{2}.$$

Voici une démonstration élémentaire (une récurrence serait tout aussi simple).

Démonstration

$$\begin{aligned}2\sum_{k=0}^{n} k &= \sum_{k=0}^{n} k + \sum_{k=0}^{n} k\\ &= \sum_{k=0}^{n} k + \sum_{l=0}^{n} (n-l) \qquad \text{avec } l = n-k\\ &= \sum_{k=0}^{n} k + \sum_{k=0}^{n} (n-k)\\ &= \sum_{k=0}^{n} n = n(n+1).\end{aligned}$$

■

Par une simple récurrence (puisque l'on connaît la forme du résultat), on obtient les égalités suivantes.

Proposition 3.17.

Pour tout $n \in \mathbb{N}$,

$$\sum_{k=0}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}, \qquad \sum_{k=0}^{n} k^3 = \left(\frac{n(n+1)}{2}\right)^2.$$

3. Formule du binôme

3.1. Coefficients binomiaux

Définition 3.18.

Le coefficient binomial $\binom{n}{p}$ est l'entier $\frac{n!}{p!(n-p)!}$ si $p \in [\![0, n]\!]$, 0 sinon.

Remarque

Comme on le verra dans un chapitre ultérieur, le coefficient binomial $\binom{n}{p}$ est le nombre de parties à p éléments d'un ensemble de cardinal n. Cette interprétation combinatoire est fort riche et permet de revisiter la plupart des résultats évoqués ici de manière plus élégante.

Remarque

Remarquons que la définition donne directement $\binom{n}{n} = \binom{n}{0} = 1$ pour tout $n \in \mathbb{N} \setminus \{0\}$ et la propriété de symétrie $\binom{n}{p} = \binom{n}{n-p}$ pour tous $0 \leq p \leq n$.

Proposition 3.19. Pascal

Pour tous $0 \leq p \leq n$,

$$\binom{n}{p} + \binom{n}{p+1} = \binom{n+1}{p+1}.$$

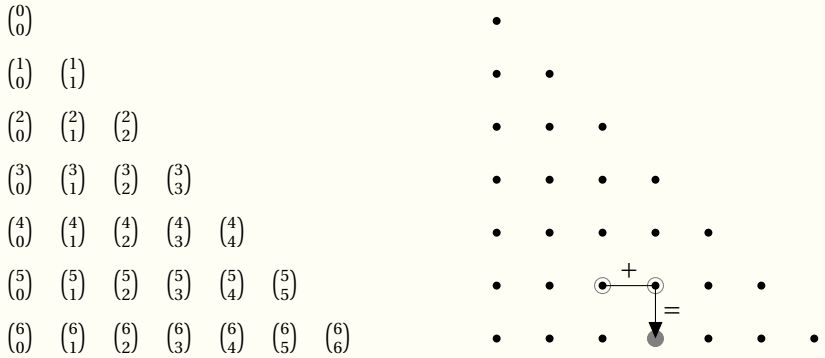
Démonstration

Fixons n et p tels que $0 \leq p \leq n$ et utilisons la définition

$$\begin{aligned}
\binom{n}{p} + \binom{n}{p+1} &= \frac{n!}{p!(n-p)!} + \frac{n!}{(p+1)!(n-p-1)!} \\
&= \frac{n!}{(p+1)!(n-p)!}(p+1+n-p) \\
&= \frac{(n+1)!}{(p+1)!(n-p)!} \\
&= \binom{n+1}{p+1}.
\end{aligned}$$

■

On représente régulièrement les coefficients binomiaux sous la forme d'un tableau à deux entrées : le coefficient $\binom{n}{p}$ se lit à l'intersection de la ligne n et de la colonne p. Les coefficients non nuls forment alors un triangle appelé triangle de Pascal. La formule précédente se visualise bien sur ce triangle comme indiqué ci-dessous.



Exemple

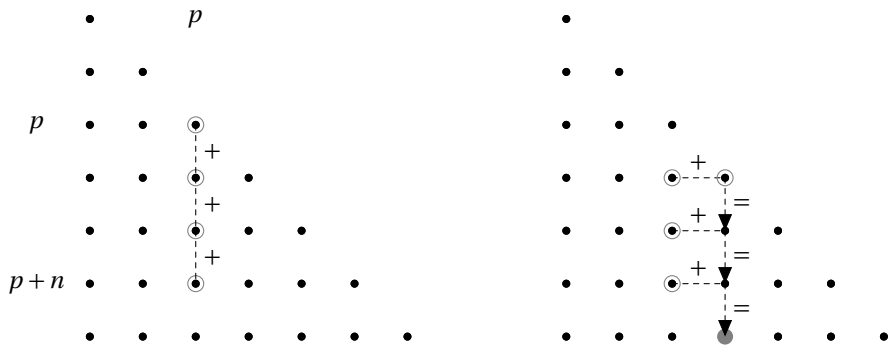
Montrons, par récurrence sur n, que

$$\forall p \in \mathbb{N}, \qquad \sum_{k=0}^{n} \binom{p+k}{p} = \binom{p+n+1}{p+1}.$$

▹ Le résultat est évident pour $n = 0$ car $\binom{p}{p} = \binom{p+1}{p+1}$ pour tout $p \in \mathbb{N}$.
▹ Soit $n \in \mathbb{N}$. Supposons que l'égalité est vérifiée au rang n et montrons-la au rang $n+1$; pour tout $p \in \mathbb{N}$,

$$\begin{aligned}\sum_{k=0}^{n+1} \binom{p+k}{p} &= \sum_{k=0}^{n} \binom{p+k}{p} + \binom{p+n+1}{p} \\ &= \binom{p+n+1}{p+1} + \binom{p+n+1}{p} = \binom{p+n+2}{p+1}.\end{aligned}$$

D'où le résultat annoncé.
On peut aussi tenter de comprendre rapidement cette identité avec le triangle de Pascal comme illustré sur le dessin suivant.



Remarquons par ailleurs que pour $p = 1$, cette identité n'est que la relation

$$\sum_{k=0}^{n} (k+1) = \frac{(n+2)(n+1)}{2}.$$

3.2. Formule de Newton

Proposition 3.20. Formule du binôme de Newton

Pour tous a, $b \in \mathbb{C}$ et $n \in \mathbb{N}$, $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$.

Démonstration

Fixons a et $b \in \mathbb{C}$ et montrons la propriété par récurrence.

▹ L'initialisation est immédiate car $(a+b)^0 = 1$.

▹ Soit $n \in \mathbb{N}$. Supposons la propriété vérifiée au rang n. Alors, avec la formule de Pascal,

$$\begin{aligned}
(a+b)^{n+1} &= (a+b)\sum_{k=0}^{n}\binom{n}{k}a^k b^{n-k} \\
&= \sum_{k=0}^{n}\binom{n}{k}a^{k+1} b^{n-k} + \sum_{k=0}^{n}\binom{n}{k}a^k b^{n-k+1} \\
&= \sum_{k=1}^{n+1}\binom{n}{k-1}a^k b^{n-k+1} + \sum_{k=0}^{n}\binom{n}{k}a^k b^{n-k+1} \\
&= \binom{n}{n}a^{n+1} b^0 + \sum_{k=1}^{n}\binom{n}{k-1}a^k b^{n-k+1} + \sum_{k=1}^{n}\binom{n}{k}a^k b^{n-k+1} + \binom{n}{0}a^0 b^{n+1} \\
&= a^{n+1} + \sum_{k=1}^{n}\left(\binom{n}{k-1}+\binom{n}{k}\right)a^k b^{n-k+1} + b^{n+1} \\
&= a^{n+1} + \sum_{k=1}^{n}\binom{n+1}{k}a^k b^{n-k+1} + b^{n+1} \\
&= \sum_{k=0}^{n+1}\binom{n+1}{k}a^k b^{n+1-k}.
\end{aligned}$$

d'où le résultat. ■

Conseils méthodologiques

Remarquons le rôle symétrique des exposants dans la formule du binôme. En effet,

$$\begin{aligned}
\sum_{k=0}^{n}\binom{n}{k}a^k b^{n-k} &= \sum_{l=0}^{n}\binom{n}{n-l}a^{n-l} b^l \qquad \text{avec } l = n-k \\
&= \sum_{l=0}^{n}\binom{n}{l}a^{n-l} b^l
\end{aligned}$$

avec la propriété de symétrie des coefficients binomiaux. Ainsi, il est pertinent de choisir le terme sur lequel on met l'exposant « compliqué » $n-k$.

En choisissant $b = 1$ dans la formule du binôme, on obtient le cas particulier suivant.

Corollaire 3.21.

Pour tout $a \in \mathbb{C}$ et tout $n \in \mathbb{N}$,

$$\sum_{k=0}^{n} \binom{n}{k} a^k = (1+a)^n.$$

En particularisant encore, on obtient le résultat suivant qui sera revisité lors du chapitre de dénombrement.

Corollaire 3.22.

Pour tout $n \in \mathbb{N}$,

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n.$$

Exemple

D'après la formule du binôme,

$$\begin{aligned}
\sum_{\substack{k=0 \\ k \text{ pair}}}^{n} \binom{n}{k} + \sum_{\substack{k=0 \\ k \text{ impair}}}^{n} \binom{n}{k} &= \sum_{k=0}^{n} \binom{n}{k} \\
&= \sum_{k=0}^{n} \binom{n}{k} 1^k 1^{n-k} = (1+1)^n = 2^n, \\
\sum_{\substack{k=0 \\ k \text{ pair}}}^{n} \binom{n}{k} - \sum_{\substack{k=0 \\ k \text{ impair}}}^{n} \binom{n}{k} &= \sum_{k=0}^{n} \binom{n}{k} (-1)^k \\
&= \sum_{k=0}^{n} \binom{n}{k} (-1)^k 1^{n-k} = (1-1)^n = 0.
\end{aligned}$$

En additionnant puis en soustrayant ces deux égalités, on obtient que

$$\sum_{\substack{k=0 \\ k \text{ pair}}}^{n} \binom{n}{k} = \sum_{\substack{k=0 \\ k \text{ impair}}}^{n} \binom{n}{k} = 2^{n-1}.$$

Remarque

On peut obtenir ce résultat d'une autre manière en utilisant astucieusement la formule de Pascal.

$$\begin{aligned}
\sum_{\substack{k=0 \\ k \text{ pair}}}^{\infty} \binom{n}{k} &= \sum_{k=0}^{\infty} \binom{n}{2k} \\
&= \sum_{k=0}^{\infty} \left(\binom{n-1}{2k} + \binom{n-1}{2k-1} \right) \\
&= \sum_{k=0}^{\infty} \binom{n-1}{k} = 2^{n-1}.
\end{aligned}$$

Pour simplifier la lecture, on a ici noté les sommes jusqu'en $+\infty$ en utilisant que les coefficients binomiaux $\binom{n}{k}$ sont nuls si $k \notin [\![0, n]\!]$.

Exemple Polynômes de Tchebychev

Pour tout $r \in \mathbb{N}$ et tout $\theta \in \mathbb{R}$,

$$\cos(r\theta) = \mathfrak{Re}\big(((e^{i\theta})^r\big) = \mathfrak{Re}((\cos\theta + i\sin\theta)^r\big) = \sum_{k=0}^{r}\binom{r}{k}\mathfrak{Re}\big((i\sin\theta)^k\big)\cos^{r-k}\theta$$

$$= \sum_{k=0}^{\lfloor \frac{r}{2}\rfloor}\binom{r}{2k}(-1)^k(\sin\theta)^{2k}\cos^{r-2k}\theta = \sum_{k=0}^{\lfloor \frac{r}{2}\rfloor}\binom{r}{2k}(-1)^k(1-\cos^2\theta)^k\cos^{r-2k}\theta$$

Ainsi, pour tout $r \in \mathbb{N}$, le polynôme

$$T_r(X) = \sum_{k=0}^{\lfloor \frac{r}{2}\rfloor}\binom{r}{2k}(-1)^k(1-X^2)^k X^{r-2k}$$

vérifie

$$\forall \theta \in \mathbb{R}, \qquad \cos r\theta = T_r(\cos\theta)$$

On peut facilement montrer que $\deg T_r = r$ et que le coefficient dominant est $\sum\limits_{\substack{k=0\\ k \text{ pair}}}^{n}\binom{n}{k}2^{r-1}$.

En dérivant cette relation par rapport à θ, on a également

$$\forall \theta \in \mathbb{R}, \qquad \sin r\theta = \sin\theta\,\frac{1}{r}T_r'(\cos\theta).$$

En dérivant la formule du binôme réelle puis en l'étendant aux complexes comme précédemment, on obtient le calcul de somme suivant.

Proposition 3.23.

Pour tout $a \in \mathbb{C}$,

$$\sum_{k=1}^{n}\binom{n}{k}ka^{k-1} = n(1+a)^{n-1}.$$

On aurait également pu démontrer cette formule en utilisant la formule suivante entre coefficients binomiaux $k\binom{n}{k} = n\binom{n-1}{k-1}$.

Démonstration

Avec cette formule, on se ramène à la formule du binôme :

$$\begin{aligned}\sum_{k=1}^{n}\binom{n}{k}ka^{k-1} &= n\sum_{k=1}^{n}\binom{n-1}{k-1}a^{k-1}\\ &= n\sum_{k=0}^{n-1}\binom{n-1}{k}a^{k}\\ &= n(1+a)^{n-1}.\end{aligned}$$

■

3.3. Sommes de puissances

Notation

Notons, pour tout $l \in \mathbb{N}$ et tout $n \geq 1$, $S_n(l) = \sum_{k=1}^{n} k^l$.

On a déjà donné les valeurs de $S_n(1)$, $S_n(2)$ et $S_n(3)$. On cherche désormais une relation de récurrence pour la suite $(S_n(l))_l$ (à n fixé).

Proposition 3.24.

Pour tout $p \geq 1$,

$$\sum_{l=1}^{p} \binom{p+1}{l} S_n(l) = (n+1)^{p+1} - n - 1.$$

Démonstration

En réécrivant la définition de $S_n(l)$, on obtient

$$\begin{aligned}
\sum_{l=1}^{p} \binom{p+1}{l} S_n(l) &= \sum_{l=1}^{p} \binom{p+1}{l} \sum_{k=1}^{n} k^l \\
&= \sum_{k=1}^{n} \sum_{l=1}^{p} \binom{p+1}{l} k^l \\
&= \sum_{k=1}^{n} \left((k+1)^{p+1} - k^{p+1} - 1\right) \qquad \text{(formule du binôme)} \\
&= \sum_{k=1}^{n} \left((k+1)^{p+1} - k^{p+1}\right) - n \\
&= (n+1)^{p+1} - 1 - n
\end{aligned}$$

par télescopage dans la dernière somme. ■

Exemple

▹ Pour $p = 1$, on obtient $\binom{2}{1} S_n(1) = (n+1)^2 - n - 1$ soit $S_n(1) = \frac{n(n+1)}{2}$.
▹ Pour $p = 2$, on obtient $\binom{3}{1} S_n(1) + \binom{3}{2} S_n(2) = (n+1)^3 - n - 1$ soit, en remplaçant $S_n(1)$,

$$3\frac{n(n+1)}{2} + 3S_n(2) = n^3 + 3n^2 + 2n$$

donc $S_n(2) = \frac{n(n+1)(2n+1)}{6}$.

Remarque

Cette relation permet de calculer les $S_n(l)$ de proche en proche. Pour la retrouver sans avoir à retenir le premier membre, on peut également procéder dans le sens contraire de cette démonstration : partir de la somme télescopique de terme général $(k+1)^{p+1} - k^{p+1}$ et développer le premier terme par la formule du binôme.

FICHE

SYNTHÈSE

Les manipulations sur les sommes et produits finis sont assez naturelles et il est important de savoir les mettre en œuvre. On se ramène souvent à des sommes connues.

Somme de puissances

Pour tout $n \in \mathbb{N}$,

$$\sum_{k=0}^{n} k = \frac{n(n+1)}{2}, \qquad \sum_{k=0}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}.$$

Somme géométrique

Pour tous a et $n \in \mathbb{N}$,

$$\sum_{k=0}^{n} a^k = \begin{cases} \dfrac{1-a^{n+1}}{1-a} & \text{si } a \neq 1, \\ n+1 & \text{sinon.} \end{cases}$$

En particulier, pour tous a et $b \in \mathbb{C}$,

$$a^{n+1} - b^{n+1} = (a-b)\sum_{k=0}^{n} a^k b^{n-k}.$$

On retiendra en particulier l'application de ces formules aux racines $n^{\text{ièmes}}$ de l'unité.

Somme de racines $n^{\text{ièmes}}$

Pour tous entiers m et n,

$$\sum_{z \in \mathbb{U}_n} z^m = \begin{cases} n & \text{si } m = 0[n], \\ 0 & \text{sinon.} \end{cases}$$

Pour tout $z \in \mathbb{U}_n$,

$$\sum_{k=0}^{n-1} z^k = \begin{cases} n & \text{si } z = 1 \\ 0 & \text{sinon.} \end{cases}$$

La généralisation de l'identité remarquable $(a+b)^2 = a^2 + 2ab + b^2$ est la formule du binôme.

Formule du binôme de Newton

Pour tous $a,\ b \in \mathbb{C}$ et $n \in \mathbb{N}$,

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k} = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k.$$

En particulier,

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n, \qquad \sum_{k=0}^{n} \binom{n}{k} (-1)^k = 0.$$

Il est nécessaire pour faire des calculs avec la formule du binôme de bien connaître les propriétés des coefficients binomiaux.

Coefficients binomiaux

Pour tous $p \leq n \in \mathbb{N}$,

$$\binom{n}{p} = \frac{n!}{p!(n-p)!}.$$

$$\binom{n}{p} = \binom{n}{n-p}, \qquad \binom{n+1}{p+1} = \binom{n}{p} + \binom{n}{p+1}.$$

EXERCICES

Vrai ou faux ?

	Vrai	Faux
a) $\sum_{k=-1789}^{1789} 1 = 2 \times 1789.$	☐	☐
b) $\sum_{\substack{k=0 \\ k \text{pair}}}^{n} 1 = \lfloor \frac{n}{2} \rfloor.$	☐	☐
c) $\prod_{\substack{k=1 \\ k \text{pair}}}^{2n} k = \frac{(2n)!}{n!}.$	☐	☐
d) Pour $n > m$, $\sum_{k=m}^{n} 2^k = 2^{n+1} - 2^m.$	☐	☐
e) Si $\omega \in \mathbb{U}_n \setminus \{1\}$, $\sum_{k=0}^{n-1} \omega^k = 0.$	☐	☐
f) Soit $k \in \mathbb{Z}$. $\sum_{\omega \in \mathbb{U}_n} \omega^k = 0.$	☐	☐
g) Le coefficient de x^k dans $\sum_{i=0}^{n} a_i x^i \sum_{j=0}^{m} b_j x^j$ est $\sum_{i=0}^{k} a_i b_{k-i}.$	☐	☐
h) Pour $0 < p < n$, $p\binom{n}{p} = n\binom{n-1}{p-1}.$	☐	☐

EXERCICES

Exercices d'application du chapitre 3

●○○ **Exercice 1**

Calculer la somme $\sum_{k=0}^{n} k(3k+1)$.

●○○ **Exercice 2**

Calculer les sommes $\sum_{k=0}^{n} \frac{3}{2^{5k+1}}$ et $\sum_{k=0}^{n} \frac{2^{k-1}}{3^{k+1}}$.

●○○ **Exercice 3**

Calculer par télescopage la somme $\sum_{k=0}^{n} \frac{k}{(k+1)!}$.

●○○ **Exercice 4**

Calculer la somme $\sum_{k=1}^{n} \frac{1}{\sqrt{k+1}+\sqrt{k}}$.

●●○ **Exercice 5**

Calculer la somme $\sum_{k=0}^{n}(-1)^k k$.

●●● **Exercice 6**

Calculer la somme $S_0 = \sum_{\substack{k=0 \\ k=0[3]}}^{n} \binom{n}{k}$.

On pourra calculer $(1+1)^n$, $(1+j)^n$ *et* $(1+j^2)^n$.

●●○ **Exercice 7**

Calculer le produit $\prod_{k=1}^{n} \frac{2k-1}{2k}$ avec des factorielles.

●○○ **Exercice 8**

Calculer, pour tout $\theta \in \mathbb{R}$, la somme $\sum_{k=0}^{n} \binom{n}{k} \cos(k\theta)$.

●●○ **Exercice 9**

Calculer la somme $\sum_{k=0}^{n} k^2 \binom{n}{k}$.

●●○ **Exercice 10**

Calculer la somme $\sum_{k=0}^{n} \frac{1}{(n-k)!(k+1)!}$.

●●○ **Exercice 11**

Soit $n \in \mathbb{N} \setminus \{0\}$, $\theta \in \mathbb{R} \setminus 2^n \pi \mathbb{Z}$. Calculer le produit $P_n = \prod_{k=1}^{n} \cos \frac{\theta}{2^k}$.

●○○ **Exercice 12**

Calculer la somme double $\sum_{l=1}^{n} \sum_{k=l+1}^{n} \frac{l}{k}$.

●○○ **Exercice 13**

Calculer la somme double $\sum_{k=0}^{n} \sum_{l=0}^{n} \min(k, l)$.

●●○ **Exercice 14**

Soit $n \in \mathbb{N} \setminus \{0\}$ et $\omega = e^{\frac{2i\pi}{n}}$. Montrer que, pour tout complexe z, $\sum_{k=1}^{n} (z + \omega^k)^n = n(z^n + 1)$.

Exercices d'approfondissement du chapitre 3

●○○ **Exercice A**

Soit $n \geq 2$ un entier. Posons, pour tout $p \leq n$, $S_p = \sum_{k=0}^{p} \binom{n}{k}$.

1. Montrer que pour tout $p \in [\![0, n-1]\!]$, $S_p + S_{n-p-1} = 2^n$.
2. En déduire $\sum_{p=0}^{n-1} S_p$.

●●○ **Exercice B**

Soit $(u_n)_n$ la suite définie par $u_1 = 1$ et, pour tout $n \in \mathbb{N}^*$,

$$u_{2n} = u_n, \qquad u_{2n+1} = (-1)^n u_n.$$

Calculer, pour tout $n \in \mathbb{N}$, $\sum_{k=1}^{4n} u_k u_{k+2}$.

●●○ **Exercice C**

Soit $x_1, x_2, \ldots, x_n \in \mathbb{R}$ et $x_{n+1} = x_1$. Montrer que $\sum_{k=1}^{n} x_k x_{k+1} \leq \sum_{k=1}^{n} x_k^2$.

●●● **Exercice D**

Soit $x_1 \leq \ldots \leq x_n$ et $y_1 \leq \ldots \leq y_n$. Montrer que

$$\Big(\frac{1}{n}\sum_{k=1}^{n} x_k\Big)\Big(\frac{1}{n}\sum_{k=1}^{n} y_k\Big) \leq \frac{1}{n}\sum_{k=1}^{n} x_k y_k.$$

●●● **Exercice E**

Soit $n \in \mathbb{N} \setminus \{0\}$.

1. Rappeler, selon les valeurs de $k \in \mathbb{N}$, $A_k = \sum_{z \in \mathbb{U}_n} z^k$.
2. Soit $P : x \mapsto \sum_{k=0}^{n-1} a_k x^k$ une application polynomiale à coefficients complexes $a_0, \ldots, a_{n-1}$ et

$$M = \max\{|P(z)|,\ z \in \mathbb{U}_n\}.$$

Montrer que, pour tout $k \leq n-1$, $|a_k| \leq M$.

●●● **Exercice F**

Soit $a_0 > a_1 > \ldots > a_n > 0$ et z un complexe tel que $\sum_{k=0}^{n} a_k z^k = 0$. Montrer que $|z| \geq 1$.

On pourra minorer le module de $(1-z)\sum_{k=0}^{n} a_k z^k$.

●●● **Exercice G**

Soit $n \in \mathbb{N} \setminus \{0\}$, $\omega = e^{\frac{2i\pi}{n}}$ et $Z = \sum_{k=0}^{n-1} \omega^{k^2}$. Calculer $|Z|^2$.

Problèmes du chapitre 3

●○○ **Problème 1**

Soit $n \geq 2$ un entier. Un n-uplet $(z_0, \ldots, z_{n-1}) \in \mathbb{C}^n$ satisfait la propriété $\mathscr{P}(n)$ si $z_0 = 1$ et, pour tout exposant $k \in [\![1, n-1]\!]$,

$$\sum_{j=0}^{n-1} z_j^k = 0.$$

1. Déterminer tous les couples de complexes satisfaisant la propriété $\mathscr{P}(2)$.
2. Soit $(z_0, z_1, z_2) \in \mathbb{C}^3$.

PROBLÈMES

a. Exprimer la quantité $z_0z_1+z_1z_2+z_2z_0$ à l'aide de $z_0+z_1+z_2$ et $z_0^2+z_1^2+z_2^2$.
b. En déduire que, si (z_0,z_1,z_2) satisfait $\mathscr{P}(3)$, alors z_0,z_1,z_2 sont les racines d'un polynôme de degré 3 que l'on explicitera.
c. Déterminer tous les triplets satisfaisant $\mathscr{P}(3)$.

3. Soit $(z_0,\dots,z_3)$ satisfaisant $\mathscr{P}(4)$. Expliciter, en utilisant les relations entre coefficients et racines, un polynôme de degré 3 dont les racines sont exactement z_1, z_2, z_3.
En déduire tous les 4-uplets satisfaisant $\mathscr{P}(4)$.
4. Soit $n\geq 2$.
a. Montrer que si $(z_0,\dots,z_{n-1})\in\mathbb{C}^n$ satisfait $\mathscr{P}(n)$, alors le n-uplet des conjugués $(\overline{z_0},\dots,\overline{z_{n-1}})$ satisfait aussi $\mathscr{P}(n)$.
b. Soit $(z_0,\dots,z_{n-1})\in\mathbb{C}^n$ qui satisfait $\mathscr{P}(n)$. Notons $z'_0=1$, $z'_{n-1}=-1$ et, pour tout $k\in[\![1,n]\!]$, z'_k et z'_{n-1+k} les racines de X^2-z_k.
Montrer que $(z'_0,\dots,z'_{2n-1})\in\mathbb{C}^{2n}$ satisfait $\mathscr{P}(2n)$.

●●○ **Problème 2 Représentation de Zeckendorf**

On considère la suite de Fibonacci $(u_n)_n$ définie par $u_0=1$, $u_1=2$ et la relation de récurrence

$$\forall n\in\mathbb{N},\qquad u_{n+2}=u_{n+1}+u_n.$$

Une représentation de Zeckendorf de l'entier $m\in\mathbb{N}$ est une suite finie $(a_k)_{k\leq N}\in\{0,1\}^{N+1}$ telle que $a_N=1$, pour tout $k<N$, $a_ka_{k+1}=0$, et

$$m=\sum_{k=0}^{N}a_ku_k.$$

On précise les valeurs suivantes

n	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
u_n	1	2	3	5	8	13	21	34	55	89	144	233	377	610	987	1597

1. Montrer que, pour tout $n\in\mathbb{N}$,
$$\sum_{k=0}^{n}u_k=u_{n+2}-2.$$
2. Posons, pour tout $n\in\mathbb{N}$,
$$\sigma_n=\sum_{k=0}^{\lfloor\frac{n}{2}\rfloor}u_{n-2k}.$$
a. Trouver une relation reliant σ_{n+2}, σ_n et u_{n+2}.
b. Montrer que, pour tout $n\in\mathbb{N}$, $\sigma_n=u_{n+1}-1$.
3. Supposons que $m\in\mathbb{N}$ admette une représentation de Zeckendorf $(a_k)_{k\leq N}$.
a. Montrer que $m\leq\sigma_N$.
b. En déduire que u_N est le plus grand terme de la suite $(u_n)_{n\in\mathbb{N}}$ inférieur ou égal à m.
4. Montrer que tout entier non nul admet une unique représentation de Zeckendorf.
5. a. Déterminer la représentation de Zeckendorf de 2016 (en précisant vos étapes de calculs).
b. Déterminer la représentation de Zeckendorf de $m+1$ où m est l'entier de représentation de Zeckendorf
$$(1,0,1,0,1,0,1,0,0,0,1,0,0,1).$$
c. Déterminer la représentation de Zeckendorf de m_1+m_2 où m_1 et m_2 sont les entiers de représentations de Zeckendorf
$$\begin{array}{c}(1,0,1,0,1,0,1,0,0,0,1,0,0,1)\\(1,0,0,1,0,0,1,0,0,1,0,0,1)\end{array}.$$

●●● Problème 3 Transformée de Fourier discrète

Dans tout ce problème, n désigne un entier naturel non nul et $\omega = e^{\frac{2i\pi}{n}}$.

L'objectif est l'étude de l'application $\mathscr{F}_n : \mathbb{C}^n \to \mathbb{C}^n$ qui associe à $u = (u_0, \dots, u_{n-1})$ l'élément $\mathscr{F}_n u = (v_0, \dots, v_{n-1})$ défini par :

$$\forall k \in [\![0, n-1]\!], \qquad v_k = \sum_{j=0}^{n-1} \overline{u}_j \omega^{jk}.$$

Partie A – Préliminaires

1. Calculer, pour tout $k \in \mathbb{N}$, les sommes $\sum_{j=0}^{n-1} \omega^{jk}$ et $\sum_{j=1}^{n} \omega^{jk}$.
2. Déterminer $\mathscr{F}_n(1, 1, \dots, 1)$ et $\mathscr{F}_n(1, \omega, \dots, \omega^{n-1})$.
3. Calculer $\mathscr{F}_n \circ \mathscr{F}_n$. En déduire que $\mathscr{F}_n$ est bijective et préciser sa bijection réciproque.
4. Déterminer $\mathscr{F}_n^{-1}(1, 1, 0, \dots, 0)$.

Partie B – Équations de convolution

Pour tous $u, v \in \mathbb{C}^n$ tels que $u = (u_0, \dots, u_{n-1})$ et $v = (v_0, \dots, v_{n-1})$, on définit :

- le produit terme à terme de u et v par $u \times v = (u_0 v_0, \dots, u_{n-1} v_{n-1})$,
- le produit de convolution de u et v par $u \otimes v = (w_0, \dots, w_{n-1})$ avec w_k la somme de tous les termes $u_i v_j$ tels que $i + j = k[n]$ que l'on note

$$\forall k \in [\![0, n-1]\!], \qquad w_k = \sum_{i+j=k[n]} u_i v_j.$$

1. Montrer que pour tous $u, v \in \mathbb{C}^n$, $\mathscr{F}_n(u \otimes v) = \mathscr{F}_n(u) \times \mathscr{F}_n(v)$.
2. En déduire une condition nécessaire et suffisante pour que l'équation $u \otimes x = v$ d'inconnue $x \in \mathbb{C}^n$ admette une solution.
3. Résoudre l'équation $(1, 1, \dots, 1) \otimes x = (1, \omega, \dots, \omega^{n-1})$.

Partie C – Algorithme de Cooley-Tukey

1. Soit $(u_0, \dots, u_3)$ et $(v_0, \dots, v_3) = \mathscr{F}_4(u_0, \dots, u_3)$.
   a. Montrer qu'il existe $(a_0, a_1, b_0, b_1) \in \mathbb{C}^4$ tel que

   $$v_0 = a_0 + b_0 \qquad v_2 = a_0 - b_0 \qquad v_1 = a_1 + \omega b_1 \qquad v_3 = a_1 - \omega b_1$$

   b. Trouver $(\alpha_0, \alpha_1) \in \mathbb{C}^2$ tel que $(a_0, a_1) = \mathscr{F}_2(\alpha_0, \alpha_1)$.
2. Supposons dans cette question que $n = 2m$ avec $m \in \mathbb{N}^*$. Pour tout $u \in \mathbb{C}^n$, on posera $u^{(P)}$ (respectivement $u^{(I)}$) l'élément de $\mathbb{C}^m$ des coefficients d'indice pair (respectivement impair) de u.
   a. Notons $\mathscr{F}_n u = (v_0, \dots, v_{n-1})$. Montrer que $(a_0, \dots, a_{m-1}) = \mathscr{F}_m(u^{(P)})$ et $(b_0, \dots, b_{m-1}) = \mathscr{F}_m(u^{(I)})$ vérifient

   $$\forall k \in [\![0, m-1]\!], \qquad v_k = a_k + \omega^k b_k \quad v_{m+k} = a_k - \omega^k b_k.$$

   b. Notons M_n le nombre de multiplications nécessaire pour calculer $\mathscr{F}_n u$ pour $u \in \mathbb{C}^n$. Que pouvez-vous dire de M_n ?

PROBLÈMES

CORRIGÉS

Corrigés des Vrai/Faux

a) Faux, il y a $2 \times 1789 + 1$ termes dans la somme.

b) Faux, on cherche à dénombrer les $k = 2l$ entre 0 et n donc les l entre 0 et $\lfloor \frac{n}{2} \rfloor$: il y en a $\lfloor \frac{n}{2} \rfloor + 1$.

c) Faux, le produit recherché est

$$\prod_{l=1}^{n} (2l) = 2^n n!.$$

d) Vrai, c'est la formule pour la suite géométrique de raison 2.

e) Vrai.

f) Faux, dès que k est un multiple de n.

g) Vrai, en effectuant un changement d'indices avec la somme des deux indices.

h) Vrai.

Corrigés des exercices d'application du chapitre 3

Exercice 1

Revenons aux sommes de puissances déjà connues :

$$\sum_{k=0}^{n} k(3k+1) = 3\sum_{k=0}^{n} k^2 + \sum_{k=0}^{n} k = 3\frac{n(n+1)(2n+1)}{6} + \frac{n(n+1)}{2} = n(n+1)^2.$$

Exercice 2

On se ramène à des calculs de sommes géométriques de raisons 2^{-5} et $\frac{2}{3}$ respectivement en mettant en facteur le terme idoine.

$$\sum_{k=0}^{n} \frac{3}{2^{5k+1}} = \frac{3}{2}\sum_{k=0}^{n} (2^{-5})^k = \frac{3}{2}.\frac{1-2^{-5(n+1)}}{1-2^{-5}},$$

$$\sum_{k=0}^{n} \frac{2^{k-1}}{3^{k+1}} = \frac{2^{-1}}{3}\sum_{k=0}^{n} \left(\frac{2}{3}\right)^k = \frac{1}{6}.\frac{1-\left(\frac{2}{3}\right)^{n+1}}{1-\frac{2}{3}} = \frac{1}{2}\left(1-\left(\frac{2}{3}\right)^{n+1}\right).$$

Exercice 3

Pour faire apparaître un télescopage, on remarque $k=(k+1)-1$; la somme s'écrit alors

$$\sum_{k=0}^{n}\frac{k}{(k+1)!}=\sum_{k=0}^{n}\left(\frac{1}{k!}-\frac{1}{(k+1)!}\right)=1-\frac{1}{(n+1)!}.$$

Exercice 4

En utilisant la « quantité conjuguée », on obtient

$$\begin{aligned}\sum_{k=1}^{n}\frac{1}{\sqrt{k+1}+\sqrt{k}}&=\sum_{k=1}^{n}\frac{\sqrt{k+1}-\sqrt{k}}{(\sqrt{k+1}+\sqrt{k})(\sqrt{k+1}-\sqrt{k})}=\sum_{k=1}^{n}\frac{\sqrt{k+1}-\sqrt{k}}{k+1-k}\\&=\sum_{k=1}^{n}\left(\sqrt{k+1}-\sqrt{k}\right)\\&=\sqrt{n+1}-1,\end{aligned}$$

par télescopage.

Exercice 5

Séparons, par associativité de la somme, les termes correspondant aux indices pairs et aux indices impairs.

$$\sum_{k=0}^{n}(-1)^k k=\sum_{\substack{k=0\\k\text{ pair}}}^{n}(-1)^k k+\sum_{\substack{k=0\\k\text{ impair}}}^{n}(-1)^k k=\sum_{l=0}^{\lfloor\frac{n}{2}\rfloor}2l-\sum_{l=0}^{\lfloor\frac{n-1}{2}\rfloor}(2l+1).$$

Pour simplifier l'écriture, séparons les cas n pair et n impair.

- Si $n=2p$ avec $p\in\mathbb{N}\setminus\{0\}$, alors

$$\begin{aligned}\sum_{k=0}^{n}(-1)^k k&=\sum_{l=0}^{p}2l-\sum_{l=0}^{p-1}(2l+1)\\&=2\frac{p(p+1)}{2}-2\frac{(p-1)p}{2}-p=p.\end{aligned}$$

- Si $n=2p+1$ avec $p\in\mathbb{N}$, alors

$$\begin{aligned}\sum_{k=0}^{n}(-1)^k k&=\sum_{l=0}^{p}2l-\sum_{l=0}^{p}(2l+1)\\&=2\frac{p(p+1)}{2}-2\frac{p(p+1)}{2}-(p+1)=-p-1.\end{aligned}$$

Exercice 6

Posons, pour compléter la notation de l'énoncé,

$$S_1=\sum_{\substack{k=0\\k=1[3]}}^{n}\binom{n}{k},\qquad S_2=\sum_{\substack{k=0\\k=2[3]}}^{n}\binom{n}{k}.$$

CORRIGÉS

Développons $(1+j)^n$ avec $j = e^{\frac{2i\pi}{3}}$ par la formule du binôme

$$\begin{aligned}(1+j)^n &= \sum_{k=0}^{n}\binom{n}{k}j^k \\ &= \sum_{\substack{k=0\\k=0[3]}}^{n}\binom{n}{k}j^k + \sum_{\substack{k=0\\k=1[3]}}^{n}\binom{n}{k}j^k + \sum_{\substack{k=0\\k=2[3]}}^{n}\binom{n}{k}j^k \\ &= S_0 + jS_1 + j^2S_2,\end{aligned}$$

en utilisant que

$$j^k = \begin{cases} 1 & \text{si } k=0[3], \\ j & \text{si } k=1[3], \\ j^2 & \text{si } k=2[3]. \end{cases}$$

On obtient de même les égalités suivantes

$$\begin{aligned}(1+j^2)^n &= S_0 + j^2S_1 + jS_2 \\ (1+1)^n &= S_0 + S_1 + S_2.\end{aligned}$$

En additionnant ces trois formules et en utilisant que $1+j+j^2=0$, on obtient que

$$3S_0 = 2^n + (1+j)^n + (1+j^2)^n.$$

On montre, de même, que

$$\begin{aligned}3S_1 &= 2^n + j^2(1+j)^n + j(1+j^2)^n; \\ 3S_2 &= 2^n + j(1+j)^n + j^2(1+j^2)^n.\end{aligned}$$

On peut calculer les sommes des coefficients binomiaux avec des indices égaux à une valeur modulo r en partant des sommes $(1+\omega)^n$ où $\omega \in \mathbb{U}_r$.

Exercice 7

Pour faire apparaître une factorielle au numérateur, on « ajoute » les termes pairs.

$$\prod_{k=1}^{n}\frac{2k-1}{2k} = \prod_{k=1}^{n}\frac{(2k-1)(2k)}{(2k)^2} = \frac{\prod_{k=1}^{n}(2k-1)(2k)}{\prod_{k=1}^{n}(2k)^2} = \frac{\prod_{k=1}^{2n}k}{\left(\prod_{k=1}^{n}2k\right)^2} = \frac{(2n)!}{(2^n n!)^2}.$$

Exercice 8

Écrivons la somme recherchée comme partie réelle d'une somme binomiale avec des exponentielles complexes.

$$\begin{aligned}\sum_{k=0}^{n}\binom{n}{k}\cos(k\theta) &= \mathfrak{Re}\left(\sum_{k=0}^{n}\binom{n}{k}e^{ik\theta}\right) = \mathfrak{Re}((1+e^{i\theta})^n) = \mathfrak{Re}\left(e^{i\frac{n\theta}{2}}\left(e^{-i\frac{\theta}{2}}+e^{i\frac{\theta}{2}}\right)^n\right) \\ &= \cos(\tfrac{n\theta}{2})(2\cos\tfrac{\theta}{2})^n.\end{aligned}$$

Exercice 9

Commençons par écrire $k^2 = k(k-1)+k$ pour tout k. L'idée est d'exploiter la formule $k\binom{n}{k} = n\binom{n-1}{k-1}$ et sa conséquence immédiate :

$$k(k-1)\binom{n}{k} = n(k-1)\binom{n-1}{k-1} = n(n-1)\binom{n-2}{k-2}.$$

Alors,

$$\begin{aligned}
\sum_{k=0}^{n} k^2\binom{n}{k} &= \sum_{k=0}^{n} k(k-1)\binom{n}{k} + \sum_{k=0}^{n} k\binom{n}{k} \\
&= \sum_{k=2}^{n} k(k-1)\binom{n}{k} + \sum_{k=1}^{n} k\binom{n}{k} \\
&= \sum_{k=2}^{n} n(n-1)\binom{n-2}{k-2} + \sum_{k=1}^{n} n\binom{n-1}{k-1} \\
&= n(n-1)\sum_{k=0}^{n-2}\binom{n-2}{k} + n\sum_{k=0}^{n-1}\binom{n-1}{k} \\
&= n(n-1)2^{n-2} + n2^{n-1}.
\end{aligned}$$

Exercice 10

Proposons deux solutions distinctes.

▹ Faisons apparaître un coefficient binomial $\binom{n+1}{k+1}$

$$\begin{aligned}
\sum_{k=0}^{n} \frac{1}{(n-k)!(k+1)!} &= \frac{1}{(n+1)!}\sum_{k=0}^{n} \frac{(n+1)!}{(n-k)!(k+1)!} \\
&= \frac{1}{(n+1)!}\sum_{k=0}^{n} \frac{(n+1)!}{(n+1-(k+1))!(k+1)!} \\
&= \frac{1}{(n+1)!}\sum_{k=0}^{n}\binom{n+1}{k+1} \\
&= \frac{1}{(n+1)!}\Big(\sum_{k=0}^{n+1}\binom{n+1}{k} - 1\Big) \qquad \text{car } \binom{n+1}{0} = 1 \\
&= \frac{1}{(n+1)!}\big(2^{n+1}-1\big).
\end{aligned}$$

CORRIGÉS

▹ Faisons apparaître un coefficient binomial $\binom{n}{k}$

$$\begin{aligned}
\sum_{k=0}^{n} \frac{1}{(n-k)!(k+1)!} &= \frac{1}{n!}\sum_{k=0}^{n} \frac{n!}{(n-k)!(k+1)!} \\
&= \frac{1}{n!}\sum_{k=0}^{n} \frac{1}{k+1}\binom{n}{k} \\
&= \frac{1}{n!}\int_0^1 \sum_{k=0}^{n}\binom{n}{k} t^k \,\mathrm{d}t \qquad \text{car } \frac{1}{k+1} = \int_0^1 t^k \,\mathrm{d}t \\
&= \frac{1}{n!}\int_0^1 (1+t)^n \,\mathrm{d}t \\
&= \frac{1}{(n+1)!}\big(2^{n+1}-1\big).
\end{aligned}$$

Exercice 11

Multiplions P_n par $\sin\frac{\theta}{2^n}$ afin d'utiliser la formule de duplication de la fonction cos.

$$\begin{aligned}
P_n.\sin\frac{\theta}{2^n} &= \Big(\prod_{k=1}^{n}\cos\frac{\theta}{2^k}\Big)\sin\frac{\theta}{2^n} \\
&= \Big(\prod_{k=1}^{n-1}\cos\frac{\theta}{2^k}\Big)\Big(\cos\frac{\theta}{2^n}\sin\frac{\theta}{2^n}\Big) \\
&= \frac{1}{2}\Big(\prod_{k=1}^{n-1}\cos\frac{\theta}{2^k}\Big)\sin\frac{\theta}{2^{n-1}} \\
&= \frac{1}{2^n}\sin\theta,
\end{aligned}$$

après une récurrence élémentaire. Ainsi, $P_n = \frac{\sin\theta}{2^n\sin\frac{\theta}{2^n}}$.

Exercice 12

Commençons par échanger les deux sommes : comme les « dépendances » s'écrivent $1 \leq l < k \leq n$, on obtient

$$\begin{aligned}
\sum_{l=1}^{n}\sum_{k=l+1}^{n}\frac{l}{k} &= \sum_{k=1}^{n}\sum_{l=1}^{k-1}\frac{l}{k} \\
&= \sum_{k=1}^{n}\frac{1}{k}\sum_{l=1}^{k-1} l \\
&= \sum_{k=1}^{n}\frac{1}{k}\frac{(k-1)k}{2} \\
&= \frac{1}{2}\sum_{j=0}^{n-1} j = \frac{(n-1)n}{4}.
\end{aligned}$$

Exercice 13

Séparons la seconde somme selon que $\min(k,l)$ est égal à k ou à l.

$$\begin{aligned}
\sum_{k=0}^{n}\sum_{l=0}^{n}\min(k,l) &= \sum_{k=0}^{n}\Big(\sum_{l=0}^{k}\min(k,l)+\sum_{l=k+1}^{n}\min(k,l)\Big)\\
&= \sum_{k=0}^{n}\Big(\sum_{l=0}^{k}l+\sum_{l=k+1}^{n}k\Big)\\
&= \sum_{k=0}^{n}\Big(\frac{k(k+1)}{2}+k(n-k)\Big)\\
&= \sum_{k=0}^{n}\Big(-\frac{1}{2}k^2+\frac{2n+1}{2}k\Big)\\
&= -\frac{1}{2}\frac{n(n+1)(2n+1)}{6}+\frac{2n+1}{2}\frac{n(n+1)}{2}\\
&= \frac{1}{6}n(n+1)(2n+1).
\end{aligned}$$

Exercice 14

Développons le terme central avec la formule du binôme puis utilisons le théorème de Fubini

$$\sum_{k=1}^{n}(z+\omega^k)^n=\sum_{k=1}^{n}\sum_{j=0}^{n}\binom{n}{j}\omega^{jk}z^{n-j}=\sum_{j=0}^{n}\binom{n}{j}z^{n-j}\sum_{k=1}^{n}\omega^{jk}.$$

Or, $\sum_{k=1}^{n}\omega^{jk}=\begin{cases} n & \text{si } j=0[n],\\ 0 & \text{sinon.}\end{cases}$

Par conséquent, $\sum_{k=1}^{n}(z+\omega^k)^n=n(1+z^n)$.

Corrigés des exercices d'approfondissement du chapitre 3

Exercice A

1. Réécrivons cette quantité en utilisant la symétrie des coefficients binomiaux

$$\begin{aligned}
S_p+S_{n-p-1} &= \sum_{k=0}^{p}\binom{n}{k}+\sum_{k=0}^{n-p-1}\binom{n}{k}\\
&= \sum_{k=0}^{p}\binom{n}{k}+\sum_{l=p+1}^{n}\binom{n}{n-l} \qquad \text{avec } l=n-k\\
&= \sum_{k=0}^{p}\binom{n}{k}+\sum_{l=p+1}^{n}\binom{n}{l}\\
&= \sum_{k=0}^{n}\binom{n}{k}=2^n.
\end{aligned}$$

2. Calculons le double de cette somme de sorte à pouvoir se ramener à la question précédente.

$$\begin{aligned}
2\sum_{p=0}^{n-1} S_p &= \sum_{p=0}^{n-1} S_p + \sum_{p=0}^{n-1} S_p \\
&= \sum_{p=0}^{n-1} S_p + \sum_{q=0}^{n-1} S_{n-1-q} \qquad \text{avec } q = n-1-p \\
&= \sum_{p=0}^{n-1} \left(S_p + S_{n-1-p}\right) \\
&= \sum_{p=0}^{n-1} 2^n = n2^n.
\end{aligned}$$

Ainsi, la somme recherchée est $\sum\limits_{p=0}^{n-1} S_p = n2^{n-1}$.

Exercice B

Séparons dans la somme les termes d'indices pairs et ceux d'indices impairs :

$$\begin{aligned}
\sum_{k=1}^{4n} u_k\, u_{k+2} &= \sum_{\substack{k=1\\ k=0[2]}}^{4n} u_k\, u_{k+2} + \sum_{\substack{k=1\\ k=1[2]}}^{4n} u_k\, u_{k+2} \\
&= \sum_{l=1}^{2n} u_{2l}\, u_{2(l+1)} + \sum_{l=0}^{2n-1} u_{2l+1}\, u_{2(l+1)+1} \\
&= \sum_{l=1}^{2n} u_{2l}\, u_{2(l+1)} + \sum_{l=1}^{2n-1} u_{2l+1}\, u_{2(l+1)+1} + u_1\, u_3 \\
&= \sum_{l=1}^{2n} u_{2l}\, u_{2(l+1)} + \sum_{l=1}^{2n-1} u_{2l+1}\, u_{2(l+1)+1} - 1.
\end{aligned}$$

Exploitons alors les relations de récurrence : pour tout $l \leq 2n$,

$$u_{2l}\, u_{2(l+1)} = u_l\, u_{l+1}, \qquad u_{2l+1}\, u_{2l+3} = (-1)^l\, u_l (-1)^{l+1}\, u_{l+1} = -u_l\, u_{l+1}.$$

Ainsi,

$$\begin{aligned}
\sum_{k=1}^{4n} u_k\, u_{k+2} &= \sum_{l=1}^{2n} u_l\, u_{l+1} - \sum_{l=1}^{2n-1} u_l\, u_{l+1} - 1 \\
&= u_{2n}\, u_{2n+1} - 1 \\
&= (-1)^n\, u_n^2 - 1 = (-1)^n - 1.
\end{aligned}$$

Exercice C

Remarquons tout d'abord que, pour tout $k \leq n$, $2x_k x_{k+1} \leq x_k^2 + x_{k+1}^2$. Ainsi,

$$\begin{aligned}
2\sum_{k=1}^{n} x_k x_{k+1} &\leq \sum_{k=1}^{n}\left(x_k^2 + x_{k+1}^2\right) \\
&= \sum_{k=1}^{n} x_k^2 + \sum_{k=2}^{n+1} x_k^2 \\
&= \sum_{k=1}^{n} x_k^2 + \sum_{k=1}^{n} x_k^2 = 2\sum_{k=1}^{n} x_k^2.
\end{aligned}$$

On obtient le résultat attendu en divisant par 2.

Exercice D

Remarquons que, pour tous i et j, $x_i - x_j$ et $y_i - y_j$ sont de même signe. Par conséquent,

$$\begin{aligned}
0 &\leq \sum_{i=1}^{n}\sum_{j=1}^{n}(x_i - x_j)(y_i - y_j) \\
&\leq \sum_{i=1}^{n}\sum_{j=1}^{n}(x_i y_i + x_j y_j - x_i y_j - x_j y_i) \\
&\leq 2n\sum_{i=1}^{n} x_i y_i - 2\sum_{i=1}^{n}\sum_{j=1}^{n} x_i y_j.
\end{aligned}$$

En conclusion,

$$2n\sum_{i=1}^{n} x_i y_i \geq 2\sum_{i=1}^{n} x_i \sum_{j=1}^{n} y_j,$$

ce qui est bien l'inégalité annoncée.

Exercice E

1. Pour tout $k \in \mathbb{N}$,

$$\begin{aligned}
A_k = \sum_{z\in\mathbb{U}_n} z^k &= \sum_{m=0}^{n-1} e^{\frac{2ikm\pi}{n}} = \sum_{m=0}^{n-1}\left(e^{\frac{2ik\pi}{n}}\right)^m \\
&= \begin{cases} n & \text{si } k = 0[n] \\ 0 & \text{sinon.} \end{cases}
\end{aligned}$$

2. Notons $P(X) = \sum_{k=0}^{n-1} a_k X^k$. Pour comprendre l'idée de la preuve, commençons par calculer

$$\begin{aligned}
\sum_{z\in\mathbb{U}_n} P(z) &= \sum_{z\in\mathbb{U}_n}\sum_{k=0}^{n-1} a_k z^k \\
&= \sum_{k=0}^{n-1} a_k \sum_{z\in\mathbb{U}_n} z^k \\
&= \sum_{k=0}^{n-1} a_k A_k = n a_0,
\end{aligned}$$

d'après la question précédente. Ainsi, l'inégalité triangulaire donne la majoration souhaitée

$$|a_0| = \frac{1}{n}\left|\sum_{z\in\mathbb{U}_n} P(z)\right| \leq \frac{1}{n}\sum_{z\in\mathbb{U}_n} |P(z)| \leq M.$$

Reprenons cette idée pour étudier les autres coefficients en effectuant un « décalage » des indices. Pour $j \in [\![0, n-1]\!]$,

$$\sum_{z\in\mathbb{U}_n} z^{-j} P(z) = \sum_{z\in\mathbb{U}_n}\sum_{k=0}^{n-1} a_k z^{k-j} = \sum_{k=0}^{n-1} a_k A_{k-j} = n a_j.$$

Comme précédemment, $|a_j| \leq M$.

Exercice F

Suivons l'indication et simplifions la quantité proposée :

$$\begin{aligned}(1-z)\sum_{k=0}^{n} a_k z^k &= \sum_{k=0}^{n} a_k z^k - \sum_{k=0}^{n} a_k z^{k+1}\\ &= \sum_{k=0}^{n} a_k z^k - \sum_{k=1}^{n+1} a_{k-1} z^k\\ &= a_0 + \sum_{k=1}^{n}(a_k - a_{k-1}) z^k - a_n z^{n+1}.\end{aligned}$$

Utilisons alors l'inégalité triangulaire sous sa « forme gauche » :

$$\forall u, v \in \mathbb{C}, \qquad |u-v| \geq |u| - |v|.$$

Ainsi,

$$\left|(1-z)\sum_{k=0}^{n} a_k z^k\right| \geq a_0 - \sum_{k=1}^{n}(a_{k-1} - a_k)|z|^k - a_n |z|^{n+1}.$$

Si $|z| < 1$, alors

$$\left|(1-z)\sum_{k=0}^{n} a_k z^k\right| > a_0 - \sum_{k=1}^{n}(a_{k-1} - a_k) - a_n = 0$$

et donc z ne vérifie pas $\sum_{k=0}^{n} a_k z^k = 0$. Par contraposition, si z vérifie l'équation, alors $|z| \geq 1$.

Exercice G

En utilisant l'expression du module au carré $|Z|^2 = Z\overline{Z}$,

$$|Z|^2 = \sum_{k=0}^{n-1} \omega^{k^2} \sum_{l=0}^{n-1} \omega^{-l^2} = \sum_{k=0}^{n-1}\sum_{l=0}^{n-1} \omega^{k^2 - l^2}.$$

Répartissons alors les couples d'indices (k, l) selon la valeur du reste r de $k+l$ modulo n. À $r \in [\![0, n-1]\!]$ fixé, les couples correspondants sont les $(k, r-k)$ pour $k \in [\![0, r]\!]$ et les $(k, r-k+n)$ pour $k \in [\![r+1, n-1]\!]$.

Ainsi, la somme recherchée est

$$\begin{aligned}
|Z|^2 &= \sum_{r=0}^{n-1}\left(\sum_{k=0}^{r}\omega^{k^2-(r-k)^2}+\sum_{k=r+1}^{n-1}\omega^{k^2-(r-k+n)^2}\right)\\
&= \sum_{r=0}^{n-1}\left(\sum_{k=0}^{r}\omega^{k^2-(r-k)^2}+\sum_{k=r+1}^{n-1}\omega^{k^2-(r-k)^2}\right) \text{ car } \omega^n=1\\
&= \sum_{r=0}^{n-1}\sum_{k=0}^{n-1}\omega^{k^2-(r-k)^2}\\
&= \sum_{r=0}^{n-1}\omega^{-r^2}\sum_{k=0}^{n-1}(\omega^{2r})^k.
\end{aligned}$$

Rappelons, par ailleurs, que $\sum_{k=0}^{n-1}(\omega^{2r})^k = \begin{cases} n & \text{si } 2r=0[n],\\ 0 & \text{sinon.}\end{cases}$

Discutons alors sur la parité de n pour trouver les entiers r tels que $2r=0[n]$.

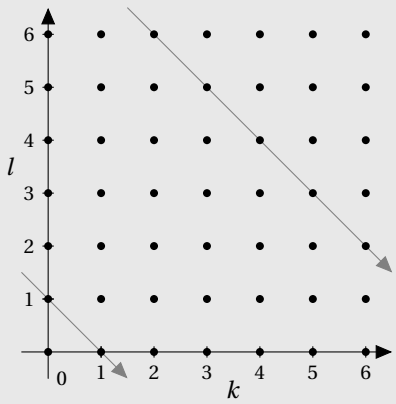
- Si $n=2p+1$ est impair, le seul entier $r\in[\![0,n-1]\!]$ qui convient est 0 donc la somme ne comporte qu'un terme : $|Z|^2=n$.
- Si $n=2p$ est pair, deux entiers $r\in[\![0,n-1]\!]$ conviennent : 0 et p. Par conséquent,

$$|Z|^2 = n(1+\omega^{-p^2}) = n(1+(-1)^{\frac{n}{2}}),$$

car $\omega^p=\omega^{\frac{n}{2}}=-1$.

Le point crucial de cet exercice est le changement d'indices pour passer de (k,l) à (k,r). Visuellement, ce changement correspond à la partition du « carré » d'indices en « diagonales continuées ».

Voici l'illustration pour $n=7$ de l'ensemble des couples (k,l) tels que $k+l=1[n]$



Corrigés des problèmes du chapitre 3

Problème 1

1. Le seul couple (z_0,z_1) satisfaisant $\mathscr{P}(2)$ est $(1,-1)$ car $0=z_0+z_1=1+z_1$.

CORRIGÉS

2. a. L'identité remarquable pour le carré d'une somme donne

$$z_0z_1+z_1z_2+z_2z_0=\frac{1}{2}\left((z_0+z_1+z_2)^2-z_0^2+z_1^2+z_2^2\right).$$

b. Si le triplet (z_0,z_1,z_2) satisfait $\mathscr{P}(3)$, alors $z_0+z_1+z_2=0$ et $z_0z_1+z_1z_2+z_2z_0=0$ donc z_0,z_1,z_2 sont les racines d'un polynôme de la forme $X^3-\alpha$ en utilisant les relations entre coefficients et racines.
Comme $z_0=1$ est racine de ce polynôme on en déduit que z_0, z_1, z_2 sont les racines du polynôme X^3-1.
c. Les triplets satisfaisant $\mathscr{P}(3)$ sont donc $(1,j,j^2)$ et $(1,j^2,j)$.

3. Soit $(z_0,\dots,z_3)$ satisfaisant $\mathscr{P}(4)$. Alors, pour tout $k\in[\![1,3]\!]$, $\sum_{j=1}^{3}z_j^k=-z_0^k=-1$. On trouve comme précédemment $z_3z_1+z_1z_2+z_2z_3=1$ puis $z_1z_2z_3=-1$ car

$$3z_1z_2z_3=z_1^3+z_2^3+z_3^3+3(z_1+z_2+z_3)(z_3z_1+z_1z_2+z_2z_3)-(z_1+z_2+z_3)^3.$$

Ainsi, les racines du polynôme X^3+X^2+X+1 (à savoir $i,-1$ et $-i$) sont exactement z_1, z_2, z_3. Les 4-uplets satisfaisant $\mathscr{P}(4)$ sont alors $(1,i,-1,-i)$ et ceux obtenus par permutation des trois dernières composantes (soit $3!=6$ 4-uplets).

4. a. Il suffit de passer au conjugué les relations définissant $\mathscr{P}_n$.
b. Pour tout $k\in[\![1,2n-1]\!]$, en utilisant que $z_j'^2=z_j$,

$$\begin{aligned}\sum_{j=0}^{2n-1}z_j'^k&=2\sum_{j=0}^{n-1}z_j^{\frac{k}{2}}=0 && \text{si } k=0[2],\\ \sum_{j=0}^{2n-1}z_j'^k&=\sum_{j=0}^{2n-1}z_j^{\frac{k-1}{2}}z_j' \\ &=\sum_{j=0}^{n-1}z_j^{\frac{k-1}{2}}(z_j'+z_{n-1+j}')=0 && \text{sinon.}\end{aligned}$$

Pour conclure dans le second cas, on a utilisé que $z_j'=-z_{n-1+j}'$ pour tout j. D'où le résultat annoncé.

Problème 2

1. Montrons le résultat par récurrence sur $n\in\mathbb{N}$.
▹ $u_0=u_2-2$ donc la propriété est vraie au rang 0.
▹ Soit $n\in\mathbb{N}$ tel que la propriété soit vraie au rang n. Alors, par hypothèse de récurrence,

$$\sum_{k=0}^{n+1}u_k=u_{n+1}+u_{n+2}-2=u_{n+3}-2.$$

2. a. Pour tout $n\in\mathbb{N}$,

$$\sigma_{n+2}=\sum_{k=0}^{[\frac{n}{2}]+1}u_{n+2-2k}=u_{n+2}+\sum_{k=1}^{[\frac{n}{2}]+1}u_{n-2(k-1)}=u_{n+2}+\sigma_n.$$

b. Montrons le résultat par récurrence sur $n\in\mathbb{N}$.
▹ $\sigma_0=u_0=u_1-1$ et $\sigma_1=u_1=u_2-1$.
▹ Soit $n\geq 1$ tel que la propriété soit vraie au rang n. Alors,

$$\sigma_{n+1}=\sigma_{n-1}+u_n=u_n+u_{n+1}-1=u_{n+2}-1.$$

3. a. Montrons le résultat par récurrence sur N.
▹ Un entier m qui admet une représentation de Zeckendorf avec $N=0$ est nécessairement égal à u_0 donc vaut 1 et $m=1=\sigma_0$.

▹ Soit $N \in \mathbb{N}$. Supposons que tout entier m qui admet une représentation de Zeckendorf d'indice maximal $k \leq N$ vérifie $m \leq \sigma_k$ et considérons un entier m qui admet une représentation de Zeckendorf d'indice maximal $N+1$. Alors, $m - u_{N+1}$ est soit nul (auquel cas, $m = u_{N+1} \leq \sigma_{N+1}$), soit admet une représentation de Zeckendorf d'indice maximal au plus $N-1$. Dans ce dernier cas, par hypothèse de récurrence, $m - u_{N+1} \leq \sigma_{N-1}$ et donc

$$m \leq u_{N+1} + \sigma_{N-1} = \sigma_{N+1}.$$

Par récurrence, le résultat est établi.

b. Par construction, $u_N \leq m$. D'après la question précédente, $m \leq \sigma_N = u_{n+1} - 1 < u_{N+1}$. Ainsi, u_N est le plus grand terme de la suite $(u_n)_{n\in\mathbb{N}}$ inférieur ou égal à m.

4. • Montrons par récurrence sur $N \geq 1$ que tous les entiers $m \in [\![1, u_N - 1]\!]$ admettent une représentation de Zeckendorf d'indice maximal $N-1$.

• Pour $N = 1$, le seul entier à considérer est 1 et il admet bien la représentation de Zeckendorf (1).

• Soit $N \geq 1$ tel que tout entier de $[\![1, u_N - 1]\!]$ admettte une représentation de Zeckendorf. Considérons $m \in [\![u_N, u_{N+1} - 1]\!]$. Alors, $m = u_N$ (auquel cas l'existence est établie) ou $m - u_N \in [\![1, u_{N-1} - 1]\!]$ et, d'après l'hypothèse de récurrence, $m - u_N$ admet une représentation de Zeckendorf d'indice maximal $N-2$: m admet alors une représentation de Zeckedorf d'indice maximal N.

• Passons à l'unicité. Si un entier m admet deux représentations de Zeckendorf, alors il existe, quitte à enlever les termes, deux représentations d'indices maximaux différents, ce qui contredit le résultat établi à la question 3.

5. a. Le plus grand terme de la suite inférieur ou égal à 2016 est 1597 : $2016 = 1597 + 419$.
Le plus grand terme de la suite inférieur ou égal à 419 est 377 : $2016 = 1597 + 377 + 42$.
Le plus grand terme de la suite inférieur ou égal à 42 est 34 : $2016 = 1597 + 377 + 34 + 8$.
Ainsi, $2016 = u_4 + u_7 + u_{12} + u_{15}$ et la représentation de Zeckendorf de 2016 est

$$(0,0,0,0,1,0,0,1,0,0,0,0,1,0,0,1).$$

b. Effectuons les calculs pas à pas en remplaçant tant que l'on peut les séquences $(1,1,0)$ et $(1,1,1)$ par $(0,0,1)$ et $(0,0,2)$.

$$\begin{aligned}
m+1 &= (2,0,1,0,1,0,1,0,0,0,1,0,0,1)\\
&= (0,1,1,0,1,0,1,0,0,0,1,0,0,1)\\
&= (0,0,0,1,1,0,1,0,0,0,1,0,0,1)\\
&= (0,0,0,0,0,1,1,0,0,0,1,0,0,1)\\
&= (0,0,0,0,0,0,0,1,0,0,1,0,0,1).
\end{aligned}$$

c. De même,

$$\begin{aligned}
m_1 + m_2 &= (2,0,1,1,1,0,2,0,0,1,1,0,1,1)\\
&= (0,1,1,1,1,0,2,0,0,1,1,0,1,1)\\
&= (0,1,1,1,1,0,2,0,0,1,1,0,0,0,1)\\
&= (0,1,1,1,1,0,2,0,0,0,0,1,0,0,1)\\
&= (0,1,1,0,0,1,2,0,0,0,0,1,0,0,1)\\
&= (0,1,1,0,0,0,1,1,0,0,0,1,0,0,1)\\
&= (0,1,1,0,0,0,0,0,1,0,0,1,0,0,1)\\
&= (0,0,0,1,0,0,0,0,1,0,0,1,0,0,1).
\end{aligned}$$

CORRIGÉS

Problème 3

Partie A – Préliminaires

1. Comme déjà calculé dans les exercices, $\sum_{j=0}^{n-1} \omega^{jk} = \sum_{j=1}^{n} \omega^{jk} = \begin{cases} n & \text{si } k = 0[n] \\ 0 & \text{sinon.} \end{cases}$
2. D'après la question précédente, $\mathscr{F}_n(1,1,\dots,1) = (n,0,\dots,0)$ et $\mathscr{F}_n(1,\omega,\dots,\omega^{n-1}) = (0,n,0,\dots,0)$.
3. Notons $w = \mathscr{F}_n \circ \mathscr{F}_n(u)$. Alors, pour tout $k \in [\![0, n-1]\!]$,

$$w_k = \sum_{j=0}^{n-1} \omega^{jk} \overline{\sum_{l=0}^{n-1} \overline{u_l} \omega^{jl}} = \sum_{l=0}^{n-1} u_l \sum_{j=0}^{n-1} \omega^{j(k-l)} = n\, u_k$$

en utilisant la question 1. D'où $\mathscr{F}_n \circ \mathscr{F}_n = n.Id$ et $\mathscr{F}_n$ est bijective de réciproque $\frac{1}{n}\mathscr{F}_n$.
4. Par linéarité $\mathscr{F}_n^{-1}(1,1,0,\dots,0) = \mathscr{F}_n^{-1}(1,0,\dots,0) + \mathscr{F}_n^{-1}(0,1,0,\dots,0)$, le résultat est donc $\frac{1}{n}(1,1,\dots,1) + \frac{1}{n}(1,\omega,\dots,\omega^{n-1})$.

Partie B – Équations de convolution

1. Notons $w = u \otimes v$ et $z = \mathscr{F}_n(w)$. Alors, pour tout $k \in [\![0, n-1]\!]$,

$$\begin{aligned} z_k = \sum_{l=0}^{n-1} w_l \omega^{lk} &= \sum_{l=0}^{n-1} \sum_{i+j=l[n]} \overline{u_i v_j} \omega^{lk} \\ &= \sum_{l=0}^{n-1} \sum_{i+j=l[n]} \overline{u_i} \omega^{ik} \overline{v_j} \omega^{jk} \\ &= \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \overline{u_i} \omega^{ik} \overline{v_j} \omega^{jk} \end{aligned}$$

D'où le résultat.
2. $u \otimes x = v$ si, et seulement si, $\mathscr{F}_n(u)\mathscr{F}_n(x) = \mathscr{F}_n(v)$ (car $\mathscr{F}_n$ est bijective). Une condition nécessaire et suffisante d'existence des solutions est que pour tout coefficient nul de $\mathscr{F}_n(u)$ correspond un coefficient nul de $\mathscr{F}_n(v)$.
3. D'après la question précédente, l'équation n'admet pas de solution.

Partie C – Algorithme de Cooley-Tukey

1. a. Avec les expressions des v_i, on trouve $a_0 = \overline{u_0} + \overline{u_2}$, $a_1 = \overline{u_0} - \overline{u_2}$, $b_0 = \overline{u_1} + \overline{u_3}$ et $b_1 = \overline{u_1} - \overline{u_3}$.
b. On en déduit que $(\alpha_0, \alpha_1) = (u_0, u_2)$ (et que $(b_0, b_1) = \mathscr{F}_2(u_1, u_3)$.
2. a. On remarque que pour tout $k \in [\![0, n-1]\!]$

$$v_k = \sum_{j=0}^{2m-1} \overline{u_j} \omega^{jk} = \sum_{j=0}^{m-1} \overline{u_{2j}} \omega^{2jk} + \omega^k \sum_{j=0}^{m-1} \overline{u_{2j-1}} \omega^{2jk}.$$

Cela suffit pour conclure en remarquant que $\omega^2 = e^{\frac{2i\pi}{m}}$.
b. La méthode de calcul naïve nous donne $M_n \leq n^2$. En utilisant la question précédente, on remarque que $M_{2^p} \leq 2.M_{2^{p-1}} + 2^{p-1}$ (ce qui nous donne une borne de l'ordre de $N \log(N)$).

Chapitre 4

Fonctions usuelles

1. Propriétés des fonctions réelles — **2.** Valeur absolue — **3.** Exponentielles, logarithmes, puissances — **4.** Exponentielle complexe, fonctions circulaires

Objectifs et compétences du programme

- Reconnaître et interpréter les symétries d'une courbe représentative de fonction.
- Apprendre les propriétés des fonctions usuelles.
- Savoir utiliser les formules trigonométriques (hyperboliques et circulaires).

Leonhard Euler, mathématicien suisse (1707-1783)
Il a contribué à toutes les mathématiques de son temps (et à beaucoup de sciences physiques) ; on peut citer quelques résultats qui ont conservé son nom dans la mémoire mathématique :

- la formule d'Euler sur les nombres complexes (voir chapitre précédent),
- les chemins eulériens sur les graphes avec l'exemple des ponts de Königsberg,
- la caractéristique d'Euler des polyèdres,
- la droite et le cercle d'Euler d'un triangle,
- le produit eulérien de la fonction ζ (qui lui permet de fournir une nouvelle preuve de l'infinité des nombres premiers),
- les équations d'Euler-Lagrange qui caractérisent les solutions extrémales à un problème géométrique...

On lui doit également les notations e pour $\exp(1)$, $\sum$ pour les sommes, i pour le complexe ou même $f(x)$ pour la valeur d'une fonction f en l'argument x. En ce qui concerne ce chapitre, il fournit notamment de nouvelles expressions pour les fonctions exponentielle et arctangente dans ces travaux sur les séries et produits infinis.

COURS

Après avoir rappelé les principales propriétés remarquables des fonctions à valeurs réelles, nous explorons les différentes familles de fonctions usuelles qui apparaissent dans les calculs en analyse celles qui sont liées à la valeur absolue, à l'exponentielle réelle puis à l'exponentielle complexe. C'est également l'occasion de reprendre le formulaire de trigonométrie.

1. Propriétés des fonctions réelles

1.1. Parité et périodicité

Définition 4.1.

Soit $f : D \to \mathbb{R}$.
▹ La fonction f est paire si pour tout $x \in D$, $-x \in D$ et $f(-x) = f(x)$.
▹ La fonction f est impaire si pour tout $x \in D$, $-x \in D$ et $f(-x) = -f(x)$.
▹ Soit $T > 0$. La fonction f est T-périodique si pour tout $x \in D$, $x + T \in D$ et $f(x+T) = f(x)$.

Remarque

Une fonction f impaire définie en 0 est toujours nulle en 0 car $f(-0) = -f(0)$.

Exemple

Une fonction polynomiale est paire (respectivement impaire) si les exposants correspondant à ses coefficients non nuls sont tous pairs (respectivement impairs).
En effet, pour $f : x \mapsto \sum_{k=0}^{N} a_k x^k$, la condition $f(-x) = f(x)$ pour tout $x \in \mathbb{R}$ se réécrit

$$\forall x \in \mathbb{R}, \qquad \sum_{\substack{k=0 \\ k \text{impair}}}^{N} 2a_k x^k = 0,$$

ce qui implique le résultat annoncé.

Remarque

Il existe des fonctions qui ne sont ni paires, ni impaires comme par exemple $x \mapsto x^3 + x^2$. Pour étudier une telle fonction f, on peut la décomposer comme somme des fonctions paire et impaire suivantes

$$x \mapsto \frac{f(x) + f(-x)}{2}, \qquad x \mapsto \frac{f(x) - f(-x)}{2}.$$

Remarquons que cette décomposition est unique : en effet, si $f = g + h$ avec g paire et h impaire, alors, pour tout x,

$$f(x) = g(x) + h(x), \qquad f(-x) = g(x) - h(x),$$

et donc

$$g(x) = \frac{f(x) + f(-x)}{2}, \qquad h(x) = \frac{f(x) - f(-x)}{2}.$$

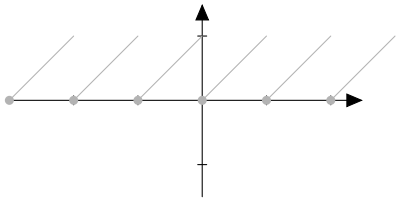
Ces deux fonctions sont respectivement appelées partie paire et partie impaire de f.

Exemple

Les fonctions cos et sin sont 2π-périodiques ; la fonction tan définie par le quotient $\frac{\sin}{\cos}$ est quant à elle π-périodique.

Exemple

Définissons la partie entière $\lfloor x \rfloor$ d'un réel x comme le plus grand entier inférieur ou égal à x puis la partie fractionnaire $\{x\}$ de x comme la différence entre x et sa partie entière. Par exemple, $\{1,789\} = 0,789$, $\{2,789\} = 0,789$ et $\{-0,789\} = 0,221$ (car $\lfloor -0.789 \rfloor = -1$).
La fonction $x \mapsto \{x\}$ est 1-périodique et sa courbe représentative est la suivante



On déduit immédiatement de la définition la proposition suivante.

Proposition 4.2.

La somme de deux fonctions paires (respectivement impaires) est paire (respectivement impaire).

En revanche la somme de deux fonctions périodiques n'est pas forcément périodique.

1.2. Courbes représentatives

Définition 4.3.

La courbe représentative d'une fonction $f : I \subset \mathbb{R} \to \mathbb{R}$ est l'ensemble des points du plan dont les coordonnées sont de la forme $(x, f(x))$ avec $x \in I$; autrement dit, la courbe représentative de f est la courbe d'équation $y = f(x)$.

Remarque

Par définition, une fonction n'associe qu'une seule image à un argument donné.
Par conséquent, la courbe représentative de $f : I \to \mathbb{R}$ ne rencontre qu'en un seul point la droite verticale d'équation $x = a$ avec $a \in I$.

Proposition 4.4.

Soit une fonction $f : I \to \mathbb{R}$ et $\mathscr{C}_f$ sa courbe représentative.
▹ La courbe représentative de $x \mapsto f(x+a)$ est déduite de $\mathscr{C}_f$ par translation de vecteur $(-a, 0)$.
▹ La courbe représentative de $x \mapsto f(x)+a$ est déduite de $\mathscr{C}_f$ par translation de vecteur $(0, a)$.
▹ La courbe représentative de $x \mapsto b - f(x)$ est déduite de $\mathscr{C}_f$ par symétrie d'axe d'équation $y = \frac{b}{2}$.
▹ La courbe représentative de $x \mapsto f(b-x)$ est déduite de $\mathscr{C}_f$ par symétrie d'axe d'équation $x = \frac{b}{2}$.

On remarque qu'en prenant $b = 0$ dans les deux derniers cas on retrouve les symétries par rapports aux axes du repère. Démontrons seulement l'un des points de cette proposition.

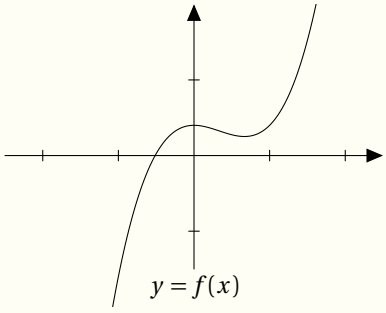
Démonstration

Soit $b \in \mathbb{R}$, $\widetilde{f} : x \mapsto f(b-x)$ et (x, y) un point du plan. Alors,

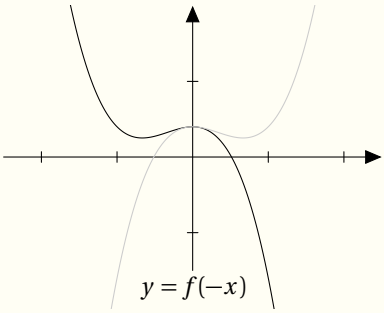
$$(x, y) \in \mathscr{C}_{\widetilde{f}} \Leftrightarrow y = f(b-x) \Leftrightarrow (b-x, y) \in \mathscr{C}_f.$$

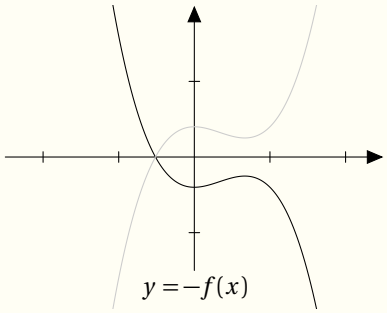
Ainsi, $(\frac{b}{2} + x, y)$ appartient à la courbe représentative de f si, et seulement si, $(\frac{b}{2} - x, y)$ appartient à la courbe représentative de $\widetilde{f}$: la courbe de $\widetilde{f}$ est déduite de celle de f par symétrie d'axe d'équation $x = \frac{b}{2}$. ■

Illustrons cette propriété avec la courbe représentative de la fonction $f : x \mapsto x^3 - x^2 + \frac{2}{5}$.

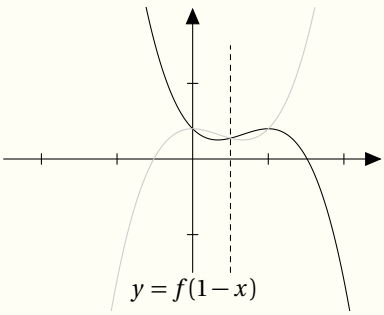


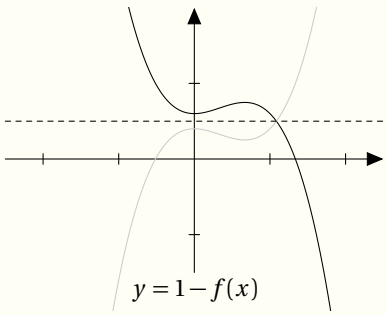
On obtient les courbes représentatives des fonctions $x \mapsto f(-x)$ et $x \mapsto -f(x)$ par symétrie par rapport à l'axe des ordonnées et des abscisses respectivement.



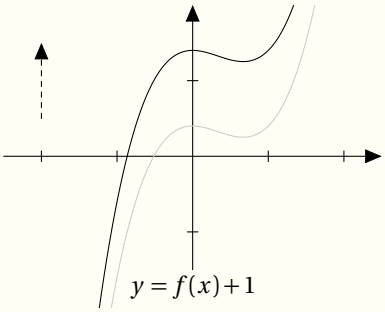


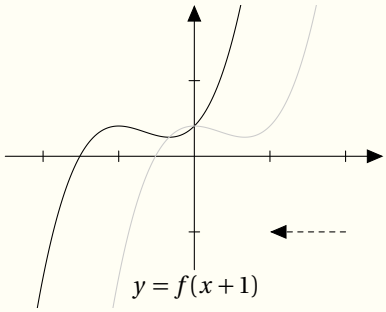
On pourrait bien sûr combiner les deux et déduire la courbe représentative de $x \mapsto -f(-x)$ de la courbe de f par symétrie centrale en l'origine. Ces résultats se généralisent aussi en « translatant » les courbes.





Enfin, on prendra garde à ne pas confondre une translation sur l'argument de f (qui donne une translation horizontale de la courbe) et une translation sur la valeur de f (qui donne une translation verticale de la courbe).





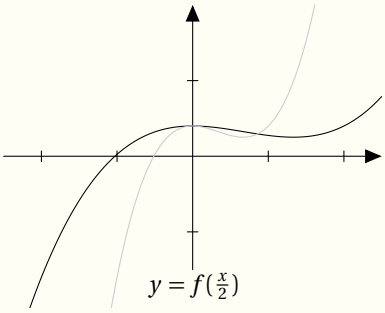
La propriété précédente permet d'obtenir un certain nombre de corollaires dont voici un exemple.

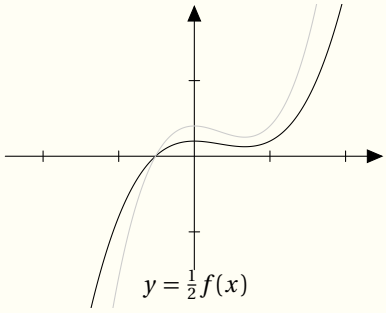
Corollaire 4.5.

La courbe représentative d'une fonction f vérifiant la relation $f(a-x) = f(x)$ est symétrique par rapport à la droite (verticale) d'équation $x = \frac{a}{2}$.

En particulier, la courbe représentative d'une fonction paire (respectivement impaire) est symétrique par rapport à l'axe des ordonnées (respectivement l'origine).

De la même façon on déduit les courbes représentatives de $x \mapsto f(ax)$ et $x \mapsto af(x)$ en dilatant selon l'axe des abscisses et selon l'axe des ordonnées respectivement.

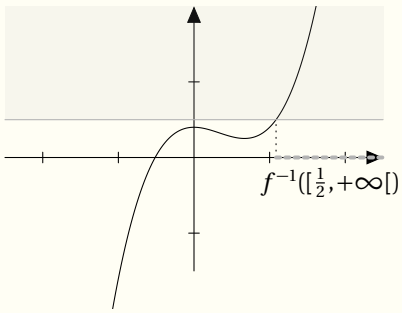




Remarque

On peut encore utiliser différemment la courbe représentative d'une fonction f.

Par exemple, la zone colorée sur le dessin ci-dessous correspond aux points d'ordonnées supérieures ou égales à $\frac{1}{2}$. En reportant sur l'axe des abscisses la partie de la courbe appartenant à cette zone, on obtient l'image réciproque $f^{-1}([\frac{1}{2}, +\infty[)$.



1.3. Fonctions dérivables

Définition 4.6.

Soit $f : I \to \mathbb{R}$.
▹ La fonction f est dérivable en $a \in I$ si la fonction $x \mapsto \frac{f(x)-f(a)}{x-a}$ admet une limite finie en a. Cette limite est alors appelée dérivée de f en a et notée $f'(a)$.
▹ La fonction f est dérivable sur I si elle est dérivable en tout point de I.
▹ La fonction f est de classe $\mathscr{C}^1$ sur I si f est dérivable sur I et si la fonction $x \mapsto f'(x)$ est continue.

Remarque

Interprétons cette définition à l'aide de la courbe représentative d'une fonction f. Pour $x \neq a$, la quantité $\frac{f(x)-f(a)}{x-a}$ est la pente de la droite joignant les points de la courbe $(x, f(x))$ et $(a, f(a))$. La « limite » de cette droite lorsque x tend vers a est, quand elle existe, la tangente en a à la courbe représentative de f ; le nombre dérivé en a est donc la pente de la tangente. Une limite infinie correspond ainsi à une tangente verticale.

On peut étendre la définition des classes de régularité
- une fonction est de classe $\mathscr{C}^0$ si elle est continue ;
- pour tout $n \geq 1$, une fonction est de classe $\mathscr{C}^n$ si elle est dérivable et que sa dérivée est de classe $\mathscr{C}^{n-1}$;
- une fonction est de classe $\mathscr{C}^\infty$ si elle est de classe $\mathscr{C}^n$ pour tout $n \in \mathbb{N}$.

On peut également étendre cette définition aux fonctions à valeurs complexes : une fonction $f : I \to \mathbb{C}$ est dérivable en $a \in I$ si $\mathfrak{Re}(f)$ et $\mathfrak{Im}(f)$ sont dérivables en a. La dérivée de f en a est alors

$$f'(a) = \mathfrak{Re}(f)'(a) + i\mathfrak{Im}(f)'(a).$$

Exemple

Étudions la dérivabilité de la fonction $x \mapsto \sqrt{x}$ sur son domaine de définition, $\mathbb{R}_+$.
▹ Remarquons tout d'abord que $\lim\limits_{x\to 0} \frac{\sqrt{x}-\sqrt{0}}{x-0} = +\infty$.
▹ Ensuite, pour tout $a > 0$,

$$\lim_{x\to a} \frac{\sqrt{x}-\sqrt{a}}{x-a} = \lim_{x\to a} \frac{1}{\sqrt{x}+\sqrt{a}} = \frac{1}{2\sqrt{a}}.$$

En conclusion, la fonction $x \mapsto \sqrt{x}$ est dérivable sur $\mathbb{R}_+^*$ mais n'est pas dérivable en 0.

Mentionnons quelques résultats opératoires sur la dérivation qui seront démontrés plus tard dans le cours.

Proposition 4.7.

Soit f, $g : I \to \mathbb{R}$ deux fonctions dérivables en $a \in I$, λ, $\mu \in \mathbb{R}$.
▹ La fonction $\lambda f + \mu g$ est dérivable en a et $(\lambda f + \mu g)'(a) = \lambda f'(a) + \mu g'(a)$.
▹ le produit $f.g$ est dérivable en a et $(f.g)'(a) = f'(a)g(a) + f(a)g'(a)$.
▹ Si $g(a) \neq 0$, alors le quotient $\frac{f}{g}$ est dérivable en a et

$$\left(\frac{f}{g}\right)'(a) = \frac{f'(a)g(a) - f(a)g'(a)}{g(a)^2}.$$

Exemple

Considérons la fonction homographique définie par $f : x \mapsto \frac{ax+b}{cx+d}$ avec a, b, c, $d \in \mathbb{R}$ tels que $ad - bc \neq 0$ (pour que f ne soit pas constante) et $c \neq 0$. Alors, f est définie et de classe $\mathscr{C}^\infty$ sur son ensemble de définition $\mathbb{R} \setminus \{\frac{-d}{c}\}$ et sa dérivée est donnée par

$$\forall x \neq -\frac{d}{c}, \qquad f'(x) = \frac{ad - bc}{(cx+d)^2}.$$

Proposition 4.8.

Soit $f : I \to J$ dérivable en a et $g : J \to \mathbb{R}$ dérivable en $f(a)$. Alors, la fonction composée $g \circ f$ est dérivable en a et $(g \circ f)'(a) = g' \circ f(a) f'(a)$.

Pour étudier les fonctions de plusieurs variables, on introduit la notion de dérivée partielle (qui correspond intuitivement à la dérivation selon une variable en considérant toutes les autres comme des constantes).

Définition 4.9. Dérivation partielle

Soit U une partie de $\mathbb{R}^n$, $f : U \to \mathbb{R}$, $(a_1, \dots, a_n) \in U$. La fonction f est dérivable selon la $k^{\text{ième}}$ variable au point $(a_1, \dots, a_n)$ si la fonction réelle

$$\varphi : x \mapsto f(a_1, \dots, a_{k-1}, x, a_{k+1}, \dots, a_n)$$

est dérivable en a_k. La quantité $\varphi'(a_k)$ est alors notée $\frac{\partial f}{\partial x_k}(a_1, \dots, a_n)$.

Comme on se ramène pour cette définition à la dérivation d'une fonction avec une seule variable, les règles de manipulations pour les dérivations partielles se déduisent des propositions précédentes.
Le principal intérêt du calcul de dérivées est le résultat suivant.

Proposition 4.10.

Soit f dérivable sur un intervalle I.
- La fonction f est constante si, et seulement si, $f' = 0$ sur I.
- La fonction f est croissante si, et seulement si, $f' \geq 0$ sur I.
- La fonction f est décroissante si, et seulement si, $f' \leq 0$ sur I.

De plus,
- Si $f' > 0$ sur I, alors f est strictement croissante sur I.
- Si $f' < 0$ sur I, alors f est strictement décroissante sur I.

Remarque

Il est ici essentiel de travailler sur un intervalle. Par exemple, la fonction $f : x \mapsto \frac{1}{x}$ est dérivable à dérivée strictement négative sur $\mathbb{R}^*$ et pourtant elle n'est pas décroissante puisque $f(-1) < f(1)$ et $-1 < 1$.

Ce résultat est évidemment essentiel pour établir des tableaux de variation et donc pour obtenir l'allure d'une courbe représentative, déterminer le signe d'une fonction ou estimer le nombre de solutions d'une équation de la forme $f(x) = b$ avec b fixé.

1.4. Bijections réelles

Définition 4.11.

Une fonction f réalise une bijection de I dans J si, tout $y \in J$ admet un unique antécédent $x \in I$ par f.

Graphiquement, la courbe représentative d'une bijection (respectivement d'une surjection ou d'une injection) de I dans J coupe une et une seule (respectivement au moins une ou au plus une) fois les droites horizontales $y = a$ où $a \in J$.

Rappelons le théorème suivant que nous utiliserons tout au long de ce chapitre.

Théorème 4.12. Théorème de la bijection

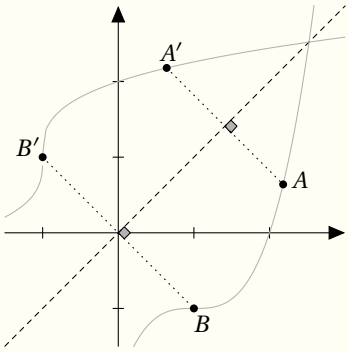
Une fonction f continue strictement croissante sur un intervalle I réalise une bijection de I dans l'intervalle $f(I)$.
Si, de plus, f est dérivable et que sa dérivée ne s'annule pas, alors f^{-1} est dérivable et

$$\forall y \in f(I), \qquad (f^{-1})'(y) = \frac{1}{f' \circ f^{-1}(y)}.$$

Remarque

Les courbes représentatives de f et de f^{-1} sont symétriques par rapport à la droite d'équation $y = x$. En effet,

$$(x, y) \in \mathscr{C}_f \quad \Leftrightarrow \quad y = f(x) \quad \Leftrightarrow \quad x = f^{-1}(y) \quad \Leftrightarrow \quad (y, x) \in \mathscr{C}_{f^{-1}}.$$



Les points A et B appartiennent à la courbe représentative de f donc leurs symétriques, A' et B', sont des points de la courbe représentative de f^{-1}. La même propriété de symétrie se trouve pour les tangentes ; ceci permet de comprendre l'hypothèse « f' ne s'annule pas » dans le théorème de bijection : si $f'(x) = 0$,

la tangente au point $(x, f(x))$ à la courbe représentative de f est horizontale donc la tangente au point $(f(x), x)$ à la courbe représentative de f^{-1} est verticale ; en conclusion, la fonction f^{-1} n'est pas dérivable en $f(x)$.

2. Valeur absolue

2.1. Relation d'ordre canonique

La propriété suivante (admise) résume les rapports entre la relation « inférieur ou égal » et les opérations sur $\mathbb{R}$.

Proposition 4.13.

La relation d'ordre total $\leq$ sur $\mathbb{R}$ vérifie, pour tous x, y, $z \in \mathbb{R}$,
- si $x \leq y$, alors $x + z \leq y + z$;
- si $x \leq y$ et $0 \leq z$, alors $xz \leq yz$.

On en déduit immédiatement les résultats suivants.

Proposition 4.14.

▹ Pour tout $x \in \mathbb{R}$, $0 \leq x$ si, et seulement si, $-x \leq 0$.
▹ Pour tout $(x, y) \in \mathbb{R}^2$, les conditions $0 \leq x$ et $0 \leq y$ entraînent $0 \leq xy$ et $0 \leq x + y$.

Démonstration

- Pour le premier point, il suffit d'appliquer la première propriété de la proposition précédente avec $z = x$ ou $z = -x$ pour avoir chacune des implications.
- La première provient de $0 = 0.x \leq y.x$. La seconde est obtenue par transitivité $0 \leq 0 + y \leq x + y$.

■

Définition 4.15. Intervalles de $\mathbb{R}$

Soit $x \leq y$. Les intervalles de bornes x et y sont les ensembles de $\mathbb{R}$ définis par

- $[x, y] = \{z \in \mathbb{R},\ x \leq z \leq y\}$
- $]x, y] = \{z \in \mathbb{R},\ x < z \leq y\}$
- $[x, y[= \{z \in \mathbb{R},\ x \leq z < y\}$
- $]x, y[= \{z \in \mathbb{R},\ x < z < y\}$

On étend sans problème les définitions d'intervalles aux cas $[x, +\infty[$, $]x, +\infty[$, $]-\infty, x]$, $]-\infty, x[$ et $]-\infty, +\infty[= \mathbb{R}$.

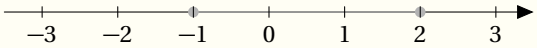
2.2. Valeur absolue

Définition 4.16.

La valeur absolue d'un élément $x \in \mathbb{R}$ est l'élément $|x|$ défini par $\max\{x, -x\}$.

Remarque

On remarque immédiatement que la condition $|x| \leq a$ est équivalente à $-a \leq x \leq a$ et que la condition $|x-\ell| \leq a$ est équivalente à $\ell - a \leq x \leq \ell + a$. Les réels x vérifiant la condition $|x-\ell| \leq a$ forment donc l'intervalle $[\ell - a, \ell + a]$: ce sont les réels à distance de ℓ inférieure ou égale à a.
Voici par exemple l'illustration de l'ensemble des réels x tels que $|x - \frac{1}{2}| \leq \frac{3}{2}$:



De manière plus générale, pour tous $a \leq b$, on a la description suivante d'un intervalle avec la valeur absolue :

$$[a, b] = \left\{x \in \mathbb{R}, \left|x - \tfrac{a+b}{2}\right| \leq \tfrac{b-a}{2}\right\}.$$

Remarque

Par définition, on a $0 \leq |x|$ pour tout réel x. Le cas d'égalité $|x| = 0$ correspond exactement à $x = 0$.

Proposition 4.17. Inégalités triangulaires

Pour tous $x,\ y \in \mathbb{R}$, $|x+y| \leq |x| + |y|$ et $||x| - |y|| \leq |x - y|$.

Démonstration

Soit $x,\ y \in \mathbb{R}$.
▹ Comme $-|x| \leq x \leq |x|$ et $-|y| \leq y \leq |y|$, on obtient $-(|x| + |y|) \leq x + y \leq |x| + |y|$, soit $|x+y| \leq |x| + |y|$.
▹ $|x| = |x - y + y| \leq |x - y| + |y|$ d'après le premier point. Par conséquent, $|x| - |y| \leq |x - y|$. Par symétrie, $|y| - |x| \leq |x - y|$ et donc $||x| - |y|| \leq |x - y|$. ■

Remarque

L'égalité $|x+y| = |x| + |y|$ n'a lieu que pour des réels x et y de même signe (il suffit de discuter selon le signe des réels x et y) :

- si x et y sont de même signe, il y a égalité ;
- sinon, on peut supposer sans perte de généralité $x > 0 > y$. L'égalité devient alors $|x+y| = x - y$ ce qui entraîne, selon le signe de $x + y$, $x = 0$ ou $y = 0$ ce qui est contradictoire avec notre hypothèse.

Proposition 4.18.

Pour tous $x,\ y \in \mathbb{R}$, $|xy| = |x||y|$.

Démonstration

Il suffit de discuter selon les signes de x et de y. ■

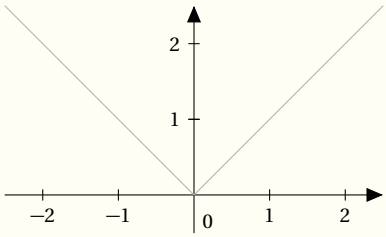
Exemple

Soit x, x', y et $y' \in \mathbb{R}$ tels que $|x-x'| \leq \varepsilon$ et $|y-y'| \leq \varepsilon$. Majorons $|xx'-yy'|$.
Pour cela, faisons apparaître un terme médian dans les sommes de sorte à ce qu'entre deux termes consécutifs seul un facteur du produit change.

$$\begin{aligned}|xy-x'y'| &= |xy-x'y+x'y-x'y'| \\ &\leq |xy-x'y|+|x'y-x'y'| \\ &\leq |y||x-x'|+|x'||y-y'| \leq (|y|+|x'|)\varepsilon.\end{aligned}$$

Remarque

La fonction $x \mapsto |x|$ est définie sur $\mathbb{R}$, continue, dérivable sauf en 0 et sa courbe représentative est



2.3. Parties positives et négatives

Définition 4.19.

La partie positive d'un réel x (respectivement la partie négative) est le réel positif x^+ (respectivement x^-) défini par $x^+ = \max(x,0)$ (respectivement $x^- = \max(-x,0)$).

Proposition 4.20.

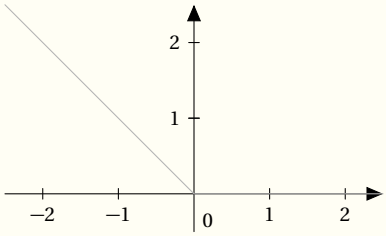
Pour tout $x \in \mathbb{R}$, $x = x^+ - x^-$ et $|x| = x^+ + x^-$.
Par conséquent, pour tout $x \in \mathbb{R}$, $x^+ = \frac{x+|x|}{2}$ et $x^- = \frac{-x+|x|}{2}$.

Démonstration

Il suffit de discuter selon que $x \geq 0$ ou $x < 0$. ■

Remarque

D'après les propriétés correspondantes de la valeur absolue, les fonctions $x \mapsto x^+$ et $x \mapsto x^-$ sont définies sur $\mathbb{R}$, continues, dérivables sauf en 0. Voici la courbe représentative de la seconde :



3. Exponentielles, logarithmes, puissances

3.1. Exponentielle réelle

Définition 4.21.

La fonction exponentielle exp est l'unique fonction solution sur $\mathbb{R}$ de l'équation $y' = y$ avec la condition initiale $y(0) = 1$.

Commençons par une propriété caractéristique de l'exponentielle.

Proposition 4.22. Propriété de morphisme

Pour tous x, $x' \in \mathbb{R}$, $\exp(x + x') = \exp(x)\exp(x')$.
En particulier, pour tout $x \in \mathbb{R}$, $(\exp(x))^{-1} = \exp(-x)$.

Démonstration
Pour tout x' fixé, les fonctions $x \mapsto \exp(x + x')$ et $x \mapsto \exp(x)\exp(x')$ sont solutions de l'équation différentielle $y' = y$ avec la même condition initiale $y(0) = \exp(x')$ donc sont égales. ■

Corollaire 4.23.

Pour tout $x \in \mathbb{R}$ et $n \in \mathbb{Z}$, $\exp(nx) = (\exp(x))^n$.

Démonstration
Pour $n \in \mathbb{N}$, il suffit de procéder par récurrence avec la propriété de morphisme. On passe ensuite à un entier n négatif avec la propriété de morphisme puisque, pour tout $x \in \mathbb{R}$ et tout $n \in \mathbb{N}$, $\exp(nx)\exp(-nx) = \exp(0) = 1$. ■

Étudions dorénavant la fonction exp. D'après l'équation différentielle, on déduit le tableau de variation suivant

x	$-\infty$		$+\infty$
$\exp'(x)$		$+$	
exp	0	$\nearrow$	$+\infty$

En effet, la fonction exp ne s'annule pas car sinon elle serait égale à la fonction nulle (par unicité de la solution à l'équation $y' = y$ avec une condition initiale nulle). Par conséquent, elle est strictement positive donc strictement croissante car égale à sa dérivée.
Pour la limite en $+\infty$, on remarque que, pour tout $x \geq 0$, $\exp'(x) = \exp(x) \geq \exp(0) = 1$ et donc en intégrant entre 0 et x, $\exp(x) \geq x + 1$ d'où la limite $+\infty$ en $+\infty$.
Pour la limite en $-\infty$, il suffit de remarquer que

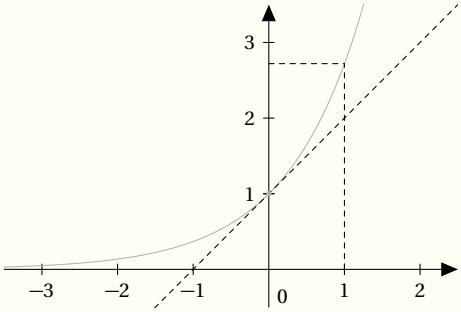
$$\lim_{x \to -\infty} \exp(x) = \lim_{x \to +\infty} \exp(-x) = \lim_{x \to +\infty} \frac{1}{\exp(x)} = 0.$$

On a établi la proposition suivante

Proposition 4.24.

La fonction exp réalise une bijection de classe $\mathscr{C}^1$, strictement monotone, entre $\mathbb{R}$ et $\mathbb{R}_+^*$.

On en déduit l'allure suivante pour la courbe représentative de exp.



On remarque que la courbe représentative de exp est au-dessus de sa tangente en 0. On peut aussi le voir comme le cas simple ($n = 1$) de la propriété suivante.

Proposition 4.25.

Pour tout $n \in \mathbb{N}$,

$$\forall x \in \mathbb{R}_+, \qquad \exp(x) \geq \sum_{k=0}^{n} \frac{x^k}{k!}.$$

Démonstration

Montrons, par récurrence sur $n \in \mathbb{N}$, que

$$\forall x \in \mathbb{R}_+, \qquad \exp(x) \geq \sum_{k=0}^{n} \frac{x^k}{k!}.$$

- Le résultat est vrai pour $n = 0$ car, par croissance de l'exponentielle, pour tout $x \in \mathbb{R}_+$, $\exp(x) \geq \exp(0) = 1$.
- Soit $n \in \mathbb{N}$ tel que, pour tout $x \in \mathbb{R}_+$,

  $$\exp(x) \geq \sum_{k=0}^{n} \frac{x^k}{k!}.$$

  Intégrons cette relation entre 0 et $x \in \mathbb{R}_+$,

  $$\exp(x) - 1 \geq \sum_{k=0}^{n} \frac{x^{k+1}}{k!(k+1)}.$$

  Ainsi,

  $$\exp(x) \geq 1 + \sum_{k=0}^{n} \frac{x^{k+1}}{(k+1)!} = \sum_{k=0}^{n+1} \frac{x^k}{k!}.$$

■

Remarque
Cette inégalité n'est pas vraie sur $\mathbb{R}$. En revanche, on dispose toujours de $e^x \geq 1 + x$ pour tout $x \in \mathbb{R}$.

Notation
On pose $e = \exp(1)$ et on note $\exp(x) = e^x$ pour tout $x \in \mathbb{R}$.

3.2. Logarithme

Définition 4.26.

La fonction logarithme népérien $\ln : \mathbb{R}_+^* \to \mathbb{R}$ est la bijection réciproque de exp.

D'après le théorème de la bijection (car exp′ ne s'annule pas), on a le résultat suivant.

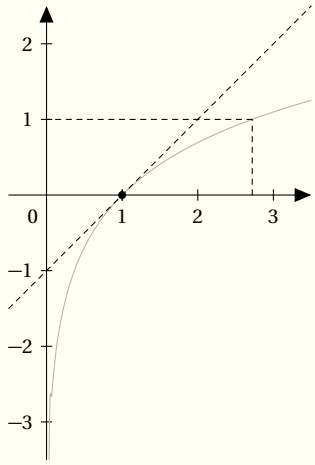
Proposition 4.27.

La fonction ln est de classe $\mathscr{C}^1$ sur $\mathbb{R}_+^*$ et sa dérivée est $x \mapsto \frac{1}{x}$.

Plus précisément, on obtient le tableau de variation suivant

x	0		$+\infty$
$\frac{1}{x}$	$\Vert$	$+$	
$\ln$	$\Vert$ $-\infty$	$\nearrow$	$+\infty$

d'où l'on déduit l'allure de la courbe représentative de ln.



Remarque
On fera attention que la fonction définie sur $\mathbb{R}^*$ par $f(x) = \ln|x|$ admet aussi pour dérivée $x \longmapsto \frac{1}{x}$.

Conseils méthodologiques

Il peut être utile de savoir utiliser la dérivée logarithmique d'une fonction f dérivable et strictement positive, c'est-à-dire la dérivée du logarithme de f qui vaut $\frac{f'}{f}$. Par exemple, pour étudier la monotonie d'une fonction f, il est équivalent d'étudier la monotonie de $\ln(f)$; or, il est plus simple de calculer la dérivée logarithmique d'un produit à l'aide de la propriété de morphisme. Réciproquement, on peut reconnaître une fonction de la forme $\frac{u'}{u}$ comme la dérivée logarithmique de u pour en obtenir une primitive.

La propriété de morphisme de l'exponentielle donne une propriété analogue pour ln.

Proposition 4.28. Propriété de morphisme

Pour tous $x,\ y \in \mathbb{R}_+^*$, $\ln(xy) = \ln(x) + \ln(y)$.
En particulier, pour tout $x \in \mathbb{R}_+^*$, $\ln \frac{1}{x} = -\ln(x)$.

Définition 4.29.

Le logarithme en base $a > 1$ est la fonction $\log_a : x \longmapsto \frac{\ln x}{\ln a}$.

Notation
La notation log (sans indice) est réservée au logarithme en base 10.

3.3. Puissances

Définition 4.30.

La fonction puissance d'exposant $r \in \mathbb{R}$ est la fonction définie sur $\mathbb{R}_+^*$ par $x \longmapsto e^{r\ln(x)}$.

Remarque
▹ Cette définition est bien cohérente avec celle déjà connue pour les entiers naturels

$$x^n = \underbrace{x \times \ldots \times x}_{n \text{ fois}}$$

et relatifs $x^{-n} = (x^{-1})^n$ pour $x \neq 0$.
▹ On peut prolonger un certain nombre de fonctions puissances à tout $\mathbb{R}$: celles qui correspondent à des exposants entiers (non nuls), aux inverses d'entiers impairs...

Les propriétés de morphisme de l'exponentielle et du logarithme permettent d'obtenir rapidement les règles de calculs suivantes.

Proposition 4.31.

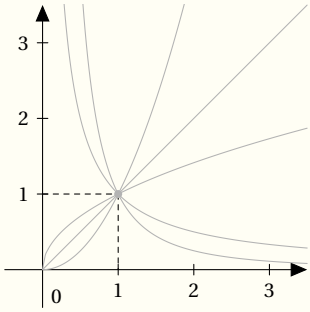
Pour tous $x,\ y \in \mathbb{R}_+^*$ et tous $r,\ s \in \mathbb{R}$,

- $x^r.x^s = x^{r+s}$
- $(x^r)^s = x^{rs}$
- $(xy)^r = x^r y^r$
- $x^{-r} = \frac{1}{x^r}$

Remarque

Il convient de faire attention aux notations : la quantité x^{a^b} signifie $x^{\left(a^b\right)}$ et cette quantité diffère en général de $(x^a)^b = x^{ab}$.

La fonction puissance d'exposant r est dérivable de dérivée $x \mapsto r\,x^{r-1}$. On obtient les allures suivantes pour les courbes correspondant aux exposants -2, -1, $\frac{1}{2}$, 1 et 2.



On peut, bien évidemment, composer les fonctions puissances avec d'autres fonctions et on obtient

Proposition 4.32.

Soit $u : \mathbb{R} \to \mathbb{R}_+^*$ et $v : \mathbb{R} \to \mathbb{R}$ deux fonctions dérivables. Alors, $x \mapsto (u(x))^{v(x)}$ est dérivable de dérivée

$$x \mapsto \left(v'(x)\ln(u(x)) + \frac{v(x)u'(x)}{u(x)}\right)(u(x))^{v(x)}.$$

Il est inutile d'apprendre cette formule mais il convient de savoir la retrouver sans erreur en revenant à la définition de $(u(x))^{v(x)} = e^{v(x)\ln(u(x))}$ et en utilisant la formule de dérivation d'une composée.

Exemple

Étudions la fonction $f : x \mapsto \left(1+\frac{1}{x}\right)^x$.
Cette fonction est définie et dérivable sur $]-\infty,-1[\cup]0,+\infty[$. Pour tout x dans ce domaine, on a

$$f'(x) = \left(\ln\left(1+\frac{1}{x}\right) - \frac{1}{1+x}\right)\left(1+\frac{1}{x}\right)^x.$$

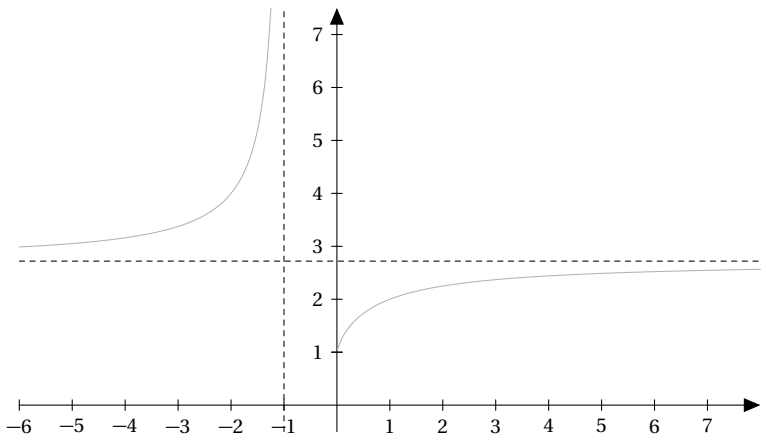
Pour étudier le signe de cette quantité, il suffit d'étudier la fonction auxiliaire $g : x \mapsto \ln\left(1+\frac{1}{x}\right) - \frac{1}{1+x}$ dont la dérivée vaut $g'(x) = -\frac{1}{x(x+1)^2}$. D'où le tableau de variation suivant :

x	$-\infty$		-1		0		$+\infty$
$g'(x)$		$+$	‖		‖	$-$	
g	0	↗	‖		‖	↘	0

g (et, par conséquent, f') est donc positive et le tableau de variation de f est

x	$-\infty$		-1		0		$+\infty$
$f'(x)$		$+$	‖		‖	$+$	
f	e	↗	$+\infty$ ‖		‖ 1	↗	e

L'allure de la courbe représentative de f est donc



Cherchons à comparer les fonctions puissances aux autres fonctions usuelles déjà étudiées.

Proposition 4.33. Croissances comparées

Pour tout $r > 0$,

$$\frac{\ln x}{x^r} \underset{x\to+\infty}{\longrightarrow} 0 \qquad \frac{x^r}{e^x} \underset{x\to+\infty}{\longrightarrow} 0.$$

$$x^r \ln x \underset{x\to 0}{\longrightarrow} 0.$$

$$|x|^r e^x \underset{x\to-\infty}{\longrightarrow} 0.$$

Démonstration
Démontrons la première relation (les suivantes découlent de la composition des limites et des propriétés de morphisme). Soit $u \in]1-r, 1[$. Alors,

$$\begin{aligned} 0 \leq \frac{\ln x}{x^r} &= \frac{1}{x^r}\int_1^x \frac{1}{t}\,\mathrm{d}t \leq \frac{1}{x^r}\int_1^x \frac{1}{t^u}\,\mathrm{d}t \\ &\leq \frac{1}{x^r}\frac{x^{1-u}}{1-u} = \frac{x^{1-u-r}}{1-u}. \end{aligned}$$

Or, $1-u-r<0$ par choix de u donc par encadrement $\frac{\ln x}{x^r} \underset{x\to+\infty}{\longrightarrow} 0$. ■

Heuristiquement, on vient d'établir qu'en $+\infty$, ln est négligeable devant les puissances qui sont elles-mêmes négligeables devant l'exponentielle. Pour calculer une limite, on mettra souvent le terme prépondérant en facteur puis on utilisera les croissances comparées.

3.4. Fonctions hyperboliques

Définition 4.34.

Les fonctions hyperboliques sont définies sur $\mathbb{R}$ par

$$\mathrm{sh} : x \mapsto \frac{e^x - e^{-x}}{2} \qquad \mathrm{ch} : x \mapsto \frac{e^x + e^{-x}}{2} \qquad \mathrm{th} : x \mapsto \frac{\mathrm{sh}(x)}{\mathrm{ch}(x)}.$$

Proposition 4.35.

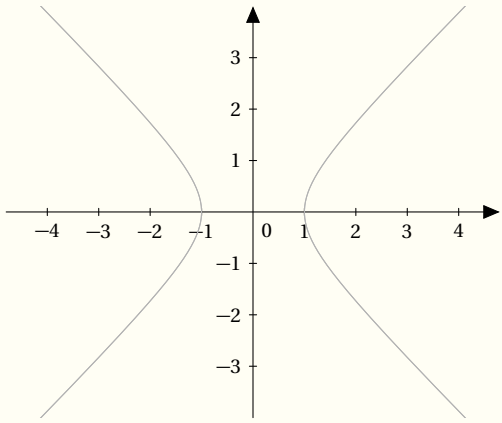
Pour tout $x \in \mathbb{R}$, $\mathrm{ch}^2(x) - \mathrm{sh}^2(x) = 1$.

Démonstration
Soit $x \in \mathbb{R}$. Alors,

$$\begin{aligned} \mathrm{ch}^2(x) - \mathrm{sh}^2(x) &= \frac{e^{2x}+2+e^{-2x}}{4} - \frac{e^{2x}-2+e^{-2x}}{4} \\ &= \frac{4}{4} = 1. \end{aligned}$$

■

Remarque
Cette propriété justifie l'appellation « hyperbolique » puisque $X^2 - Y^2 = 1$ est l'équation réduite d'une hyperbole. Plus précisément, cette courbe est composée de deux branches symétriques par rapport à l'axe des ordonnées et les points du plan de coordonnées de la forme $(\mathrm{ch}(x), \mathrm{sh}(x))$ décrivent la branche « droite » de celle-ci (de la même manière que les points du plan de coordonnées de la forme $(\cos(x), \sin(x))$ décrivent le cercle unité).



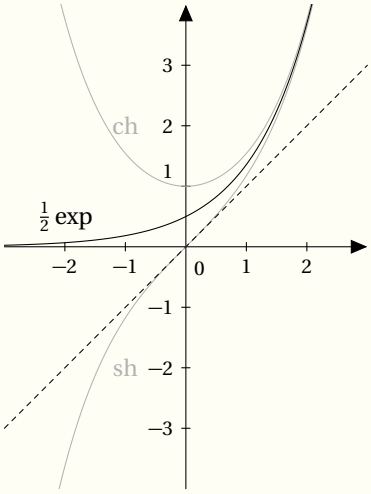
On obtient directement de la définition les propriétés suivantes.

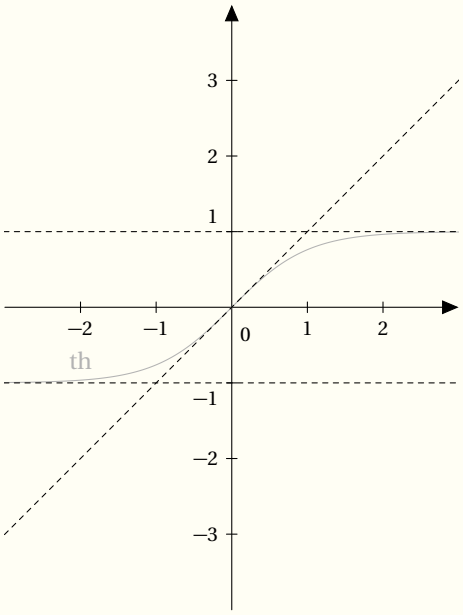
Proposition 4.36.

▹ La fonction ch est une fonction paire, alors que les fonctions sh et th sont impaires.
▹ Les trois fonctions hyperboliques sont de classe $\mathscr{C}^\infty$ et

$$\mathrm{ch}' = \mathrm{sh} \qquad \mathrm{sh}' = \mathrm{ch} \qquad \mathrm{th}' = \frac{1}{\mathrm{ch}^2} = 1 - \mathrm{th}^2 .$$

Voici le tracé des trois fonctions de trigonométrie hyperbolique. On remarque notamment le comportement asymptotique de ch, sh et $x \mapsto \frac{1}{2}\exp(x)$ en $+\infty$.





4. Exponentielle complexe, fonctions circulaires

4.1. Exponentielle complexe

Définition 4.37.

Pour tout complexe $z = x + iy$ avec $x, y \in \mathbb{R}$, on pose $e^z = e^x(\cos(y) + i\sin(y))$.

Remarque

L'application $\exp : \mathbb{C} \to \mathbb{C}$ n'est ni injective (pour tout $z \in \mathbb{C}$, $\exp(z) = \exp(z + i2\pi)$), ni surjective (0 n'est pas atteint).

Proposition 4.38.

L'application $\exp : \mathbb{C} \to \mathbb{C}^*$ est surjective.

Démonstration

Pour tout $z = \rho.e^{i\theta}$ avec $\rho > 0$, $z = \exp(\ln(\rho) + i\theta)$. ■

Remarque

La non-injectivité de la fonction exponentielle pose problème lorsqu'il s'agit de définir un « logarithme complexe » ; il faut alors se demander comment gérer les arguments de sorte à conserver la propriété de morphisme : l'utilisation d'une telle fonction est en dehors du programme.

4.2. Fonctions trigonométriques circulaires

Commençons par rappeler les relations avec l'exponentielle complexe.

Proposition 4.39. Formules d'Euler

Pour tout $x \in \mathbb{R}$,

$$\cos(x) = \frac{e^{ix} + e^{-ix}}{2}, \qquad \sin(x) = \frac{e^{ix} - e^{-ix}}{2i}.$$

Remarque

On peut alors établir un lien « formel » avec les fonctions hyperboliques qui permet entre autres de retrouver les formules de trigonométrie

$$\cos(x) = \mathrm{ch}(ix) \qquad \sin(x) = -i\,\mathrm{sh}(ix).$$

On veillera à ne pas utiliser ces relations dans nos raisonnements et à leur conserver leur rôle mnémotechnique.

Nous avons déjà vu les formules d'addition. En voici une nouvelle application.

Exemple

Calculons $\cos\frac{2\pi}{5}$. En utilisant que la somme des racines 5ièmes de l'unité est nulle, on obtient

$$1+2\cos\frac{2\pi}{5}+2\cos\frac{4\pi}{5}=0.$$

D'après la formule de duplication du cosinus,

$$1+2\cos\frac{2\pi}{5}+2\left(2\cos^2\frac{2\pi}{5}-1\right)=0,$$

donc $\cos\frac{2\pi}{5}$ est une racine de $4X^2+2X-1$. Les racines de ce polynôme sont $\frac{-1\pm\sqrt{5}}{4}$. Or, $\frac{2\pi}{5}\in[0,\frac{\pi}{2}]$ donc $\cos\frac{2\pi}{5}\geq 0$ et par conséquent, $\cos\frac{2\pi}{5}=\frac{-1+\sqrt{5}}{4}$.

On utilise aussi abondamment les formules suivantes pour le calcul intégral.

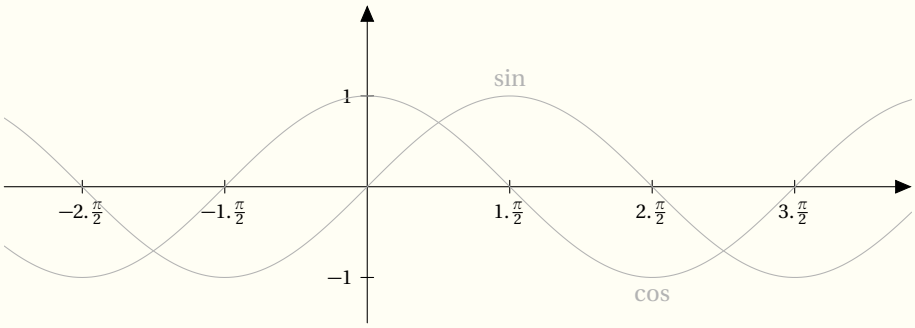
Proposition 4.40.

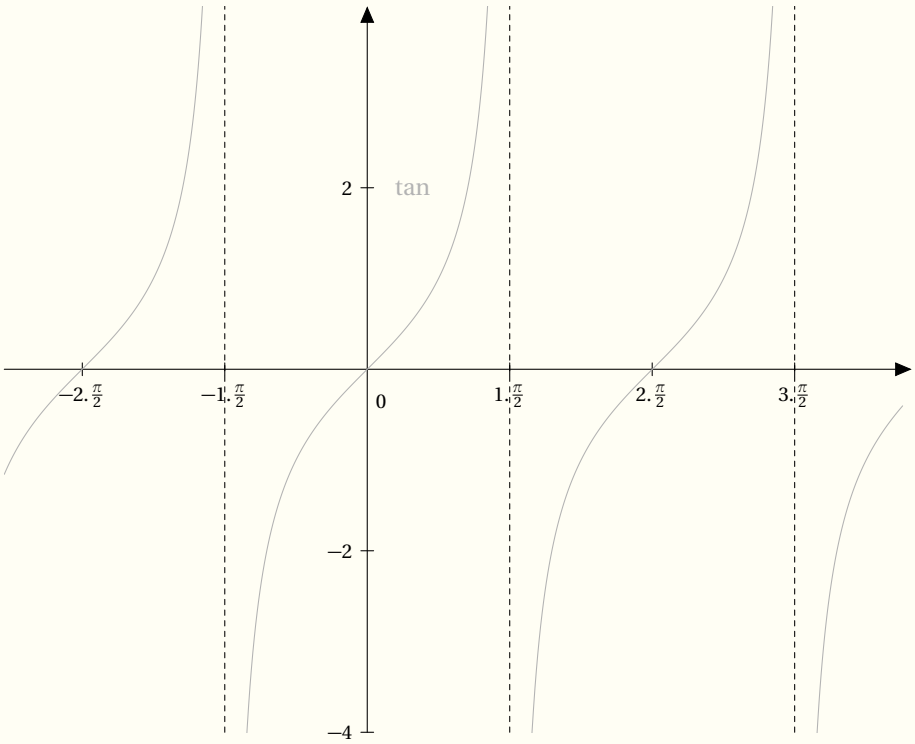
Pour tout $x\in\mathbb{R}\setminus\pi\mathbb{Z}$,

$$\sin(x)=\frac{2t}{1+t^2},\qquad \cos(x)=\frac{1-t^2}{1+t^2},\qquad \tan(x)=\frac{2t}{1-t^2}.$$

où $t=\tan\frac{x}{2}$.

On rappelle également l'expression des dérivées $\cos'=-\sin$, $\sin'=\cos$ et $\tan'=\frac{1}{\cos^2}=1+\tan^2$ qui permettent d'obtenir les courbes suivantes.





4.3. Fonctions trigonométriques réciproques

Définition 4.41.

▹ La fonction $\arccos : [-1,1] \to [0,\pi]$ est la bijection réciproque de $\cos : [0,\pi] \to [-1,1]$.
▹ La fonction $\arcsin : [-1,1] \to [-\frac{\pi}{2}, \frac{\pi}{2}]$ est la bijection réciproque de $\sin : [-\frac{\pi}{2}, \frac{\pi}{2}] \to [-1,1]$.
▹ La fonction $\arctan : \mathbb{R} \to]-\frac{\pi}{2}, \frac{\pi}{2}[$ est la bijection réciproque de $\tan :]-\frac{\pi}{2}, \frac{\pi}{2}[\to \mathbb{R}$.

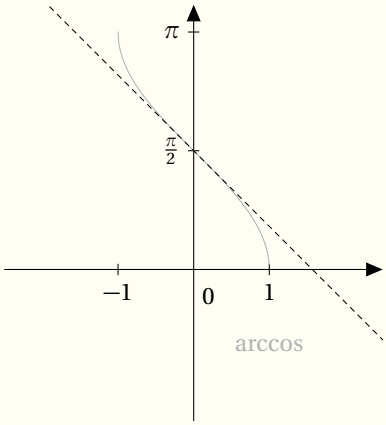
En appliquant le théorème de la bijection, on a le résultat suivant.

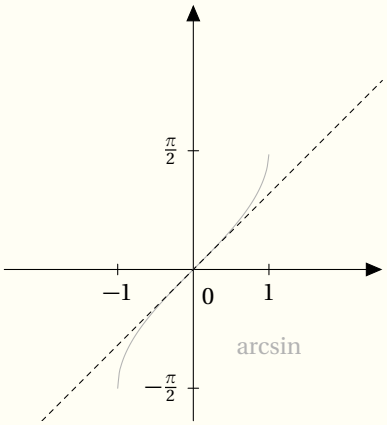
Proposition 4.42.

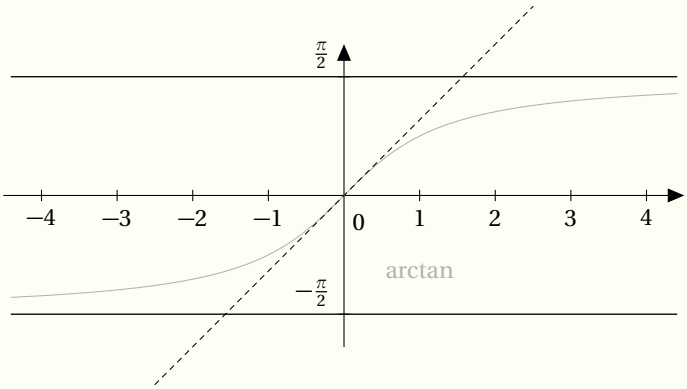
Les fonctions arccos et arcsin sont dérivables sur $]-1,1[$, arctan est dérivable sur $\mathbb{R}$. Les dérivées correspondantes sont

$$\arcsin'(y) = \frac{1}{\sqrt{1-y^2}} \qquad \arccos'(y) = \frac{-1}{\sqrt{1-y^2}} \qquad \arctan'(y) = \frac{1}{1+y^2}.$$

On en déduit les allures suivantes pour les courbes des fonctions circulaires réciproques







Il existe de nombreuses identités reliant ces fonctions entre elles.

Proposition 4.43.

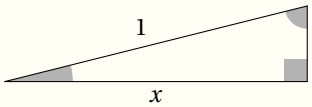
$$\forall x \in [-1,1], \qquad \arccos(x) + \arcsin(x) = \frac{\pi}{2}.$$

Démonstration

La fonction définie par le membre de gauche est continue sur $[-1,1]$, dérivable sur $]-1,1[$. Sa dérivée est nulle sur un intervalle, elle est donc constante ; or, elle vaut $\frac{\pi}{2}$ en 0 d'où l'égalité annoncée. ■

Remarque

Ce résultat peut se retrouver géométriquement pour $x \in [0,1]$ avec de la trigonométrie élémentaire dans un triangle rectangle dont l'hypoténuse est de longueur 1 et un autre côté de longueur x :



Proposition 4.44.

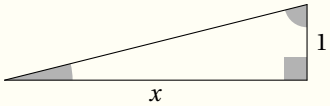
$$\forall x \in \mathbb{R}^*, \qquad \arctan(x) + \arctan\left(\frac{1}{x}\right) = \begin{cases} \frac{\pi}{2} & \text{si } x > 0, \\ -\frac{\pi}{2} & \text{si } x < 0. \end{cases}.$$

Démonstration

La fonction définie par le membre de gauche est continue sur $\mathbb{R}^*$, dérivable sur $\mathbb{R}^*$. Sa dérivée est nulle sur $\mathbb{R}_+^*$ et $\mathbb{R}_-^*$, elle est donc constante sur chacun de ces intervalles. On obtient le résultat en évaluant en 1 et en -1. ■

Remarque

Une fois encore, ce résultat peut se retrouver pour $x \in [0,1]$ avec de la trigonométrie élémentaire dans un triangle rectangle dont les deux côtés autre que l'hypoténuse sont de longueur x et 1 :



Par ailleurs, la propriété de bijection permet les simplifications suivantes (on fera bien attention aux intervalles).

Proposition 4.45.

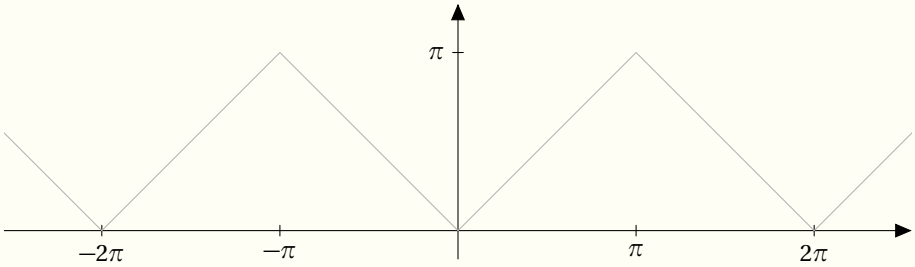
- $\forall y \in [-1,1], \quad \cos(\arccos(y)) = y$
- $\forall y \in [-1,1], \quad \sin(\arcsin(y)) = y$
- $\forall y \in \mathbb{R}, \quad \tan(\arctan(y)) = y$
- $\forall x \in [0,\pi], \quad \arccos(\cos(x)) = x$
- $\forall x \in [-\frac{\pi}{2}, \frac{\pi}{2}], \quad \arcsin(\sin(x)) = x$
- $\forall x \in]-\frac{\pi}{2}, \frac{\pi}{2}[, \quad \arctan(\tan(x)) = x$

Conseils méthodologiques

Pour simplifier une fonction trigonométrique et une fonction trigonométrique réciproque, on veillera à se ramener systématiquement à l'intervalle où elles sont effectivement réciproques.

Remarque

On remarque bien qu'en général, $\arccos(\cos(x)) \neq x$ comme on peut le voir sur la courbe représentative suivante de $\arccos \circ \cos$.



Pour justifier l'allure de cette courbe, trouvons, pour chaque x, un élément $y \in [0,\pi]$ tel que $\arccos(\cos x) = \arccos(\cos y)$ et donc $\arccos(\cos x) = y$.

- Pour tout $x \in [0, \pi]$, $\arccos(\cos x) = x$.
- Pour tout $x \in [-\pi, 0]$, $\arccos(\cos x) = \arccos(\cos(-x)) = -x$.

Plus généralement, fixons $k \in \mathbb{Z}$.

- Pour tout $x \in [2k\pi, (2k+1)\pi]$,
$$\arccos(\cos x) = \arccos(\cos(x - 2k\pi)) = x - 2k\pi.$$
- Pour tout $x \in [(2k+1)\pi, (2k+2)\pi]$,
$$\arccos(\cos x) = \arccos(\cos((2k+2)\pi - x)) = (2k+2)\pi - x.$$

Exemple

Simplifions la fonction $x \mapsto \sin(\arccos x)$.
Pour cela remarquons tout d'abord que arccos est à valeurs dans $[-\frac{\pi}{2}, \frac{\pi}{2}]$ et que sur cet intervalle la fonction sin est positive. Ainsi, pour tout $x \in [-1, 1]$,
$$\begin{aligned}\sin(\arccos x) &= \sqrt{1 - \cos^2(\arccos x)} \\ &= \sqrt{1 - x^2}.\end{aligned}$$

FICHE

SYNTHÈSE

Les premières compétences à maîtriser après ce chapitre sont la manipulation d'inégalités, des valeurs absolues, la lecture précise d'une courbe représentative de fonction et l'obtention de la courbe représentative à partir d'un tableau de variation.

Le principal théorème du chapitre est le théorème de la bijection.

Théorème de la bijection

Une fonction f continue strictement croissante sur un intervalle I réalise une bijection de I dans l'intervalle $f(I)$. Si, de plus, f est dérivable et que sa dérivée ne s'annule pas, alors f^{-1} est dérivable et

$$\forall y \in f(I), \qquad (f^{-1})'(y) = \frac{1}{f' \circ f^{-1}(y)}.$$

Parmi les fonctions usuelles, certaines jouissent d'une propriété de morphisme qui permet de les utiliser pour simplifier les calculs.

Propriétés de morphisme

- $\forall x, y \in \mathbb{R}$,
$$\exp(x+y) = \exp(x)\exp(y).$$
- $\forall x, y \in \mathbb{R}_+^*$,
$$\ln(xy) = \ln(x) + \ln(y).$$
- $\forall x, y \in \mathbb{R}_+^*$, $\forall r, s \in \mathbb{R}$,
$$x^r.x^s = x^{r+s}, \qquad (x^r)^s = x^{rs}, \qquad (xy)^r = x^r y^r, \qquad x^{-r} = \frac{1}{x^r}.$$

Les croissances comparées permettent de calculer un certain nombre de limites.

Croissances comparées

Pour tout $r > 0$,

$$\lim_{x \to +\infty} \frac{\ln x}{x^r} = 0 \qquad \lim_{x \to +\infty} \frac{x^r}{e^x} = 0.$$

$$\lim_{x \to 0} x^r \ln x = 0.$$

$$\lim_{x \to -\infty} |x|^r e^x = 0.$$

Les fonctions trigonométriques hyperboliques sont définies à partir de l'exponentielle réelle.

Fonctions trigonométriques hyperboliques

$$\mathrm{sh} : x \mapsto \frac{e^x - e^{-x}}{2} \qquad \mathrm{ch} : x \mapsto \frac{e^x + e^{-x}}{2} \qquad \mathrm{th} : x \mapsto \frac{\mathrm{sh}(x)}{\mathrm{ch}(x)}.$$

$$\forall x \in \mathbb{R}, \qquad \mathrm{ch}^2(x) - \mathrm{sh}^2(x) = 1.$$

Rappelons le parallèle avec les fonctions trigonométriques (circulaires).

Fonctions trigonométriques

$$\sin : x \mapsto \frac{e^{ix} - e^{-ix}}{2i} \qquad \cos : x \mapsto \frac{e^{ix} + e^{-ix}}{2} \qquad \tan : x \mapsto \frac{\sin(x)}{\cos(x)}.$$

$$\forall x \in \mathbb{R}, \qquad \cos^2(x) + \sin^2(x) = 1.$$

De nombreuses formules de trigonométrie complètent les formules d'additions déjà étudiées dans les chapitres précédents. Ici, on s'intéresse à leurs réciproques obtenues en se restreignant aux intervalles idoines.

Fonctions circulaires réciproques

- ▹ La fonction $\arccos : [-1,1] \to [0,\pi]$ est la bijection réciproque de $\cos : [0,\pi] \to [-1,1]$.
- ▹ La fonction $\arcsin : [-1,1] \to [-\frac{\pi}{2}, \frac{\pi}{2}]$ est la bijection réciproque de $\sin : [-\frac{\pi}{2}, \frac{\pi}{2}] \to [-1,1]$.
- ▹ La fonction $\arctan : \mathbb{R} \to]-\frac{\pi}{2}, \frac{\pi}{2}[$ est la bijection réciproque de $\tan :]-\frac{\pi}{2}, \frac{\pi}{2}[\to \mathbb{R}$.

Il est important de connaître les courbes représentatives des fonctions usuelles et de savoir y reconnaître un certain nombre de propriétés.

EXERCICES

Vrai ou faux ?

		Vrai	Faux
a)	Le produit de deux fonctions impaires est une fonction paire.	☐	☐
b)	Il n'existe pas de fonction à la fois paire et impaire.	☐	☐
c)	Le domaine de définition de la fonction $x \mapsto \frac{1}{x+\sqrt{x^2}}$ est $\mathbb{R}_+^*$.	☐	☐
d)	La fonction partie entière $x \mapsto \lfloor x \rfloor$ est impaire.	☐	☐
e)	Une fonction T-périodique est aussi $1789T$-périodique.	☐	☐
f)	La somme d'une fonction $3T$-périodique et $2T$-périodique est $6T$-périodique.	☐	☐
g)	La courbe représentative de la fonction cos est symétrique par rapport à l'axe d'équation $x = \frac{\pi}{2}$.	☐	☐
h)	Le point $(\pi, 0)$ est centre de symétrie de la courbe représentation de la fonction sin.	☐	☐
i)	La courbe représentative de $x \mapsto -f(1-x)$ est déduite de celle de f par symétrie autour du point $(\frac{1}{2}, 0)$.	☐	☐
j)	La fonction partie entière est 1-périodique.	☐	☐
k)	Le produit de deux fonctions croissantes est une fonction croissante.	☐	☐
l)	La fonction $f : \mathbb{R} \to \mathbb{R}$ est surjective si, et seulement si, toute droite horizontale coupe la courbe représentative de f.	☐	☐
m)	Pour tout $x \in [0, \pi]$, $\arcsin(\sin(x)) = x$.	☐	☐
n)	Pour tout $x \in [-1, 1]$, $\cos(\arccos(x)) = x$.	☐	☐

Exercices d'application du chapitre 4

●○○ **Exercice 1**

Montrer que la fonction $f : x \mapsto \ln(\sqrt{x^2+1} + x)$ est impaire.

●○○ **Exercice 2**

Considérons la fonction $f : x \mapsto \frac{1}{e^x - 1}$.

1. Calculer, pour tout $x > 0$, $f(x) + f(-x)$.
2. En déduire que la courbe représentative de f admet un centre de symétrie.

●○○ **Exercice 3**

Soit $f : x \mapsto \begin{cases} x+1 & \text{si } x \in \mathbb{Q}, \\ x & \text{sinon.} \end{cases}$

La fonction est-elle surjective de $\mathbb{R}$ dans $\mathbb{R}$? injective ? bijective ?

●●○ **Exercice 4**

Soit $a \in \mathbb{R}$. Définissons $f_a : \mathbb{R} \to \mathbb{R}$ par

$$f_a(x) = \begin{cases} x + a & \text{si } x \geq 0, \\ x - a & \text{sinon.} \end{cases}$$

1. Déterminer les réels a pour lesquelles f_a est surjective.
2. Déterminer les réels a pour lesquelles f_a est injective.

●○○ **Exercice 5**

Déterminer les réels x et y tels que $(x+y)^+ = x^+ + y^+$.

●●○ **Exercice 6**

Déterminer les réels x tels que $\cos\left(x - \frac{\pi}{4}\right) = \sin\left(2x - \frac{\pi}{3}\right)$.

●●○ **Exercice 7**

Déterminer les réels x tels que $(\cos x)^4 + (\sin x)^6 = 1$.

●○○ **Exercice 8**

Calculer $\frac{1}{\sin \frac{\pi}{18}} - \frac{\sqrt{3}}{\cos \frac{\pi}{18}}$.

●●○ **Exercice 9**

Soit $n \in \mathbb{N}$ et $x \in \mathbb{R}$ tels que $x \neq \frac{\pi}{2}[\pi]$ et $nx \neq \frac{\pi}{2}[\pi]$. Exprimer $\tan(nx)$ en fonction de $\tan(x)$.

●○○ **Exercice 10**

Déterminer les $x \in \mathbb{R}$ tels que $\ln(1 + x\,\mathrm{th}(x)) = x$.

●●○ **Exercice 11**

Trouver une condition nécessaire et suffisante sur a, $b \in \mathbb{R}$ pour qu'il existe $A \in \mathbb{R}^*$ et $\varphi \in R$ tels que

$$\forall\ x \in \mathbb{R}, \qquad a\,\mathrm{ch}(x) + b\,\mathrm{sh}(x) = A\,\mathrm{ch}(x + \varphi).$$

●●○ **Exercice 12**

Déterminer les réels $a > b$ pour lesquels le système

$$\begin{cases} \mathrm{ch}(x) & + & \mathrm{ch}(y) & = a \\ \mathrm{sh}(x) & + & \mathrm{sh}(y) & = b \end{cases}$$

admet des solutions.

●○○ **Exercice 13**

Montrer que, pour tout $x \neq 0$,

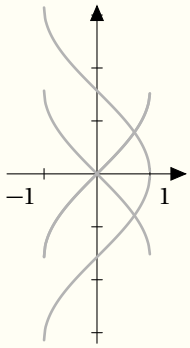
$$\frac{1}{\mathrm{sh}(x)} = \frac{1}{\mathrm{th}\,\frac{x}{2}} - \frac{1}{\mathrm{th}\,x}.$$

En déduire une expression simple de $\sum_{k=0}^{n} \frac{1}{\mathrm{sh}(2^k x)}$ puis sa limite quand n tend vers $+\infty$.

EXERCICES

●○○ **Exercice 14**

Reconnaître les courbes représentatives de $x \mapsto \arccos(x)$, $x \mapsto -\arccos(x)$, $x \mapsto \arcsin(x)$, $x \mapsto \arcsin(-x)$ et $x \mapsto \frac{\pi}{2} - \arccos(x)$ sur la figure suivante



●○○ **Exercice 15**

Calculer $\arcsin \frac{3}{5} + \arcsin \frac{5}{13}$.

●○○ **Exercice 16**

Calculer $f(x) = \arcsin x + \arcsin \sqrt{1-x^2}$ pour $x \in [0,1]$.

●●○ **Exercice 17**

Montrer que, pour tout $x > 0$, $\arctan \frac{1}{\sqrt{x}} = \arcsin \frac{1}{\sqrt{x+1}}$.

●○○ **Exercice 18**

1. Montrer que $\arctan(2\sqrt{2}) + 2\arctan(\sqrt{2}) = \pi$.
2. En déduire la valeur de $\arctan\left(\frac{1}{2\sqrt{2}}\right) + 2\arctan\left(\frac{1}{\sqrt{2}}\right)$.

●●○ **Exercice 19**

On admet l'égalité suivante entre complexes

$$(8+i)^6(57+i)^2(239+i) = 150837781250 + 150837781250i.$$

En déduire la formule

$$\frac{\pi}{4} = 6\arctan\frac{1}{8} + 2\arctan\frac{1}{57} + \arctan\frac{1}{239}.$$

Exercices d'approfondissement du chapitre 4

●●● **Exercice A**

Soit $k > 0$. Une fonction $f : \mathbb{R} \to \mathbb{R}$ est k-lipschitzienne si

$$\forall x, y \in \mathbb{R}, \qquad |f(x) - f(y)| \leq k|x-y|.$$

On admet qu'une fonction k-lipschitzienne est continue. Par ailleurs, pour une fonction $f : \mathbb{R} \to \mathbb{R}$, on définit les fonctions partie positive et partie négative de f par $f_+ : x \mapsto (f(x))^+$ et $f_- : x \mapsto (f(x))^-$

1. Montrer que si $f : \mathbb{R} \to \mathbb{R}$ est k-lipschitzienne, alors les fonctions f_+ et f_- sont aussi k-lipschitziennes.
2. Soit $f : \mathbb{R} \to \mathbb{R}$ telle que f_+ et f_- sont k-lipschitziennes. Montrer que f est k-lipschitzienne.

●●○ **Exercice B**

Déterminer l'abscisse du premier maximum local sur $\mathbb{R}_+^*$ de $f : x \mapsto \sum_{k=1}^{n} \frac{\sin(kx)}{k}$.

●●● **Exercice C**

Soit $t_1, \dots, t_{13} \in \mathbb{R}$. Montrer qu'il existe deux entiers distincts $i, j \in [\![1, 13]\!]$ tels que

$$0 \leq \frac{t_i - t_j}{1 + t_i t_j} \leq 2 - \sqrt{3}.$$

On pourra calculer $\tan \frac{\pi}{12}$.

●●● **Exercice D**

Notons dans cet exercice f la fonction arctan.

1. Montrer que, pour tout $n \in \mathbb{N} \setminus \{0\}$ et tout $x \in \mathbb{R}$,
$$f^{(n)}(x) = (n-1)!\big(\cos f(x)\big)^n \sin\left(n\left(f(x) + \frac{\pi}{2}\right)\right).$$
2. Déterminer les zéros de $f^{(n)}$ pour tout $n \in \mathbb{N} \setminus \{0\}$.

PROBLÈMES

Problèmes du chapitre 4

●○○ **Problème 1 Majoration de Huygens**

1. Soit $f : x \mapsto 2\sin x + \tan x - 3x$.
   a. Justifier que f est dérivable et calculer sa dérivée.
   b. Déterminer le signe pour $x \in [0, 1]$ de la quantité $2x^3 - 3x^2 + 1$.
   c. En déduire que
   $$\forall x \in [0, \frac{\pi}{2}[, \qquad 2\sin x + \tan x \geq 3x.$$
   d. L'inégalité est-elle encore vraie pour $x \in]\frac{\pi}{2}, \frac{3\pi}{2}[$?
2. A-t-on l'inégalité $2\operatorname{sh} x + \operatorname{th} x \geq 3x$ pour tout $x \in \mathbb{R}_+$?

●●○ **Problème 2**

1. Rappeler (sans démonstration) la formule donnant $\tan(x + y)$ en fonction de $\tan x$ et $\tan y$ pour des réels x et y pour lesquels ces trois termes sont définis.
2. Calculer $\arctan a + \arctan b$ pour
   a. $a, b \in [0, 1[$
   b. $a > 0 > b$
3. a. Préciser, pour tout $k \in \mathbb{N}$, $\arctan(k+1) - \arctan(k)$.
   b. En déduire $\lim \sum_{k=0}^{n} \arctan \frac{1}{k^2+k+1}$.
4. Soit $(F_n)_n$ la suite de Fibonacci définie par $F_0 = 0$, $F_1 = 1$ et
$$\forall n \in \mathbb{N}, \qquad F_{n+2} = F_n + F_{n+1}.$$
   a. Montrer que, pour tout $n \in \mathbb{N}$,
   $$F_{n+1}^2 - F_n F_{n+2} = (-1)^n.$$

b. Montrer que, pour tout $n \in \mathbb{N} \setminus \{0\}$,

$$\arctan \frac{1}{F_{2n+1}} + \arctan \frac{1}{F_{2n+2}} = \arctan \frac{1}{F_{2n}}.$$

c. En déduire $\lim \sum_{k=0}^{n} \arctan \frac{1}{F_{2k+1}}$.

●●● Problème 3 Trigonométrie lemniscatique

Partie A

Notons, pour tout $x \in [0,1[$, $\ell(x) = \int_0^x \frac{1}{\sqrt{1-t^4}}\,\mathrm{d}t$.

1. Justifier que l'application $x \mapsto \ell(x)$ définit une application strictement croissante sur $[0,1[$ que l'on notera ℓ.
2. Montrer que $\ell(x) \leq \arcsin x$ pour tout $x \in [0,1[$ puis en déduire que ℓ admet une limite finie en 1 que l'on notera σ (mais que l'on ne calculera pas).
   On étend ℓ en une nouvelle application de $[-1,1]$, encore notée ℓ, définie par

$$x \mapsto \begin{cases} \ell(x) & \text{si } x \in [0,1[, \\ \sigma & \text{si } x = 1, \\ -\ell(-x) & \text{sinon.} \end{cases}$$

3. Montrer que ℓ est une bijection entre $[-1,1]$ et un intervalle I que l'on précisera.

Partie B – Sinus lemniscatique

Définissons la fonction sinus lemniscatique sl comme la bijection réciproque de ℓ sur l'intervalle I, prolongée par symétrie par rapport à l'axe d'équation $x = \sigma$ puis par 4σ-périodicité.

1. Calculer $\mathrm{sl}(k\sigma)$ pour tout $k \in \mathbb{Z}$.
2. Montrer que sl est continue, dérivable puis deux fois dérivable.
   On pourra admettre la dérivabilité aux points $k\sigma$ si l'on n'a pas encore repris la notion de dérivabilité.

Partie C – Une équation différentielle

Considérons (E) l'équation différentielle $y'' + 2y^3 = 0$.

1. Montrer que la fonction sl est solution de (E).
   Considérons y une solution de (E).
2. Montrer que la fonction $(y')^2 + y^4$ est constante. Notons-la γ.
3. Déterminer l'application $x \mapsto \ell(\gamma^{-\frac{1}{4}} y(x))$ sur les intervalles où $y' \geq 0$.
4. En déduire les solutions de (E) dont la dérivée reste positive.

Partie D – Cosinus lemniscatique

Définissons la fonction cosinus lemniscatique cl par

$$\forall x \in \mathbb{R}, \qquad \mathrm{cl}(x) = \frac{\mathrm{sl}'(x)}{1 + \mathrm{sl}^2(x)}.$$

1. Simplifier la fonction $x \mapsto \mathrm{cl}^2(x) + \mathrm{sl}^2(x) + \mathrm{cl}^2(x)\mathrm{sl}^2(x)$.
2. Calculer cl' puis montrer que cl est solution de l'équation (E).
3. Montrer que, pour tout $x \in \mathbb{R}$, $\mathrm{cl}(x) = \mathrm{sl}(\sigma - x)$.
4. Calculer, pour tout $(x,y) \in \mathbb{R}^2$, $\mathrm{sl}(x+y)$ en fonction de $\mathrm{cl}(x)$, $\mathrm{cl}(y)$, $\mathrm{sl}(x)$ et $\mathrm{sl}(y)$.

CORRIGÉS

Corrigés des Vrai/Faux

a) Vrai.

b) Faux : il y a la fonction nulle.

c) Vrai.

d) Faux : $\lfloor 1.5 \rfloor = 1$ et $\lfloor -1.5 \rfloor = -2 \neq \pm 1$.

e) Vrai et d'ailleurs, pour tout $n \in \mathbb{Z}$, elle est aussi nT-périodique.

f) Vrai.

g) Faux. Par exemple, $\cos \pi \neq \cos 0$.

h) Vrai car, pour tout $x \in \mathbb{R}$, $\sin(2\pi - x) = -\sin x$.

i) Vrai.

j) Faux car, pour tout $x \in \mathbb{R}$, $\lfloor x+1 \rfloor = \lfloor x \rfloor + 1$.

k) Faux, cela dépend du signe des fonctions.

l) Vrai.

m) Faux car arcsin est à valeurs dans $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ donc $\arcsin(\sin(x)) \neq x$ pour $x > \frac{\pi}{2}$.

n) Vrai.

Corrigés des exercices d'application du chapitre 4

Exercice 1

La fonction f est définie sur l'intervalle $\mathbb{R}$ qui est symétrique par rapport à 0. De plus, pour tout $x \in \mathbb{R}$, en multipliant par la quantité conjuguée,

$$\begin{aligned} f(-x) &= \ln(\sqrt{x^2+1} - x) = \ln\left(\frac{(\sqrt{x^2+1} - x)(\sqrt{x^2+1} + x)}{(\sqrt{x^2+1} + x)}\right) \\ &= \ln \frac{1}{\sqrt{x^2+1} + x} = -f(x). \end{aligned}$$

En conclusion, f est impaire.

Remarque

La fonction f considérée n'est autre que la bijection réciproque de sh, habituellement notée argsh. En effet, pour tous x, $y \in \mathbb{R}$,

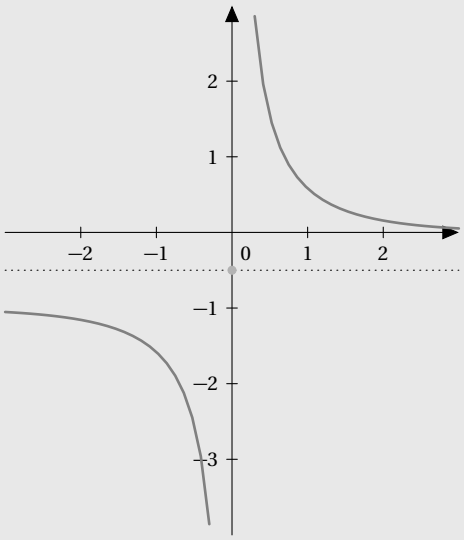
$$\begin{aligned} y = \mathrm{sh}(x) &\Leftrightarrow 2y = e^x - e^{-x} \\ &\Leftrightarrow e^{2x} - 2ye^x - 1 = 0 \\ &\Leftrightarrow e^x \in \{y + \sqrt{y^2+1}, y - \sqrt{y^2+1}\} \\ &\Leftrightarrow e^x = y + \sqrt{y^2+1} \qquad \text{car } y - \sqrt{y^2+1} < 0 \\ &\Leftrightarrow x = f(y). \end{aligned}$$

Exercice 2

1. Pour tout $x > 0$,

$$\begin{aligned} f(x) + f(-x) &= \frac{1}{e^x - 1} + \tfrac{1}{e^{-x}-1} \\ &= \frac{1}{e^x - 1} + \tfrac{e^x}{1-e^x} = -1. \end{aligned}$$

2. Pour tout $x > 0$, $f(0+x) + \frac{1}{2} = -\frac{1}{2} - f(0-x)$ donc le point $(0, -\frac{1}{2})$ est centre de symétrie de la courbe représentative de f.



Exercice 3

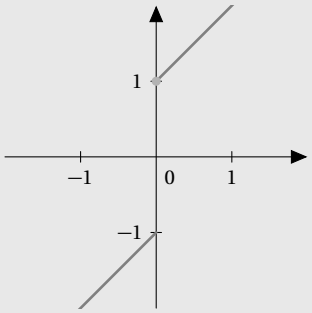
Montrons que la fonction f est bijective. Remarquons tout d'abord que $f(\mathbb{Q}) \subset \mathbb{Q}$ et $f(\mathbb{R} \setminus \mathbb{Q}) \subset \mathbb{R} \setminus \mathbb{Q}$. Soit $y \in \mathbb{R}$. Cherchons les antécédents de y.

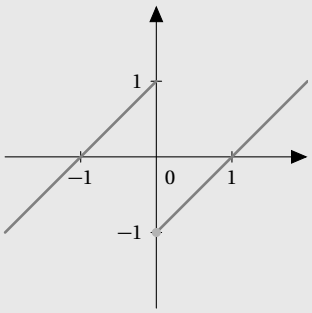
- Si $y \in \mathbb{Q}$, alors les antécédents de y appartiennent nécessairement à $\mathbb{Q}$ d'après la remarque liminaire et on trouve que $f(x) = y$ si, et seulement si, $x = y - 1$.
- Sinon, les antécédents de y appartiennent nécessairement à $\mathbb{R} \setminus \mathbb{Q}$ et on trouve que $f(x) = y$ si, et seulement si, $x = y$.

Ainsi, tout réel y admet un unique antécédent et donc f est bijective de $\mathbb{R}$ dans $\mathbb{R}$.

Exercice 4

Pour bien visualiser la situation il est bon de commencer par tracer la courbe représentative de ces fonctions. Voici, par exemple, le tracé des courbes représentatives de f_1 et de f_{-1}.





Nous allons montrer que f_a est surjective si, et seulement si, $a \leq 0$ et injective si, et seulement si, $a \geq 0$.

1. Remarquons que $f_a(\mathbb{R}_+) = [a, +\infty[$ et $f_a(\mathbb{R}_-^*) =]-\infty, -a[$: ainsi,

$$f_a(\mathbb{R}) = f_a(\mathbb{R}_+) \cup f_a(\mathbb{R}_-^*) =]-\infty, -a[\cup [a, +\infty[.$$

Si $a \leq 0$, alors $-a \geq a$ et donc $f(\mathbb{R}) = \mathbb{R}$: f_a est surjective. Sinon, $0 \in]-a, a[$ donc $0 \notin f_a(\mathbb{R})$: f_a n'est pas surjective.

2. Si $a < 0$, alors 0 admet deux antécédents par f : $a \in \mathbb{R}_-$ et $-a \in \mathbb{R}_+^*$. Sinon, le seul antécédent de $y \geq a$ est $x = y - a \in \mathbb{R}_+$; le seul antécédent de $y < a$ est $x = y + a \in \mathbb{R}_-^*$; enfin les réels de l'intervalle $[-a, a[$ n'admettent aucun antécédent. En conclusion, f_a est injective dans ce cas.

Exercice 5

On rappelle que pour tout $z \in \mathbb{R}$, $z^+ = \frac{1}{2}(z + |z|)$. Par conséquent, l'équation se réécrit

$$x + y + |x + y| = x + |x| + y + |y|,$$

soit encore $|x + y| = |x| + |y|$. D'après le cours, cette égalité n'est réalisée que lorsque x et y sont de même signe.

Exercice 6

Rappelons que, pour tout $\theta \in \mathbb{R}$, $\sin\theta = \cos\left(\frac{\pi}{2} - \theta\right)$. Ainsi, l'équation se réécrit

$$\cos\left(x - \frac{\pi}{4}\right) = \cos\left(\frac{\pi}{2} - 2x + \frac{\pi}{3}\right).$$

Par conséquent, les solutions sont les x tels que

$$x - \frac{\pi}{4} = \frac{\pi}{2} - 2x + \frac{\pi}{3}[2\pi]$$

ou

$$x - \frac{\pi}{4} = -\left(\frac{\pi}{2} - 2x + \frac{\pi}{3}\right)[2\pi].$$

On trouve donc l'ensemble des réels x tels que

$$x = \frac{13\pi}{36}\left[\frac{2\pi}{3}\right] \qquad \text{ou} \qquad x = \frac{7\pi}{12}[2\pi].$$

CORRIGÉS

Exercice 7

Pour tout $x \in \mathbb{R}$, $(\cos x)^4 \leq (\cos x)^2$ et $(\sin x)^6 \leq (\sin x)^2$. Par conséquent,

$$1 \leq (\cos x)^4 + (\sin x)^6 \leq (\cos x)^2 + (\sin x)^6 \leq (\cos x)^2 + (\sin x)^2 = 1.$$

Ainsi, pour une solution x de l'équation, chaque inégalité intermédiaire est une égalité soit $(\cos x)^4 = (\cos x)^2$ et $(\sin x)^6 = (\sin x)^2$ ce qui se traduit encore par $\cos x \in \{0, \pm 1\}$ et $\sin x \in \{0, \pm 1\}$. Les seules solutions sont donc les éléments x de $\frac{\pi}{2}\mathbb{Z}$.

Exercice 8

Commençons par mettre l'expression au même dénominateur ; le numérateur est alors de la forme $a \cos \theta + b \sin \theta$ et il est factorisé par $\sqrt{a^2 + b^2}$ pour reconnaître une formule d'addition.

$$\begin{aligned}\frac{1}{\sin \frac{\pi}{18}} - \frac{\sqrt{3}}{\cos \frac{\pi}{18}} &= \frac{2}{\sin \frac{\pi}{18} \cos \frac{\pi}{18}} \left(\frac{1}{2} \cos \frac{\pi}{18} - \frac{\sqrt{3}}{2} \sin \frac{\pi}{18} \right) \\ &= \frac{4}{\sin 2\frac{\pi}{18}} \sin\left(\frac{\pi}{6} - \frac{\pi}{18} \right) = \frac{4}{\sin \frac{\pi}{9}} \sin \frac{\pi}{9} = 4.\end{aligned}$$

Exercice 9

Par définition de la fonction tan et avec les formules d'Euler,

$$\tan(nx) = \frac{\sin(nx)}{\cos(nx)} = -i \frac{e^{inx} - e^{-inx}}{e^{inx} + e^{-inx}}.$$

De plus, la formule du binôme de Newton donne

$$\begin{aligned}e^{inx} &= (\cos(x) + i \sin(x))^n \\ &= \sum_{k=0}^{n} \binom{n}{k} \cos(x)^{n-k} i^k \sin(x)^k \\ &= \cos(x)^n \sum_{k=0}^{n} \binom{n}{k} i^k \tan(x)^k.\end{aligned}$$

De même,

$$e^{-inx} = \cos(x)^n \sum_{k=0}^{n} \binom{n}{k} (-i)^k \tan(x)^k.$$

En combinant ces résultats, on obtient

$$\tan(nx) = -i \frac{\sum_{k=0}^{n} \binom{n}{k} \left(i^k - (-i)^k \right) \tan(x)^k}{\sum_{k=0}^{n} \binom{n}{k} (i^k + (-i)^k) \tan(x)^k}$$

Ces sommes se simplifient alors en

$$\tan(nx) = -i \frac{\sum_{\substack{k=0 \\ k=1[2]}}^{n} \binom{n}{k} 2 i^k \tan(x)^k}{\sum_{\substack{k=0 \\ k=0[2]}}^{n} \binom{n}{k} 2 i^k \tan(x)^k} = \frac{\sum_{\substack{k=0 \\ k=1[2]}}^{n} \binom{n}{k} (-1)^{\frac{k-1}{2}} \tan(x)^k}{\sum_{\substack{k=0 \\ k=0[2]}}^{n} \binom{n}{k} (-1)^{\frac{k}{2}} \tan(x)^k}.$$

Exercice 10

Remarquons tout d'abord que le membre de droite est bien défini pour tout $x \in \mathbb{R}$ puisque $x \mapsto x\,\mathrm{th}(x)$ est à valeurs positives.
L'équation se réécrit $1 + x\,\mathrm{th}\,x = e^x$. Étudions $f : x \mapsto 1 + x\,\mathrm{th}(x)$. Comme th est à valeurs dans $[-1, 1]$, on obtient $1 \leq f(x) \leq 1 + |x|$.

- Pour $x = 0$, l'équation est satisfaite.
- Pour $x \in \mathbb{R}_+^*$, il n'y a pas de solution puisque $e^x > 1 + x \geq f(x)$.
- Pour $x \in \mathbb{R}_-^*$, il n'y a pas de solution puisque $e^x < 1 \leq f(x)$.

La seule solution est $x = 0$.

Exercice 11

Si l'équation est satisfaite, alors $a = A\mathrm{ch}(\varphi)$ (en évaluant en $x = 0$) et $b = A\mathrm{sh}(\varphi)$ (en dérivant puis en évaluant en $x = 0$). Par conséquent, $|a| > |b|$ d'après l'inégalité correspondante entre les fonctions ch et sh.
Réciproquement, supposons $|a| > |b|$. Définissons $A \in \mathbb{R}^*$ par $A^2 = a^2 - b^2$ et $aA > 0$ (donc A et a sont de même signe). Comme $A^2 \leq a^2$, $\frac{a}{A} \geq 1$ donc il existe $\varphi \in \mathbb{R}$ tel que $\frac{a}{A} = \mathrm{ch}(\varphi)$. Ensuite $a = A\mathrm{ch}(\varphi)$ puis

$$b^2 = a^2 - A^2 = A^2(1 - \mathrm{ch}^2(\varphi)) = A^2\mathrm{sh}^2(\varphi).$$

Quitte à changer φ en $-\varphi$ (ce qui ne perturbe pas la relation $a = A\mathrm{ch}(\varphi)$), on a $b = A\mathrm{sh}(\varphi)$. Ainsi, pour tout $x \in \mathbb{R}$,

$$\begin{aligned} a\mathrm{ch}(x) + b\mathrm{sh}(x) &= A\mathrm{ch}(\varphi)\mathrm{ch}(x) + A\mathrm{sh}(\varphi)\mathrm{sh}(x) \\ &= A\mathrm{ch}(x + \varphi). \end{aligned}$$

Remarque

D'après le cours, pour factoriser un terme de la forme $a\cos(x) + b\sin(x)$, on met en facteur $\sqrt{a^2 + b^2}$. L'analogue en trigonométrie hyperbolique est donc : pour factoriser un terme de la forme $a\mathrm{ch}(x) + b\mathrm{sh}(x)$, on met en facteur $\sqrt{a^2 - b^2}$.

Exercice 12

Posons $X = e^x$ et $Y = e^y$. Le système se traduit alors par

$$\left\{ \begin{array}{lcll} X + \frac{1}{X} & + & Y + \frac{1}{Y} & = 2a \\ X - \frac{1}{X} & + & Y - \frac{1}{Y} & = 2b \end{array} \right.$$

Avec l'opération élémentaire $L_1 \leftarrow L_1 + L_2$, ce système est équivalent à

$$\left\{ \begin{array}{lcll} X & + & Y & = a + b \\ X - \frac{1}{X} & + & Y - \frac{1}{Y} & = 2b \end{array} \right.$$

En substituant la première équation dans la seconde, on obtient

$$\left\{ \begin{array}{lcll} X & + & Y & = a + b \\ \frac{1}{X} & + & \frac{1}{Y} & = a - b \end{array} \right.$$

Ainsi, le système initial est équivalent à la résolution dans $(\mathbb{R}_+^*)^2$ de $X + Y = a + b$ et $XY = \frac{a+b}{a-b}$, donc à l'existence de racines dans $\mathbb{R}_+^*$ du polynôme

$$X^2 - (a + b)X + \frac{a + b}{a - b} = 0.$$

Le discriminant est $\Delta = \frac{a+b}{a-b}(a^2 - b^2 - 4)$. Les racines sont réelles si, et seulement si, $a^2 \geq b^2 + 4$ et, dans ce cas, les racines sont strictement positives (car $a + b > \sqrt{\Delta}$).
La condition recherchée est $a^2 \geq b^2 + 4$.

Exercice 13

▹ Pour tout $x \neq 0$, $\operatorname{th} x \neq 0$ et $\operatorname{th} \frac{x}{2} \neq 0$. Ainsi,

$$\frac{1}{\operatorname{th} \frac{x}{2}} - \frac{1}{\operatorname{th} x} = \frac{\operatorname{ch} \frac{x}{2} \operatorname{sh} x - \operatorname{sh} \frac{x}{2} \operatorname{ch} x}{\operatorname{sh} \frac{x}{2} \operatorname{sh} x} = \frac{\operatorname{sh}\left(x - \frac{x}{2}\right)}{\operatorname{sh} \frac{x}{2} \operatorname{sh} x} = \frac{1}{\operatorname{sh}(x)}.$$

▹ Pour tout $x \neq 0$, la somme est télescopique

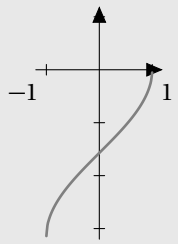
$$\begin{aligned}\sum_{k=0}^{n} \frac{1}{\operatorname{sh}(2^k x)} &= \sum_{k=0}^{n} \frac{1}{\operatorname{th} \frac{2^k x}{2}} - \sum_{k=0}^{n} \frac{1}{\operatorname{th} 2^k x} \\ &= \frac{1}{\operatorname{th} \frac{x}{2}} - \frac{1}{\operatorname{th} 2^n x}.\end{aligned}$$

Or, $\lim\limits_{x \to +\infty} \operatorname{th} x = 1$ donc $\lim\limits_{n} \sum\limits_{k=0}^{n} \frac{1}{\operatorname{sh}(2^k x)} = \frac{1}{\operatorname{th} \frac{x}{2}} - 1$.

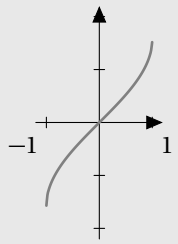
Exercice 14

▹ Les tracés des courbes des fonctions $x \mapsto \arccos(x)$ et $x \mapsto \arcsin(x)$ sont donnés dans le cours. Par ailleurs, la fonction $x \mapsto \frac{\pi}{2} - \arccos(x)$ est $x \mapsto \arcsin x$ donc cette courbe est déjà connue.

▹ La courbe de la fonction $x \mapsto -\arccos(x)$ est symétrique par rapport à l'axe des abscisses de la courbe de $x \mapsto \arccos(x)$ soit



▹ La courbe de la fonction $x \mapsto \arcsin(-x) = -\arcsin(x)$ est symétrique par rapport à l'axe des abscisses de la courbe de $x \mapsto \arcsin(x)$ soit



Exercice 15

Notons $\alpha = \arcsin \frac{3}{5} + \arcsin \frac{5}{13}$ et calculons $\sin \alpha$ avec la formule d'addition.

$$\sin \alpha = \frac{3}{5} \cos \arcsin \frac{5}{13} + \frac{5}{13} \cos \arcsin \frac{3}{5}.$$

Par ailleurs, pour tout $t \in [-1,1]$, $\cos \arcsin t = \sqrt{1-t^2}$ (car $\cos^2 + \sin^2 = 1$ et cos est positif sur $\arcsin[-1,1] = [-\frac{\pi}{2}, \frac{\pi}{2}]$) donc

$$\sin\alpha = \frac{3}{5}\sqrt{1-\frac{5^2}{13^2}} + \frac{5}{13}\sqrt{1-\frac{3^2}{5^2}} = \frac{56}{65}.$$

Remarquons ensuite que $0 \leq \frac{3}{5} \leq \frac{1}{\sqrt{2}}$ et $0 \leq \frac{5}{13} \leq \frac{1}{\sqrt{2}}$ (simple vérification en élevant au carré) donc, par croissance de la fonction arcsin,

$$0 \leq \alpha \leq 2\arcsin\frac{1}{\sqrt{2}} = \frac{\pi}{2}.$$

Par conséquent, $\alpha = \arcsin(\sin\alpha) = \arcsin\frac{56}{65}$.

Exercice 16

Remarquons tout d'abord que $f(x)$ est entre 0 et π car chacun des termes appartient à $[0, \frac{\pi}{2}]$. Par ailleurs, pour tout $x \in [0,1]$, la formule d'addition donne

$$\begin{aligned}\sin(f(x)) &= \sin(\arcsin x)\cos(\arcsin\sqrt{1-x^2}) + \cos(\arcsin x)\sin(\arcsin\sqrt{1-x^2})\\ &= x\sqrt{1-\sqrt{1-x^2}^2} + \sqrt{1-x^2}\sqrt{1-x^2} = x^2 + 1 - x^2 = 1.\end{aligned}$$

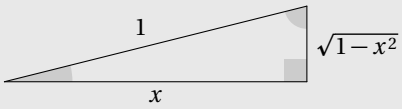
En conclusion, $f(x) = \frac{\pi}{2}$ pour tout $x \in [0,1]$.

Remarque

D'autres moyens de résolution étaient envisageables et tout aussi efficaces :

▹ On aurait également pu montrer que la fonction f est continue sur $[0,1]$ et dérivable sur $]0,1[$ à dérivée nulle ; ainsi, f est constante et $f(0) = \frac{\pi}{2}$.

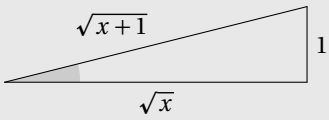
▹ On aurait enfin pu considérer la somme des angles dans un triangle rectangle dont les côtés ont pour longueur 1, x et $\sqrt{1-x^2}$ (d'après le théorème de Pythagore) :



La quantité $f(x)$ est la somme des deux angles contre l'hypoténuse donc vaut $\pi - \frac{\pi}{2} = \frac{\pi}{2}$.

Exercice 17

Considérons un triangle rectangle dont les côtés ont pour longueur 1, $\sqrt{x}$ et $\sqrt{x+1}$ (pour l'hypoténuse d'après le théorème de Pythagore) :



D'après les relations trigonométriques dans un triangle rectangle, l'angle indiqué sur cette figure admet pour tangente $\frac{1}{\sqrt{x}}$ et pour sinus $\frac{1}{\sqrt{x+1}}$ d'où l'égalité recherchée.

Exercice 18

1. D'après la formule d'addition de tan,

$$\tan\big(2\arctan(\sqrt{2})\big) = \frac{2\sqrt{2}}{1-\sqrt{2}^2} = -2\sqrt{2}.$$

Par ailleurs,

$$\tan\big(\pi - \arctan(2\sqrt{2})\big) = \tan\big(-\arctan(2\sqrt{2})\big) = -\tan\big(\arctan(2\sqrt{2})\big) = -2\sqrt{2}.$$

Or, par définition de arctan, les deux valeurs $2\arctan(\sqrt{2})$ et $\pi - \arctan(2\sqrt{2})$ appartiennent à $]0,\pi[$ donc sont égales.

2. D'après le lien entre $\arctan x$ et $\arctan \frac{1}{x}$, on a

$$\begin{aligned}\arctan\left(\tfrac{1}{2\sqrt{2}}\right) + 2\arctan\left(\tfrac{1}{\sqrt{2}}\right) &= \frac{\pi}{2} - \arctan(2\sqrt{2}) + 2\Big(\frac{\pi}{2} - \arctan(\sqrt{2})\Big)\\ &= \frac{3\pi}{2} - \big(\arctan(2\sqrt{2}) + 2\arctan(\sqrt{2})\big)\\ &= \frac{\pi}{2}.\end{aligned}$$

Exercice 19

Rappelons tout d'abord que si $x + iy = re^{i\theta}$ avec $x \in \mathbb{R}^*$, $y \in \mathbb{R}$, $r \in \mathbb{R}_+^*$, $\theta \in \mathbb{R}$, alors $\tan\theta = \frac{y}{x}$.

▹ Le résultat proposé par l'énoncé indique que la quantité

$$\tan\left(6\arctan\frac{1}{8} + 2\arctan\frac{1}{57} + \arctan\frac{1}{239}\right)$$

vaut

$$\frac{150837781250}{150837781250} = 1,$$

donc $6\arctan\frac{1}{8} + 2\arctan\frac{1}{57} + \arctan\frac{1}{239} = \frac{\pi}{4}[\pi]$.

Pour obtenir une égalité à la place d'une congruence on va chercher à localiser le membre de gauche dans un intervalle de longueur inférieure à π.

▹ Soit $t = \tan\frac{\pi}{12}$. D'après la formule d'addition de la fonction tan,

$$\frac{2t}{1-t^2} = \tan\frac{\pi}{6} = \frac{1}{\sqrt{3}},$$

donc en résolvant cette équation de degré 2, $t = 2 - \sqrt{3} = \frac{1}{2+\sqrt{3}} > \frac{1}{8}$, c'est-à-dire $\arctan\frac{1}{8} < \frac{\pi}{12}$.

Immédiatement, avec la croissance de la fonction arctan,

$$\begin{aligned}0 \le 6\arctan\frac{1}{8} + 2\arctan\frac{1}{57} + \arctan\frac{1}{239} &\le 9\arctan\frac{1}{8}\\ &< 9\frac{\pi}{12} = \frac{3\pi}{4}.\end{aligned}$$

En conclusion,

$$6\arctan\frac{1}{8} + 2\arctan\frac{1}{57} + \arctan\frac{1}{239} = \frac{\pi}{4}.$$

Corrigés des exercices d'approfondissement du chapitre 4

Exercice A

1. Supposons que la fonction f est k-lipschitzienne. D'après la définition de partie positive, $f_+ = \frac{f+|f|}{2}$. Par conséquent, pour tous $x, y \in \mathbb{R}$,

$$\begin{aligned}|f_+(x)-f_+(y)| &\leq \frac{1}{2}|f(x)-f(y)+|f(x)|-|f(y)||\\ &\leq \frac{1}{2}|f(x)-f(y)|+\frac{1}{2}||f(x)|-|f(y)||\\ &\leq |f(x)-f(y)| \leq k|x-y|.\end{aligned}$$

Ainsi, f_+ est k-lipschitzienne. On procède de même pour f_-.

2. Supposons que les fonctions f_+ et f_- sont k-lipschitziennes. Alors, $f = f_+ - f_-$ est continue comme différence de deux fonctions continues. Considérons $x, y \in \mathbb{R}$.

- Si les valeurs $f(x)$ et $f(y)$ sont positives, alors

$$|f(x)-f(y)| = |f_+(x)-f_+(y)| \leq k|x-y|.$$

On procède de même si les valeurs de $f(x)$ et $f(y)$ sont négatives.

- Sinon, on peut supposer, sans perte de généralité, que $f(x) > 0$ et $f(y) < 0$. D'après le théorème des valeurs intermédiaires appliqué à f, il existe z entre x et y tel que $f(z) = 0$. Introduisons cette valeur dans notre calcul

$$\begin{aligned}|f(x)-f(y)| &= |f_+(x)+f_-(y)|\\ &= |f_+(x)-f_+(z)+f_-(y)-f_-(z)|\\ &\leq |f_+(x)-f_+(z)|+|f_-(y)-f_-(z)|\\ &\leq k|x-z|+k|y-z| \leq k|x-y|,\end{aligned}$$

où pour la dernière inégalité, nous avons utilisé que z est entre x et y.
En conclusion, la fonction f est k-lipschitzienne.

Exercice B

Remarquons tout d'abord que la fonction f est dérivable de dérivée $f' : x \mapsto \sum_{k=1}^{n} \cos(kx)$.

Calculons cette dérivée sur $]0, 2\pi[$ en repassant par les parties réelles :

$$\begin{aligned}f'(x) &= \sum_{k=1}^{n} \mathfrak{Re}\left(e^{ikx}\right)\\ &= \mathfrak{Re}\left(\sum_{k=1}^{n} e^{ikx}\right)\\ &= \mathfrak{Re}\left(\frac{e^{ix}-e^{i(n+1)x}}{1-e^{ix}}\right)\\ &= \frac{\sin\left(\frac{n}{2}x\right)}{\sin\left(\frac{1}{2}x\right)}\cos\left(\frac{n+1}{2}x\right).\end{aligned}$$

CORRIGÉS

Ainsi, le premier zéro de f' sur $\mathbb{R}_+^*$ est en $x_n = \frac{\pi}{n+1}$ et la dérivée est positive avant et y change de signe (car dans l'expression seul le cosinus change de signe).

L'extremum local recherché est donc $f(x_n) = \sum_{k=1}^{n} \frac{\sin(\frac{k\pi}{n+1})}{k}$.

Exercice C

Commençons par montrer l'indication. On sait que $\tan\frac{\pi}{6} = \frac{1}{\sqrt{3}}$ donc $t = \tan\frac{\pi}{12}$ vérifie

$$\frac{1}{\sqrt{3}} = \frac{2t}{1-t^2}.$$

Comme $t \geq 0$, on a, en résolvant cette équation de degré 2, $t = 2-\sqrt{3}$.

Revenons au problème de départ. Considérons des réels $\theta_1, \dots, \theta_{13} \in \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$ tels que, pour tout $k \in [\![1, 13]\!]$, $t_k = \tan\theta_k$. On cherche donc deux entiers distincts i, $j \in [\![1, 13]\!]$ tels que $0 \leq \tan(\theta_i - \theta_j) \leq \tan\frac{\pi}{12}$ soit encore, avec la croissance de la fonction tan sur $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$, tels que $0 \leq \theta_i - \theta_j \leq \frac{\pi}{12}$.

En recouvrant l'intervalle $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$ en 12 sous-intervalles de longueur $\frac{\pi}{12}$, on obtient que deux des 13 réels $\theta_1, \dots, \theta_{13}$ sont dans le même petit intervalle donc séparés par moins de $\frac{\pi}{12}$.

Exercice D

1. Procédons par récurrence.

- Pour $n = 1$ et tout $x \in \mathbb{R}$,

$$\begin{aligned}\cos f(x)\sin\left(f(x)+\frac{\pi}{2}\right) &= \cos^2 f(x) = \frac{1}{1+\tan^2 f(x)}\\ &= \frac{1}{1+x^2} = f'(x).\end{aligned}$$

- Soit $n \in \mathbb{N} \setminus \{0\}$ telle que la propriété soit vraie au rang n. Alors, pour tout $x \in \mathbb{R}$,

$$\begin{aligned}f^{(n+1)}(x) &= (f^{(n)})'(x)\\ &= (n-1)!\left(-nf'(x)\sin f(x)(\cos f(x))^{n-1}\sin\left(n\left(f(x)+\frac{\pi}{2}\right)\right)\right)\\ &+ n!f'(x)(\cos f(x))^n\cos\left(n\left(f(x)+\frac{\pi}{2}\right)\right)\\ &= n!f'(x)(\cos f(x))^{n-1}\cos\left((n+1)f(x)+n\frac{\pi}{2}\right)\\ &= n!(\cos f(x))^{n+1}\sin\left((n+1)\left(f(x)+\frac{\pi}{2}\right)\right).\end{aligned}$$

2. Comme f est à valeurs dans $]-\frac{\pi}{2}, \frac{\pi}{2}[$, $\cos\circ f$ ne s'annule jamais. Par conséquent, $f^{(n)}(x) = 0$ si, et seulement si, $\sin\left(n\left(f(x)+\frac{\pi}{2}\right)\right) = 0$, c'est-à-dire $f(x)+\frac{\pi}{2} \in \frac{\pi}{n}\mathbb{Z}$. Les zéros de $f^{(n)}$ sont donc les

$$\tan\left(-\frac{\pi}{2}+k\frac{\pi}{n}\right) = -\frac{1}{\tan\left(k\frac{\pi}{n}\right)}$$

avec $k \in [\![1, n-1]\!]$.

Corrigés des problèmes du chapitre 4

Problème 1

1. a. La fonction f est dérivable comme somme de fonctions dérivables. On trouve immédiatement, pour tout $x \in [0, \frac{\pi}{2}[$

$$f'(x) = 2\cos x + \frac{1}{\cos^2 x} - 3 = \frac{2\cos^3 x - 3\cos^2 x + 1}{\cos^2 x}.$$

b. La fonction $\varphi : x \mapsto 2x^3 - 3x^2 + 1$ est dérivable de dérivée $\varphi'(x) = 6(x^2 - x) \leq 0$. Ainsi, φ est décroissante avec $\varphi(1) = 0$. Par conséquent,

$$\forall x \in [0,1], \qquad 2x^3 - 3x^2 + 1 \geq 0.$$

c. D'après les deux questions précédentes, la fonction f est croissante sur $[0, \frac{\pi}{2}[$. Comme $f(0) = 0$, f est positive sur cet intervalle et donc

$$\forall x \in [0, \frac{\pi}{2}[, \qquad 2\sin x + \tan x \geq 3x.$$

d. L'inégalité n'est pas vraie pour $x \in]\frac{\pi}{2}, \frac{3\pi}{2}[$ car la limite de f en $-\frac{\pi}{2}$ par valeurs positives est $-\infty$: la fonction n'est donc pas positive sur cet intervalle.

2. La même méthode permet de vérifier cette inégalité.

Problème 2

1. Comme vu dans le cours, $\tan(x+y) = \frac{\tan x + \tan y}{1 - \tan x \tan y}$.
2. Notons $x = \arctan a + \arctan b$. Alors, avec la question précédente

$$\tan x = \frac{a+b}{1-ab}.$$

Or, dans les deux cas, $x \in]-\frac{\pi}{2}, \frac{\pi}{2}[$ car

a. $\arctan a$ et $\arctan b$ appartiennent à $[0, \frac{\pi}{4}[$.

b. $\arctan a \in [0, \frac{\pi}{2}[$ et $\arctan b \in]-\frac{\pi}{2}, 0]$.

Par conséquent, $x = \arctan \tan x = \arctan \frac{a+b}{1-ab}$.

3. a. Avec la question précédente,

$$\begin{aligned}\arctan(k+1) - \arctan(k) &= \arctan(k+1) + \arctan(-k) \\ &= \arctan \frac{k+1-k}{1+k(k+1)} = \arctan \frac{1}{k^2+k+1}.\end{aligned}$$

b. Pour tout $n \in \mathbb{N}$, on obtient par télescopage

$$\sum_{k=0}^{n} \arctan \frac{1}{k^2+k+1} = \arctan(n+1) - \arctan(0).$$

Par conséquent,

$$\sum_{k=0}^{n} \arctan \frac{1}{k^2+k+1} \longrightarrow \frac{\pi}{2}.$$

CORRIGÉS

4. a. Notons, pour tout $n \in \mathbb{N}$, $u_n = F_{n+1}^2 - F_n F_{n+2}$. Alors, $u_0 = 1$ et, pour tout $n \in \mathbb{N}$,

$$\begin{aligned} u_{n+1} &= F_{n+2}^2 - F_{n+1}(F_{n+2} + F_{n+1}) \\ &= F_{n+2}(F_{n+2} - F_{n+1}) - F_{n+1}^2 \\ &= F_{n+2}F_n - F_{n+1}^2 = -u_n. \end{aligned}$$

Le résultat suit immédiatement pour cette suite géométrique de raison -1.
b. On applique encore la formule d'addition : pour tout $n \in \mathbb{N} \setminus \{0\}$,

$$\begin{aligned} \arctan \frac{1}{F_{2n+1}} + \arctan \frac{1}{F_{2n+2}} &= \arctan \frac{\frac{1}{F_{2n+1}} + \frac{1}{F_{2n+2}}}{1 - \frac{1}{F_{2n+1}} \frac{1}{F_{2n+2}}} \\ &= \arctan \frac{F_{2n+1} + F_{2n+2}}{F_{2n+1}F_{2n+2} - 1} \\ &= \arctan \frac{1}{F_{2n}} . \frac{F_{2n}F_{2n+1} + F_{2n}F_{2n+2}}{F_{2n+1}F_{2n+2} - 1} \\ &= \arctan \frac{1}{F_{2n}} . \frac{F_{2n}F_{2n+1} + F_{2n+1}^2 - 1}{F_{2n+1}F_{2n+2} - 1} \\ &= \arctan \frac{1}{F_{2n}}. \end{aligned}$$

c. Notons, pour tout n, $S_n = \sum_{k=0}^{n} \arctan \frac{1}{F_{2k+1}}$. Pour tout $n \in \mathbb{N} \setminus \{0\}$,

$$\begin{aligned} S_n + \arctan \frac{1}{F_{2n+2}} &= S_{n-1} + \arctan \frac{1}{F_{2n+1}} + \arctan \frac{1}{F_{2n+2}} \\ &= S_{n-1} + \arctan \frac{1}{F_{2n}}. \end{aligned}$$

Par récurrence, $S_n = S_0 + \arctan \frac{1}{F_2} - \arctan \frac{1}{F_{2n+2}}$. Comme $F_n \to +\infty$ et $S_0 = \frac{\pi}{4}$, on obtient $S_n \to \frac{\pi}{2}$.

Problème 3

Partie A

1. L'application $x \mapsto \ell(x)$ est strictement croissante car, pour tous $y > x$,

$$\ell(y) - \ell(x) = \int_x^y \frac{1}{\sqrt{1-t^4}} \, \mathrm{d}t > 0.$$

2. Pour tout $x \in [0,1[$, $x^4 \leq x^2$ et donc

$$\frac{1}{1-x^4} \leq \frac{1}{1-x^2}.$$

En passant à la racine puis en intégrant de 0 à x, on obtient l'inégalité $\ell(x) \leq \arcsin x$ pour tout $x \in [0,1[$. Par conséquent, ℓ est majorée par $\frac{\pi}{2}$. Comme ℓ est croissante, ℓ admet un limite finie en 1.
3. Par construction, ℓ réalise une bijection entre $[0,1]$ et $[0,\sigma]$. Comme la fonction ℓ est impaire, elle réalise une bijection entre $[-1,1]$ et $[-\sigma,\sigma]$.

Partie B – Sinus lemniscatique

1. La 4σ-périodicité permet de réduire à quatre les cas à étudier.
- Pour $k = 0[4]$, $\mathrm{sl}(k\sigma) = \mathrm{sl}(0) = 0$ car $\ell(0) = 0$.
- Pour $k = 1[4]$, $\mathrm{sl}(k\sigma) = \mathrm{sl}(\sigma) = 1$ car $\ell(1) = \sigma$.

- Pour $k=2[4]$, $\mathrm{sl}(k\sigma)=\mathrm{sl}(2\sigma)=\mathrm{sl}(0)=0$ par la symétrie de sl.
- Pour $k=3[4]$, $\mathrm{sl}(k\sigma)=\mathrm{sl}(-\sigma)=-1$ car $\ell(-1)=-\ell(1)=-\sigma$.

2. ▹ L'application sl est clairement continue sur $\mathbb{R}$ et de classe $\mathscr{C}^2$ sur $\mathbb{R}\setminus\{(2k+1)\sigma,\ k\in\mathbb{Z}\}$. De plus, pour tout $x\in]-\sigma,\sigma[$,

$$\mathrm{sl}'(x)=\frac{1}{\ell'(\mathrm{sl}(x))}=\sqrt{1-\mathrm{sl}^4(x)}.$$

Par conséquent, $\mathrm{sl}'\to 0$ en σ. Comme sl est continue en σ et que sa dérivée admet la limite finie 0 en σ, sl est dérivable en σ et $\mathrm{sl}'(0)=0$. On procède de même en $-\sigma$.
▹ Pour tout $x\in]-\sigma,\sigma[$,

$$\mathrm{sl}''(x)=\frac{-2\mathrm{sl}^3(x)\mathrm{sl}'(x)}{\sqrt{1-\mathrm{sl}^4(x)}}=-2\mathrm{sl}^3(x).$$

De plus, le taux d'accroissement admet une limite finie :

$$\begin{aligned}\frac{\mathrm{sl}'(x)-1}{x-\sigma}&=\frac{\sqrt{1-\mathrm{sl}^4(x)}-1}{x-\sigma}\\&=\frac{\mathrm{sl}(x)}{x-\sigma}\cdot\frac{-\mathrm{sl}^3(x)}{\sqrt{1-\mathrm{sl}^4(x)}+1}\\&\longrightarrow \mathrm{sl}'(\sigma)\frac{-3}{1}=0\end{aligned}$$

On conclut ainsi pour la dérivée seconde en σ. On procède de même en $-\sigma$.

Partie C – Une équation différentielle

1. Ceci est une conséquence directe de l'expression trouvée à la question 2 de la partie précédente.
2. Dérivons la fonction $(y')^2+y^4$.

$$\begin{aligned}((y')^2+y^4)'&=2y'y''+4y'y^3\\&=2y'(y''+2y^3)=0.\end{aligned}$$

En conclusion, la fonction $(y')^2+y^4$ est constante.
3. L'application $\varphi : x\mapsto \ell(\gamma^{-\frac{1}{4}}y(x))$ est dérivable de dérivée

$$\varphi'(x)=\frac{\gamma^{-\frac{1}{4}}y'(x)}{\sqrt{1-\gamma^{-1}y(x)^4}}=\frac{\gamma^{-\frac{1}{4}}\sqrt{\gamma-y(x)^4}}{\sqrt{1-\gamma^{-1}y(x)^4}}=\gamma^{\frac{1}{4}}.$$

En intégrant, on obtient que $\varphi : x\mapsto \gamma^{\frac{1}{4}}x+C$ avec $C\in\mathbb{R}$.
4. En inversant ℓ, on obtient que les solutions de E dont la dérivée reste positive sont de la forme

$$x\mapsto \gamma^{\frac{1}{4}}\mathrm{sl}\left(\gamma^{\frac{1}{4}}x+C\right)$$

avec $C\in\mathbb{R}$.

CORRIGÉS

Partie D – Cosinus lemniscatique

1. Pour tout $x \in \mathbb{R}$,

$$\begin{aligned}\mathrm{cl}^2(x)+\mathrm{sl}^2(x)+\mathrm{cl}^2(x)\mathrm{sl}^2(x) &= (1+\mathrm{sl}^2(x))\mathrm{cl}^2(x)+\mathrm{sl}^2(x)\\ &= \frac{\mathrm{sl}'^2(x)}{1+\mathrm{sl}^2(x)}+\mathrm{sl}^2(x)\\ &= \frac{1-\mathrm{sl}^4(x)+\mathrm{sl}^2(x)+\mathrm{sl}^4(x)}{1+\mathrm{sl}^2(x)} = 1 \qquad \text{car } \mathrm{sl}'^2+\mathrm{sl}^4 = 1.\end{aligned}$$

2. ▹ Par définition, la fonction cl est dérivable et, pour tout $x \in \mathbb{R}$,

$$\begin{aligned}\mathrm{cl}'(x) &= \frac{\mathrm{sl}''(x)(1+\mathrm{sl}^2(x))-2\mathrm{sl}'^2(x)\mathrm{sl}(x)}{(1+\mathrm{sl}^2(x))^2}\\ &= \frac{-2\mathrm{sl}(x)}{1+\mathrm{sl}^2(x)}.\end{aligned}$$

▹ En dérivant à nouveau, on obtient

$$\begin{aligned}\mathrm{cl}''(x) &= -2\mathrm{sl}'(x)\frac{1-\mathrm{sl}^2(x)}{(1+\mathrm{sl}^2(x))^2}\\ &= -2\frac{\mathrm{sl}'^3(x)}{(1+\mathrm{sl}^2(x))^3} = -2\mathrm{cl}^3(x).\end{aligned}$$

Par conséquent, cl est solution de (E).

3. Remarquons que $(\mathrm{cl}')^2+(\mathrm{cl})^4$ doit être une fonction constante γ d'après la question D2. Or cette constante γ vaut ici $(\mathrm{cl}')^2(0)+(\mathrm{cl})^4(0) = 1$. Ainsi, la dérivée de cl ne s'annulant pas, on obtient qu'il existe $C \in \mathbb{R}$ tel que

$$\forall x \in \mathbb{R}, \qquad \mathrm{cl}(x) = \mathrm{sl}(x+C).$$

Par ailleurs, en évaluant en $x = 0$, on obtient $\mathrm{sl}(C) = 1$ donc $C = \sigma[4\sigma]$. Ainsi, pour tout $x \in \mathbb{R}$,

$$\mathrm{cl}(x) = \mathrm{sl}(x+\sigma) = \mathrm{sl}(\sigma - x).$$

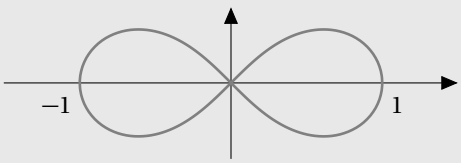
4. Pour tout $(x, y) \in \mathbb{R}^2$, posons

$$\varphi(x, y) = \frac{\mathrm{sl}(x)\mathrm{sl}'(y)+\mathrm{sl}'(x)\mathrm{sl}(y)}{1+\mathrm{sl}^2(x)\mathrm{sl}^2(y)}.$$

On vérifie rapidement que, à s fixé, $x \mapsto \varphi(x, s-x)$ est constante et égale à $\varphi(0, s-0) = \mathrm{sl}(s)$. Par conséquent,

$$\mathrm{sl}(x+y) = \frac{\mathrm{sl}(x)\mathrm{sl}'(y)+\mathrm{sl}'(x)\mathrm{sl}(y)}{1+\mathrm{sl}^2(x)\mathrm{sl}^2(y)}.$$

Ces fonctions trigonométriques sont dites lemniscatiques car la fonction ℓ correspondant à un calcul de longueur d'arc sur la courbe appelée Lemniscate de Bernoulli



Chapitre 5

Équations différentielles

1. Généralités — **2.** Primitives — **3.** Équations linéaires d'ordre 1 — **4.** Équations linéaires d'ordre 2

Objectifs et compétences du programme

- Connaître le vocabulaire des équations différentielles.
- Déterminer une primitive d'une fonction simple.
- Résoudre une équation différentielle linéaire d'ordre 1.
- Résoudre une équation différentielle linéaire d'ordre 2 à coefficients constants.

Josef Hoëné-Wronski, mathématicien et scientifique polonais (1776-1853)
Il demeure un personnage à part dans l'histoire des mathématiques. Initialement soldat dans l'armée polonaise, il entreprend des études universitaires en philosophie et en mathématiques.
Après une révélation mystique, il consacre l'essentiel de son énergie à établir une théorie messianique mais préserve un peu de temps pour écrire quelques mémoires mathématiques où l'on trouve notamment des déterminants, désormais appelés « wronskiens », qui permettent d'étudier certaines propriétés quantitatives des solutions d'équations différentielles linéaires. Toutefois, son tempérament et son style confus compliquent la réception de ses travaux. Il finira par se mettre à dos quasiment tous les mathématiciens de l'Académie avant de s'exiler à Londres (où il entreprend notamment de construire des véhicules tout terrain).

COURS

Dans ce chapitre, on apprend à résoudre des équations différentielles linéaires simples : les équations d'ordre 1 à coefficients quelconques et les équations d'ordre 2 à coefficients constants. Partant d'équation d'inspiration physique, on découvre peu à peu la structure de l'ensemble des solutions et les méthodes pour déterminer cet ensemble. On insiste particulièrement sur le résultat d'existence et d'unicité pour une solution vérifiant des conditions initiales données.

1. Généralités

1.1. Définitions générales

Une équation différentielle est une équation (définie sur un intervalle donné) dont l'inconnue est une fonction y, reliant les dérivées successives de l'inconnue prises avec le même argument. L'ordre de la plus importante dérivation apparaissant dans l'équation est appelé ordre de l'équation.

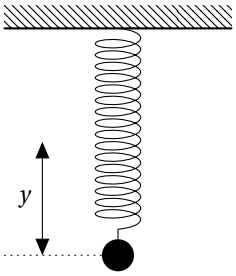
Une équation différentielle d'ordre k est sous forme résolue si elle peut s'écrire sous la forme

$$y^{(k)}(x) = f(x, y(x), \dots, y^{(k-1)}(x)).$$

Exemple

L'ordonnée y d'un point de masse m à l'extrémité d'un ressort vertical de coefficient k est solution de

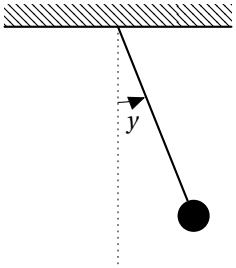
$$m y''(x) = -k y(x).$$



Exemple

L'angle d'un pendule simple de longueur l est solution de

$$y''(x) = -\frac{g}{l}\sin(y(x)) - k y'(x).$$



Le terme $-ky'$ est un terme de frottement.
Si on oublie le terme de frottement (c'est-à-dire $k=0$), on peut réécrire différemment l'équation ; multiplions par y' et intégrons

$$\frac{1}{2}(y'(x))^2 = \frac{g}{l}\cos(y(x)) + C, \qquad \text{avec } C \in \mathbb{R}.$$

On reconnaît l'équation de conservation de l'énergie (potentielle et cinétique).

Remarque

Toutes les équations de ces exemples sont d'ordres 1 ou 2. La variable x n'y apparaît pas : on dit de ces équations qu'elles sont autonomes.
La dernière équation n'est pas sous forme résolue.

1.2. Équations différentielles linéaires

Définition 5.1.

Une équation différentielle linéaire est une équation de la forme

$$\sum_{i=0}^{k} a_i(x) y^{(i)}(x) = b(x).$$

Les fonctions $a_0, \ldots, a_k$ sont appelées coefficients de l'équation.
La fonction b est appelée second membre de l'équation. Une équation dont le second membre est identiquement nul est homogène.

Remarque

Cette équation est appelée linéaire car l'application $\varphi : y \mapsto \sum_{i=0}^{k} a_i . y^{(i)}$ est linéaire, c'est-à-dire que pour toutes fonctions y et z et tous réels λ et μ,

$$\varphi(\lambda y + \mu z) = \lambda\varphi(y) + \mu\varphi(z).$$

Pour étudier une équation différentielle linéaire $\varphi(y) = b$, on commence toujours par étudier l'équation homogène associée $\varphi(y) = 0$.
En physique, on comprend le second membre comme le signal d'entrée et l'équation homogène décrit donc le régime libre du système.

Définition 5.2.

Un problème de Cauchy consiste à chercher les solutions d'une équation différentielle d'ordre k sur l'intervalle I étant données les valeurs de $y(a), \dots, y^{(k-1)}(a)$ pour un $a \in I$ donné, appelées conditions initiales.

Un problème de Cauchy en cinématique consiste à déterminer la trajectoire d'un mobile à partir de l'équation du mouvement et de conditions initiales : la position et la vitesse à un instant donné.

1.3. Structure de l'ensemble des solutions

Les trois propositions suivantes découlent d'un simple calcul exploitant la linéarité de la dérivation.

Proposition 5.3.

Soit f et g solutions d'une équation linéaire homogène (E), λ, μ des réels. Alors $\lambda f + \mu g$ est encore solution de (E).

En d'autres termes, l'ensemble des solutions d'une équation différentielle linéaire homogène est non vide et stable par combinaison linéaire (on verra plus loin que la structure algébrique correspondante est celle d'espace vectoriel).

Proposition 5.4.

Soit (E) une équation différentielle linéaire avec second membre et z une solution de (E).
Alors, toute solution de (E) s'écrit comme somme d'une solution de l'équation homogène associée à (E) et de z.

Proposition 5.5. Principe de superposition

Soit λ, $\mu \in \mathbb{R}$ et y_1, y_2 des solutions de deux équations linéaires de même équation homogène et de seconds membres f_1 et f_2. Alors, la fonction $\lambda y_1 + \mu y_2$ est solution de l'équation linéaire de même équation homogène avec le second membre $\lambda f_1 + \mu f_2$.

Cette dernière proposition permet de se ramener à deux problèmes plus simples lorsque le second membre est somme de deux fonctions élémentaires distinctes.

2. Primitives

Dans cette partie, on admet les propriétés de l'intégrale (linéarité, lien avec primitivisation) ; on développe les manipulations et calculs qui permettent de résoudre des équations de la forme $y' = b$ où b est une fonction continue (par morceaux).

Remarque
Il ne s'agit pas de fonder le calcul intégral mais seulement de compléter le cours de terminale pour disposer des outils nécessaires.

2.1. Généralités

Définition 5.6.

Une primitive d'une fonction f est une fonction F dérivable telle que $F' = f$.

Proposition 5.7.

Deux primitives d'une fonction définies sur un même intervalle ne diffèrent que d'une constante.

Démonstration

La différence de deux primitives est à dérivée nulle donc constante sur un intervalle. ■

Remarque

Le résultat est faux si l'on ne précise pas que l'on travaille sur un intervalle puisque la fonction non constante

$$f : x \mapsto \begin{cases} 0 & \text{si } x \in [0,1] \\ 1 & \text{si } x \in [1,2] \end{cases}$$

est à dérivée nulle.

Exemple

Dans les calculs de primitives simples, il est important de reconnaître les fonctions de la forme $u'u$, $\frac{u'}{u}$ et plus généralement de la forme $u'u^n$ avec $n \in \mathbb{Z}$. Par exemple, avec $u = \ln(x)$, on trouve, toujours à constante près, les primitives suivantes, pour $n \geq 2$,

$f(x)$	$F(x)$
$\frac{1}{x}\ln(x)$	$\frac{1}{2}(\ln(x))^2$
$\frac{1}{x\ln(x)}$	$\ln\lvert\ln(x)\rvert$
$\frac{1}{x\ln(x)^n}$	$\frac{1}{1-n}\cdot\frac{1}{(\ln(x))^{n-1}}$

Dressons un rapide tableau de primitives usuelles. La colonne de gauche contient l'expression de $f(x)$, celle de droite $F(x)$ où F est une primitive de f.

$f(x)$	$F(x)$
x^α	$\frac{1}{\alpha+1}x^{\alpha+1}$
$\frac{1}{x}$	$\ln\lvert x\rvert$
$\cos(x)$	$\sin(x)$
$\sin(x)$	$-\cos(x)$
$\exp(x)$	$\exp(x)$
$\frac{1}{1+x^2}$	$\arctan(x)$
$\frac{1}{\sqrt{1-x^2}}$	$\arcsin(x)$

Plus généralement, avec les fonctions réciproques, on obtient les primitives suivantes pour $a > 0$ (et à constante près)

$f(x)$	$F(x)$
$\frac{1}{a^2+x^2}$	$\frac{1}{a}\arctan\left(\frac{x}{a}\right)$
$\frac{1}{\sqrt{a^2-x^2}}$	$\arcsin\left(\frac{x}{a}\right)$

On admet (temporairement jusqu'à la construction rigoureuse de l'intégrale) la propriété suivante.

Proposition 5.8.

Soit f continue définie sur un intervalle I et $a \in I$. L'unique primitive de f nulle en a est la fonction

$$x \mapsto \int_a^x f(t)\,\mathrm{d}t\,.$$

Corollaire 5.9.

Soit f continue définie sur un intervalle $[a,b]$ et F une primitive de f sur $[a,b]$. Alors,

$$\int_a^b f(t)\,\mathrm{d}t = F(b) - F(a).$$

Cette quantité est encore notée $[F(t)]_a^b$.

Démonstration

Les fonctions F et $G : x \mapsto \int_a^x f(t)\,\mathrm{d}t$ ne diffèrent que d'une constante d'où

$$G(b) - F(b) = G(a) - F(a).$$

Ainsi,

$$\int_a^b f(t)\,\mathrm{d}t = G(b) - G(a) = F(b) - F(a).$$

■

Exemple

Soit $(a,b) \in \mathbb{R}^2 \setminus \{(0,0)\}$. Déterminons une primitive de $x \mapsto e^{ax}\cos(bx)$ en utilisant une fonction dont la variable apparaît comme borne d'une intégrale. Pour tout $x \in \mathbb{R}$,

$$\begin{aligned}
\int_0^x e^{at}\cos(bt)\,\mathrm{d}t &= \int_0^x \mathfrak{Re}(e^{(a+ib)t})\,\mathrm{d}t \\
&= \mathfrak{Re}\left(\int_0^x e^{(a+ib)t}\,\mathrm{d}t\right) \\
&= \mathfrak{Re}\left(\frac{e^{(a+ib)x}-1}{a+ib}\right) \\
&= \frac{a e^{ax}\cos(bx) - a + b e^{ax}\sin(bx)}{a^2+b^2}.
\end{aligned}$$

2.2. Intégration par parties

Proposition 5.10. Intégration par parties

Soit u, v deux fonctions de classe $\mathscr{C}^1$ sur $[a, b]$. Alors

$$\int_a^b u'(t)v(t)\,\mathrm{d}t = \big[u(t)v(t)\big]_a^b - \int_a^b u(t)v'(t)\,\mathrm{d}t\,.$$

Démonstration

Il suffit d'intégrer la relation $(u.v)'(t) = u'(t)v(t) + u(t)v'(t)$ entre a et b. ■

La technique de l'intégration par parties est très utile dès que l'on doit intégrer le produit d'un polynôme par une fonction exponentielle, une fonction trigonométrique ou une fonction logarithme puisque l'on peut faire baisser le degré du polynôme.

Exemple

Pour calculer $I_n(x) = \int_0^x t^n e^{-t}\,\mathrm{d}t$, on écrit par intégration par parties (avec $u'(t) = e^{-t}$ et $v(t) = t^n$ pour les notations ci-dessus)

$$I_n(x) = \big[-t^n e^{-t}\big]_0^x + nI_{n-1}(x) = -x^n e^{-x} + nI_{n-1}(x).$$

Avec la valeur de $I_0(x)$, on en déduit par récurrence un mode de calcul de $I_n(x)$.

Conseils méthodologiques

On peut utiliser l'intégration par parties pour faire disparaître un terme compliqué et le remplacer dans le calcul par sa dérivée.

Exemple

Les primitives de ln (sur $\mathbb{R}_+^*$) sont $x \mapsto x\ln x - x + C$ avec $C \in \mathbb{R}$. Il suffit pour cela d'appliquer la formule d'intégration par parties avec $u'(t) = 1$ et $v(t) = \ln(t)$.

Exemple

Les primitives de arctan (sur $\mathbb{R}$) sont $x \mapsto x\arctan x - \frac{1}{2}\ln(x^2+1) + C$ avec $C \in \mathbb{R}$. Il suffit pour cela d'appliquer la formule d'intégration par parties avec $u'(t) = 1$ et $v(t) = \arctan(t)$.

Exemple

Déterminons une primitive $x \mapsto \frac{1}{(x^2+1)^n}$. Calculons, pour tout $x \in \mathbb{R}$, $I_n(x) = \int_0^x \frac{1}{(1+t^2)^n}\,\mathrm{d}t$. On trouve immédiatement $I_1(x) = \arctan x$. De plus, par intégration par parties avec $u'(t) = 1$ et $v(t) = (1+t^2)^{-n}$,

$$\begin{aligned} I_n(x) &= \left[\frac{t}{(1+t^2)^n}\right]_0^x + 2n\int_0^x \frac{t^2}{(1+t^2)^{n+1}}\,\mathrm{d}t \\ &= \frac{x}{(1+x^2)^n} + 2n(I_n(x) - I_{n+1}(x)). \end{aligned}$$

Par conséquent, $I_{n+1}(x) = \frac{2n-1}{2n} I_n(x) + \frac{1}{2n} \cdot \frac{x}{(1+x^2)^n}$. À partir de cette relation de récurrence et de l'expression de $I_1(x)$, on peut calculer $I_n(x)$.

Exemple Intégrales de Wallis

Pour tout $n \in \mathbb{N}$, on pose $I_n = \int_0^{\pi/2} \cos^n t \, \mathrm{d}t$. Calculons I_n à l'aide d'un intégration par parties. Pour tout entier n,

$$\begin{aligned} I_n - I_{n+2} &= \int_0^{\pi/2} \cos^n t \sin^2 t \, \mathrm{d}t \\ &= \left[\frac{-\cos^{n+1} t}{n+1} \sin t \right]_0^{\frac{\pi}{2}} - \int_0^{\pi/2} \frac{-\cos^{n+1} t}{n+1} \cos t \, \mathrm{d}t = \frac{1}{n+1} I_{n+2}. \end{aligned}$$

D'où, $I_{n+2} = \frac{n+1}{n+2} I_n$ pour tout $n \in \mathbb{N}$. Comme on a $I_0 = \frac{\pi}{2}$ et $I_1 = 1$, pour tout $p \in \mathbb{N}$,

$$\begin{aligned} I_{2p} &= \frac{2p-1}{2p} \cdots \frac{1}{2} I_0 = \frac{(2p)!}{(2^p p!)^2} \cdot \frac{\pi}{2}; \\ I_{2p+1} &= \frac{2p}{2p+1} \cdots \frac{2}{3} I_1 = \frac{(2^p p!)^2}{(2p+1)!}. \end{aligned}$$

2.3. Changement de variables

Proposition 5.11.

Soit f une fonction continue sur un intervalle I et φ une fonction de classe $\mathscr{C}^1$ de $[a, b]$ dans I. Alors,

$$\int_a^b f(\varphi(x))\varphi'(x) \mathrm{d}x = \int_{\varphi(a)}^{\varphi(b)} f(y) \mathrm{d}y \, .$$

Démonstration

Les deux fonctions $b \mapsto \int_a^b f(\varphi(x))\varphi'(x) \mathrm{d}x$ et $b \mapsto \int_{\varphi(a)}^{\varphi(b)} f(y) \mathrm{d}y$ sont des primitives de $x \mapsto f(\varphi(x))\varphi'(x)$ nulles en a donc sont égales. ■

Exemple

Calculons, pour tout $a > 0$, $\int_0^x \frac{1}{a^2+t^2} \mathrm{d}t$.

$$\begin{aligned} \int_0^x \frac{1}{a^2+t^2} \mathrm{d}t &= \frac{1}{a^2} \int_0^x \frac{1}{1+\left(\frac{t}{a}\right)^2} \mathrm{d}t \\ &= \frac{1}{a^2} \int_0^{\frac{x}{a}} \frac{1}{1+u^2} a \, \mathrm{d}u = \frac{1}{a} \arctan \frac{x}{a}. \end{aligned}$$

Ainsi, une primitive de $x \mapsto \frac{1}{a^2+x^2}$ est $x \mapsto \frac{1}{a} \arctan \frac{x}{a}$. On montre de même qu'une primitive de $x \mapsto \frac{1}{\sqrt{a^2-x^2}}$ est $x \mapsto \arcsin \frac{x}{a}$.

Exemple

En posant $y = x^2$ avec x entre 0 et 1, on obtient

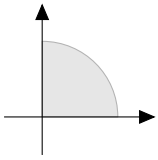
$$\int_0^1 x\tan(x^2)\,\mathrm{d}x = \frac{1}{2}\int_0^1 \tan y\,\mathrm{d}y = \frac{1}{2}[-\ln|\cos(y)|]_0^1 = -\frac{1}{2}\ln(\cos 1).$$

Exemple

En posant $y = \sin x$ avec x entre $\frac{\pi}{2}$ et 0, on obtient

$$\int_0^1 \sqrt{1-y^2}\,\mathrm{d}y = \int_0^{\frac{\pi}{2}} \cos^2(x)\,\mathrm{d}x = \int_0^{\frac{\pi}{2}} \frac{\cos(2x)+1}{2}\,\mathrm{d}x = \frac{\pi}{4}.$$

La courbe représentative de la fonction à intégrer étant un quart de cercle, on retrouve bien évidemment l'aire d'un quart du disque unité.



Proposition 5.12.

L'intégrale d'une fonction impaire sur un segment symétrique par rapport à 0 est nulle.

Démonstration

Soit f impaire. En posant $y = -x$,

$$\int_{-a}^{a} f(x)\,\mathrm{d}x = \int_{a}^{-a} f(-y)(-\mathrm{d}y) = -\int_{-a}^{a} f(y)\,\mathrm{d}y\,.$$

donc $\int_{-a}^{a} f(x)\,\mathrm{d}x = 0$. ■

3. Équations linéaires d'ordre 1

Remarquons que nous avons déjà résolu des équations différentielles de la forme $y'(x) = b(x)$. En effet, résoudre cette équation correspond à déterminer les primitives de la fonction b. C'était l'objet de la partie précédente.

3.1. $y'(x) = a\,y(x)$

Cette équation est la plus simple à résoudre car on reconnaît la définition de l'exponentielle. Plus précisément, cette équation est équivalente à $e^{-ax}(y'(x) - a\,y(x)) = 0$ donc $x \mapsto e^{-ax}\,y(x)$ est constante.

Proposition 5.13.

Les solutions de l'équation différentielle linéaire $y'(x) = a\,y(x)$ sont les fonctions $x \mapsto C\,e^{a\,x}$ avec $C \in \mathbb{R}$.

Remarque

Si l'on cherche les solutions à valeurs complexes d'une telle équation, on obtient le même résultat mais avec une constante $C \in \mathbb{C}$.

3.2. $y'(x) = a(x)y(x)$

On procède comme pour l'équation précédente en multipliant par $e^{-A(x)}$ où A désigne une primitive de a. On obtient que $x \mapsto e^{-A(x)}y(x)$ est constante.

Proposition 5.14.

Les solutions de l'équation différentielle linéaire $y'(x) = a(x)y(x)$ sur l'intervalle I contenant x_0 sont les fonctions $x \mapsto C\,e^{\int_{x_0}^{x} a(u)\,d\,u}$ avec $C \in \mathbb{R}$.

Conseils méthodologiques

Dans de nombreux exercices, on cherche à faire apparaître un terme de la forme $(e^{\text{truc}}\,y)'$ puis à intégrer.

Exemple

L'équation $y' = \frac{\alpha}{x}\,y$ est définie sur $\mathbb{R}_+^*$ et sur $\mathbb{R}_-^*$. Sur chacun de ces intervalles, les solutions sont les fonctions $x \mapsto C\,e^{\alpha \ln|x|} = C|x|^{\alpha}$ avec C une constante réelle qui dépend de l'intervalle sur lequel on se place pour résoudre l'équation.

Une conséquence de cette résolution explicite est le théorème d'existence et d'unicité suivant.

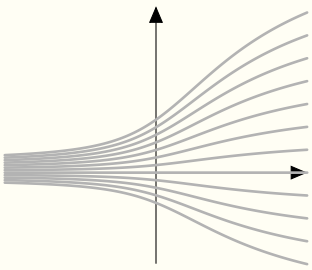
Proposition 5.15.

Pour tout $(x_0, y_0) \in \mathbb{R}^2$, il existe une unique solution y de l'équation $y'(x) = a(x)y(x)$ vérifiant la condition initiale $y(x_0) = y_0$ et cette solution est

$$x \mapsto y_0\,e^{\int_{x_0}^{x} a(t)\,\mathrm{d}t}.$$

Remarque

Interprétons géométriquement cette propriété : les courbes représentatives de deux solutions distinctes ne peuvent se couper (sinon il y aurait deux solutions correspondant aux conditions initiales indiquées par le point d'intersection) ; par ailleurs, par tout point du plan passe la courbe représentative d'une solution. Voici en guise d'illustration quelques solutions d'une équation linéaire d'ordre 1 :



En conclusion, les courbes représentatives des solutions distinctes d'une équation différentielle d'ordre 1 forment une partition du plan.

3.3. $y'(x) = a(x)y(x) + b(x)$

Remarquons que si z est une solution (particulière) de l'équation $y'(x) = a(x)y(x) + b(x)$, alors f est une solution de cette équation si, et seulement si, $f - z$ est solution de l'équation homogène $y'(x) = a(x)y(x)$ (que l'on sait déjà résoudre). Il suffit donc de chercher une solution particulière.

▹ Dans le cas où a est constante et b est le produit d'un polynôme par une fonction exponentielle de paramètre λ, on cherche une solution particulière sous la forme du produit d'un polynôme par une fonction exponentielle de paramètre λ (car l'ensemble des fonctions de cette forme est stable par dérivation et combinaison linéaire).

Exemple

Résolvons $y'(x) - y(x) = x^2 - 1$ avec la condition initiale $y(0) = \alpha$ en cherchant une solution polynomiale $ax^2 + bx + c$. En remplaçant dans l'équation on trouve les conditions

$$a = -1, \qquad -b + 2a = 0, \qquad b - c = -1.$$

Donc une solution particulière est $x \mapsto -x^2 - 2x - 1$ et la solution recherchée est

$$x \mapsto -x^2 - 2x - 1 + (\alpha + 1)e^x.$$

Précisons davantage le cas des équations à coefficients constants avec second membre produit d'une exponentielle et d'un polynôme.

Proposition 5.16.

L'équation $y'(x) = a\,y(x) + e^{\alpha x}P(x)$ avec $P \in \mathbb{R}[X]$ admet une solution de la forme $x \mapsto e^{\alpha x}Q(x)$ avec $Q \in \mathbb{R}[X]$ tel que

- $\deg Q = \deg P + 1$ si $\alpha = a$ (c'est-à-dire si l'exponentielle au second membre correspond aux solutions de l'équation homogène) ;
- $\deg Q = \deg P$ sinon.

Démonstration

Posons $z : x \mapsto e^{-ax}y(x)$. La fonction y vérifie $y'(x) = a\,y(x) + e^{\alpha x}P(x)$ si, et seulement si, $z'(x) = e^{(\alpha - a)x}P(x)$.

- Si $\alpha = a$, il suffit de prendre pour z une primitive de P.
- Sinon, on obtient une primitive de $x \mapsto e^{(\alpha - a)x}P(x)$ par intégrations par parties successives d'où l'on déduit le résultat.

■

▷ Dans le cas général, on utilise la méthode de variation de la constante : on cherche la solution sous la forme $x \mapsto C(x)e^{A(x)}$ où A est une primitive de a et C est une fonction dérivable que l'on cherche à déterminer (on a remplacé la constante dans l'expression de la solution générale de l'équation homogène par une fonction d'où le nom de la méthode). En reportant dans l'équation, on obtient la condition nécessaire et suffisante suivante

$$C'(x)e^{A(x)} = b(x).$$

Par conséquent, une solution particulière est donnée par

$$z(x) = e^{A(x)} \int_0^x b(v)e^{-A(v)}\,\mathrm{d}v\,.$$

Exemple

Résolvons $y'(x) - y(x) = e^{-2x}$ avec la condition initiale $y(0) = a$ par la méthode de variation de la constante. En posant $y(x) = C(x)e^x$, on obtient $C'(x) = e^{-3x}$ et donc $C(x) = -\frac{1}{3}e^{-3x}$. Une solution particulière est donnée par $z(x) = -\frac{1}{3}e^{-2x}$ et les solutions de l'équation homogène sont les fonctions $x \mapsto Ce^x$. La fonction recherchée est donc

$$x \mapsto \left(a + \frac{1}{3}\right)e^x - \frac{1}{3}e^{-2x}.$$

Comme on est sûr de trouver une solution particulière (au moins par la méthode de variation de la constante), on obtient

Proposition 5.17.

L'équation $y'(x) = a(x)y(x) + b(x)$ admet une unique solution satisfaisant la condition initiale $y(x_0) = y_0$ et cette solution est donnée par

$$y(x) = e^{\int_{x_0}^x a(u)\mathrm{d}u}\left(y_0 + \int_{x_0}^x b(v)e^{-\int_{x_0}^v a(u)\mathrm{d}u}\,\mathrm{d}v\right).$$

Remarque

L'unicité dans cette proposition traduit graphiquement que les courbes représentatives de deux solutions distinctes ne se coupent pas. L'existence assure que ces courbes représentatives recouvrent le plan.

4. Équations linéaires d'ordre 2

4.1. Équations homogènes à coefficients constants

Commençons par chercher les solutions exponentielles de l'équation linéaire à coefficients constants $ay'' + by' + cy = 0$. On vérifie alors que $x \mapsto e^{rx}$ est solution si, et seulement si, $ar^2 + br + c = 0$. Cela nous amène à la proposition suivante.

Proposition 5.18.

Considérons l'équation $ay'' + by' + cy = 0$ avec $a,\ b,\ c \in \mathbb{R}$ et $a \neq 0$. Soit Δ le discriminant de l'équation caractéristique $ar^2 + br + c = 0$ et r_1, r_2 les racines (éventuellement égales) de celle-ci.

- Si $\Delta > 0$, alors les solutions sont les fonctions $x \mapsto Ae^{r_1 x} + Be^{r_2 x}$ avec $A, B \in \mathbb{R}$.
- Si $\Delta = 0$, alors les solutions sont les fonctions $x \mapsto (Ax + B)e^{r_1 x}$ avec $A, B \in \mathbb{R}$.
- Si $\Delta < 0$, alors les solutions sont les fonctions

$$x \mapsto (A\cos(\mathfrak{Im}(r_1)x) + B\sin(\mathfrak{Im}(r_1)x))e^{\mathfrak{Re}(r_1)x},$$

avec $A, B \in \mathbb{R}$.

Remarque

Dans le dernier cas, on peut aussi écrire (en posant $C = \sqrt{A^2 + B^2}$) les solutions sous la forme

$$x \mapsto C\cos(\mathfrak{Im}(r_1)x + \varphi)e^{\mathfrak{Re}(r_1)x}$$

avec $C, \varphi \in \mathbb{R}$, φ étant alors appelé déphasage de la solution.

Démonstration

- Commençons par considérer le cas où l'équation caractéristique admet une racine réelle r_1 et posons $z : x \mapsto e^{-r_1 x}y(x)$. Un rapide calcul donne que y est une solution de $ay'' + by' + cy = 0$ si, et seulement si, z est une solution de $az'' + (2ar_1 + b)z' = 0$. Remarquons enfin que $2r_1 + \frac{b}{a} = r_1 - r_2$.
  - Si $\frac{2ar_1+b}{a} = 0$ (c'est-à-dire si $\Delta = 0$), alors $z'' = 0$ et donc z est un polynôme de degré au plus 1 et y a bien la forme recherchée.
  - Si $\frac{2ar_1+b}{a} \neq 0$ (c'est-à-dire si $\Delta \neq 0$), alors $z'' = (r_2 - r_1)z'$ donc (en intégrant l'équation en z' puis en réintégrant) on obtient $z(x) = Ae^{(r_2-r_1)x} + B$ avec $A, B \in \mathbb{R}$. On en déduit directement la forme de y.

- Si l'équation caractéristique n'admet pas de racine réelle, on procède de même avec une racine complexe r_1 et on montre de même (en considérant des fonctions à valeurs complexes) que y est de la forme $x \mapsto Ae^{r_1 x} + Be^{\overline{r}_1 x}$ avec $A, B \in \mathbb{C}$. Reste maintenant à exprimer que la solution est réelle, c'est-à-dire $y = \overline{y}$. Cette dernière équation s'écrit

$$\forall x, \qquad Ae^{r_1 x} + Be^{\overline{r}_1 x} = \overline{A}e^{\overline{r}_1 x} + \overline{B}e^{r_1 x}$$

donc $\overline{A} = B$. On obtient ainsi la forme recherchée $y(x) = 2\mathfrak{Re}(Ae^{r_1 x})$.

■

De la même manière, on obtient la résolution sur $\mathbb{C}$.

Corollaire 5.19.

Considérons l'équation $ay'' + by' + cy = 0$ avec $a, b, c \in \mathbb{C}$ et $a \neq 0$. Soit Δ le discriminant de l'équation caractéristique $ar^2 + br + c = 0$ et r_1, r_2 les racines (éventuellement égales) de celle-ci.

- Si $\Delta \neq 0$, alors les solutions sont les fonctions $x \mapsto Ae^{r_1 x} + Be^{r_2 x}$ avec $A, B \in \mathbb{C}$.
- Si $\Delta = 0$, alors les solutions sont les fonctions $x \mapsto (Ax + B)e^{r_1 x}$ avec $A, B \in \mathbb{C}$.

Exemple

Les solutions de $y'' + y' + y = 0$ sont les fonctions

$$x \mapsto e^{-\frac{1}{2}x}\left(A\cos\frac{\sqrt{3}}{2}x + B\sin\frac{\sqrt{3}}{2}x\right),$$

avec $A, B \in \mathbb{R}$. Par conséquent, elles tendent vers 0 en $+\infty$.
Ce résultat se généralise dès que les racines de l'équation caractéristique (ici, j et j^2) sont de partie réelle strictement négative.

On déduit de ce résultat le théorème d'existence et d'unicité suivant :

Proposition 5.20.

Étant donnés trois réels x_0, y_0, y_0', il existe une unique solution de l'équation $a\,y'' + b\,y' + c\,y = 0$ satisfaisant les conditions initiales $y(x_0) = y_0$ et $y'(x_0) = y_0'$.

Démonstration

Il suffit de considérer les trois cas.

▹ Si $b^2 - 4ac > 0$, alors une solution $x \mapsto Ae^{r_1 x} + Be^{r_2 x}$ avec $A, B \in \mathbb{R}$ et r_1, r_2 les racines de l'équation caractéristique vérifie les conditions initiales si, et seulement si,

$$\begin{cases} Ae^{r_1 x_0} & + & Be^{r_2 x_0} & = y_0 \\ Ar_1 e^{r_1 x_0} & + & Br_2 e^{r_2 x_0} & = y_0' \end{cases}$$

Ce système admet une unique solution

$$\begin{cases} A & = \frac{r_2 y_0 - y_0'}{r_2 - r_1} e^{-r_1 x_0} \\ B & = \frac{r_1 y_0 - y_0'}{r_1 - r_2} e^{-r_2 x_0} \end{cases}$$

▹ Si $b^2 - 4ac = 0$, alors une solution $x \mapsto (Ax + B)e^{r_1 x}$ avec $A, B \in \mathbb{R}$ et r_1 la racine de l'équation caractéristique vérifie les conditions initiales si, et seulement si,

$$\begin{cases} x_0 A & + & B & = y_0 e^{-r_1 x_0} \\ (r_1 x_0 + 1)A & + & r_1 B & = y_0' e^{-r_1 x_0} \end{cases}$$

Ce système admet une unique solution car il est de déterminant -1.

▹ Si $b^2 - 4ac < 0$, alors une solution $x \mapsto (A\cos(\beta x) + B\sin(\beta x))e^{\alpha x}$, avec $A, B \in \mathbb{R}$ et $\alpha \pm i\beta$ les racines de l'équation caractéristique, vérifie les conditions initiales si, et seulement si,

$$\begin{cases} \cos(\beta x_0)A & + & \sin(\beta x_0)B & = y_0 e^{-\alpha x_0} \\ (\alpha\cos(\beta x_0) - \beta\sin(\beta x_0))A & + & (\beta\cos(\beta x_0) + \alpha\sin(\beta x_0))B & = y_0' e^{-\alpha x_0} \end{cases}$$

Ce système admet une unique solution car il est de déterminant $\beta \neq 0$. ■

Remarque

Dans le cas des équations d'ordre 2, les courbes représentatives des solutions peuvent se croiser. Toutefois, cette proposition indique qu'en un point de concours, deux de ces courbes n'ont pas la même tangente.

4.2. Exemples d'équations avec second membre

Proposition 5.21.

L'équation $a\,y'' + b\,y' + c\,y = e^{\alpha x}P(x)$ avec $P \in \mathbb{R}[X]$ admet une solution de la forme $x \mapsto e^{\alpha x}Q(x)$ avec $Q \in \mathbb{R}[X]$ tel que

- $\deg Q = \deg P + 2$ si α est racine double de l'équation caractéristique ;
- $\deg Q = \deg P + 1$ si α est racine simple de l'équation caractéristique ;
- $\deg Q = \deg P$ sinon.

Démonstration

Soit r_1 une racine de l'équation caractéristique ; posons $z : x \mapsto e^{-r_1 x}\,y(x)$. Alors, y est une solution de $a\,y'' + b\,y' + c\,y = e^{\alpha x}P(x)$ si, et seulement si, z est une solution de $a\,z'' + (2a\,r_1 + b)z' = e^{(\alpha - r_1)x}P(x)$, c'est-à-dire $z'' - (r_2 - r_1)z' = \frac{1}{a}e^{(\alpha - r_1)x}P(x)$. On applique alors le résultat de la proposition 5.16.

- Si $\alpha - r_1 = r_2 - r_1$, alors z' est de la forme $x \mapsto e^{(\alpha - r_1)x}Q(x)$ avec $\deg Q = \deg P + 1$.
  - Si $\alpha - r_1 = 0$ (c'est-à-dire $\alpha = r_1 = r_2$), alors z est de la forme $x \mapsto e^{(\alpha - r_1)x}R(x)$ avec $\deg R = \deg Q + 1 = \deg P + 2$.
  - Sinon (c'est-à-dire $\alpha = r_2 \neq r_1$), z est de la forme $x \mapsto e^{(\alpha - r_1)x}R(x)$ avec $\deg R = \deg Q = \deg P + 1$.

- Si $\alpha - r_1 \neq r_2 - r_1$, alors z' est de la forme $x \mapsto e^{(\alpha - r_1)x}Q(x)$ avec $\deg Q = \deg P$.
  - Si $\alpha - r_1 = 0$ (c'est-à-dire $\alpha = r_1 \neq r_2$), alors z est de la forme $x \mapsto e^{(\alpha - r_1)x}R(x)$ avec $\deg R = \deg Q + 1 = \deg P + 1$.
  - Sinon (c'est-à-dire $\alpha \notin \{r_1, r_2\}$), z est de la forme $x \mapsto e^{(\alpha - r_1)x}R(x)$ avec $\deg R = \deg Q = \deg P$.

En se ramenant à y, on obtient le résultat annoncé. ■

Exemple

Cherchons une solution particulière à l'équation $y'' - 2y' + y = x\,e^x$. D'après la proposition précédente, on sait qu'il existe une solution $z : x \mapsto (a\,x^3 + b\,x^2 + c\,x + d)e^x$ avec a, b, c et $d \in \mathbb{R}$. Pour déterminer ces constantes a, b, c et d, calculons $z'' - 2z' + z$.

$$z'' - 2z' + z = (6a\,x + 2b)e^x.$$

Comme z est solution, on a $b = 0$ et $a = \frac{1}{6}$ (et aucune condition sur c et d). Ainsi, une solution particulière est $x \mapsto \frac{1}{6}x^3e^x$ et, en ajoutant les solutions de l'équation homogène, les solutions sont les fonctions

$$x \mapsto \frac{1}{6}x^3e^x + (Ax + B)e^x,$$

avec A et $B \in \mathbb{R}$.

Remarque

On remarque que dans l'exemple précédent, on n'obtient aucune condition sur les constantes c et d. Ceci est normal puisque la fonction $f : x \mapsto (c\,x + d)e^x$ est solution de l'équation homogène. En écrivant $z = f + g$, on obtient alors $z'' - 2z' + z = g'' - 2g' + g$. On aurait donc pu se dispenser des coefficients c et d dès le départ puisque l'on savait qu'ils n'interviendraient pas dans le calcul.

De manière plus générale, lorsque l'on cherche une solution sous la forme d'un produit exponentielle par polynôme, il est inutile de faire figurer les facteurs qui correspondent aux solutions de l'équation homogène.

4.3. Étude des oscillateurs linéaires amortis

Reprenons, en termes plus physiques, l'étude d'équations linéaires à coefficients constants fréquentes.

Définition 5.22.

Un oscillateur linéaire (ou harmonique) est un système gouverné par une équation différentielle de la forme

$$y'' + 2\lambda y' + \omega_0^2 y = e(x),$$

avec $\lambda \geq 0$ et $\omega_0 > 0$. Le second membre est appelé excitation du système ; si l'équation est homogène, le système est appelé oscillateur libre.
Le terme $2\lambda y'$ est le terme d'amortissement.

L'objectif est désormais de décrire quelques oscillateurs linéaires de conditions initiales

$$y(0) = a \quad \text{et} \quad y'(0) = 0.$$

1. Oscillateur libre ($e = 0$), non amorti ($\lambda = 0$)
   Dans ce cas simple, la solution au problème de Cauchy est $x \mapsto a\cos(\omega_0 x)$; a est l'amplitude, ω_0 est la pulsation qui est reliée à la période par la relation $T_0 = \frac{2\pi}{\omega_0}$.
   Remarquons que l'on peut obtenir une « intégrale première » ($(y')^2 + \omega_0^2 y^2$ est constante) en multipliant par y' et en intégrant. En termes physiques, cette dernière égalité traduit la conservation de l'énergie (le système est fermé car il n'y a pas d'excitation et conservatif car il n'y a pas d'amortissement).
2. Oscillateur libre ($e = 0$), amorti
   On discute selon le signe du discriminant de l'équation caractéristique $\Delta = 4(\lambda^2 - \omega_0^2)$. On introduit quelquefois la quantité $Q = \frac{\omega_0}{2\lambda}$ appelée facteur de qualité ; on remarque que Q est décroissant en λ donc qu'un fort taux d'amortissement implique un faible facteur de qualité.
   a. $\Delta < 0$ ou $Q > \frac{1}{2}$ (amortissement faible)
      La solution au problème de Cauchy est donnée par
      $$y(x) = a e^{-\lambda x}\left(\cos(x\sqrt{-\Delta}) + \frac{\lambda}{\sqrt{-\Delta}}\sin(x\sqrt{-\Delta})\right).$$
      On remarque en particulier que y tend vers 0 en $+\infty$ en oscillant : sans excitation et avec amortissement, le système revient à une position d'équilibre.
   b. $\Delta > 0$ ou $Q < \frac{1}{2}$
      La solution au problème de Cauchy est donnée par
      $$y(x) = \frac{a}{2}e^{-\lambda x}\left(\left[1 + \frac{\lambda}{\sqrt{\Delta}}\right]\exp(x\sqrt{\Delta}) + \left[1 - \frac{\lambda}{\sqrt{\Delta}}\right]\exp(-x\sqrt{\Delta})\right).$$
      On remarque en particulier que y tend vers 0 en $+\infty$ (sans oscillation cette fois) : sans excitation et avec amortissement, le système revient à une position d'équilibre.
   c. $\Delta = 0$ ou $Q = \frac{1}{2}$ (cas critique)
      La solution au problème de Cauchy est donnée par
      $$y(x) = a(\lambda x + 1)e^{-\lambda x}.$$
      On remarque encore que y tend vers 0 en $+\infty$.
3. Oscillateur excité, amorti

Dans ce cas de figure, on sait que la solution sera la somme d'une solution de l'équation homogène qui tend vers 0 en $+\infty$ (régime transitoire) et d'une solution particulière (régime stationnaire).
Dans le cas d'une excitation sinusoïdale $e(x) = \cos(\omega x)$ avec $\omega \neq 0$, on obtient une solution particulière (en la cherchant sous la forme d'un polynôme trigonométrique) définie par

$$y(x) = \frac{\omega_0^2 - \omega^2}{(\omega_0^2 - \omega^2)^2 + 4\lambda^2\omega^2}\cos(\omega x) + \frac{2\lambda\omega}{(\omega_0^2 - \omega^2)^2 + 4\lambda^2\omega^2}\sin(\omega x).$$

On remarque que cette solution n'existerait pas pour $\lambda = 0$ (système non amorti) et $\omega = \omega_0$: ce cas très particulier où le régime propre du système a la même pulsation que l'excitation correspond au phénomène de résonance.

FICHE

SYNTHÈSE

Dans ce chapitre, nous avons appris à résoudre certaines équations différentielles. La plus simple est $y' = f$ et sa résolution correspond à la détermination des primitives de f.

Primitive et intégration

Soit f continue définie sur un intervalle I et $a \in I$. L'unique primitive de f nulle en a est la fonction

$$x \mapsto \int_a^x f(t)\,\mathrm{d}t\,.$$

Les autres primitives de f se déduisent de celle-ci par l'ajout d'une constante.

Voici un petit tableau récapitulatif des primitives usuelles.

Primitives usuelles

La colonne de gauche contient l'expression de $f(x)$, celle de droite $F(x)$ où F est une primitive de f.

x^α avec $\alpha \neq -1$	$\frac{1}{\alpha+1}x^{\alpha+1}$
$\frac{1}{x}$	$\ln\lvert x\rvert$
$\ln(x)$	$x\ln(x)-x$
$\cos(x)$	$\sin(x)$
$\sin(x)$	$-\cos(x)$
$\exp(x)$	$\exp(x)$
$\frac{1}{1+x^2}$	$\arctan(x)$
$\frac{1}{\sqrt{1-x^2}}$	$\arcsin(x)$ $-\arccos(x)$

On peut obtenir de nombreuses primitives en utilisant les manipulations sur les intégrales : intégration par parties, changement de variables...

Équations différentielles linéaires d'ordre 1

Les solutions de l'équation différentielle linéaire homogène $y'(x) = a(x)y(x)$ sont les fonctions $x \mapsto Ce^{A(x)}$ avec $C \in \mathbb{R}$ et A une primitive de a.
Pour résoudre l'équation avec second membre, on ajoute une solution particulière que l'on cherche sous la forme $x \mapsto C(x)e^{A(x)}$ avec C dérivable (méthode de variation de la constante).
Chaque problème de Cauchy admet une unique solution.

Dans le cas particulier d'un second membre produit d'une exponentielle et d'un polynôme, on peut chercher la solution particulière de la même forme et on connaît *a priori* le degré du polynôme.

Équations différentielles linéaires d'ordre 2 à coefficients constants

Considérons l'équation $a\,y'' + b\,y' + c\,y = 0$ avec $a,\ b,\ c \in \mathbb{R}$ et $a \neq 0$. Soit Δ le discriminant de l'équation caractéristique $a\,r^2 + b\,r + c = 0$ et r_1, r_2 les racines (éventuellement égales) de celle-ci.

- Si $\Delta > 0$, alors les solutions sont les fonctions $x \longmapsto Ae^{r_1 x} + Be^{r_2 x}$ avec $A, B \in \mathbb{R}$.
- Si $\Delta = 0$, alors les solutions sont les fonctions $x \longmapsto (Ax + B)e^{r_1 x}$ avec $A, B \in \mathbb{R}$.
- Si $\Delta < 0$, alors les solutions sont les fonctions

$$x \longmapsto (A\cos(\mathfrak{Im}(r_1)x) + B\sin(\mathfrak{Im}(r_1)x))e^{\mathfrak{Re}(r_1)x},$$

  avec $A, B \in \mathbb{R}$.

Chaque problème de Cauchy admet une unique solution.

On sait comme précédemment résoudre des équations à coefficients constants et second membre produit d'une exponentielle et d'un polynôme.

Cas d'un second membre exponentielle-polynôme

Soit $a,\ b$ et $c \in \mathbb{R}$. L'équation $a\,y'' + b\,y' + c\,y = e^{\alpha x}P(x)$ avec $P \in \mathbb{R}[X]$ admet une solution de la forme $x \longmapsto e^{\alpha x}Q(x)$ avec $Q \in \mathbb{R}[X]$ tel que

- $\deg Q = \deg P + 2$ si α est racine double de l'équation caractéristique ;
- $\deg Q = \deg P + 1$ si α est racine simple de l'équation caractéristique ;
- $\deg Q = \deg P$ sinon.

EXERCICES

Vrai ou faux ?

		Vrai	Faux
a)	Les solutions de $y'(x)+a\,y(x)=0$ sont de la forme $x\mapsto C\,e^{a\,x}$ avec $C\in\mathbb{R}$.	☐	☐
b)	Les solutions de $y'(x)=a(x)y(x)$ sont de la forme $x\mapsto C\,e^{a(x)x}$ avec $C\in\mathbb{R}$.	☐	☐
c)	Les solutions de $y''(x)-2y'(x)+2y(x)=0$ sont de la forme $x\mapsto e^{-x}(A\cos x+B\sin x)$ avec $A,\,B\in\mathbb{R}$.	☐	☐
d)	Soit $a>0$. L'équation $y'(x)=a\,y(x)$ admet une unique solution bornée sur $\mathbb{R}_+$.	☐	☐
e)	L'équation $y'(x)=a\,y(x)+\sin(3x)$ admet au plus une solution périodique définie sur $\mathbb{R}$.	☐	☐
f)	Les solutions de $y'+a\,y=0$ sont deux à deux proportionnelles.	☐	☐
g)	Les solutions de $y''+a\,y'=0$ sont deux à deux proportionnelles.	☐	☐
h)	Les solutions de $y''+a\,y=0$ sont deux à deux proportionnelles.	☐	☐
i)	La fonction $x\mapsto e^{r\,x}$ est solution de $a\,y''+b\,y'+c\,y=0$ si, et seulement si, r est racine de l'équation caractéristique.	☐	☐
j)	La fonction $x\mapsto\arccos x$ est solution de $$y'+\frac{1}{\sqrt{1-x^2}}y=0.$$	☐	☐
k)	Les fonctions $x\mapsto\sin(x)$ et $x\mapsto\sin(2x)$ sont solutions d'une même équation linéaire d'ordre 2 à coefficients constants réels.	☐	☐
l)	L'équation $y''-2y'+2y=x.e^x$ admet une solution de la forme $x\mapsto P(x).e^x$ avec P un polynôme.	☐	☐
m)	L'équation $y''-3y'+2y=\ln(x)$ admet une solution de la forme $x\mapsto P(x).e^{\alpha x}$ avec P un polynôme et $\alpha\in\mathbb{R}$.	☐	☐
n)	Il y a une unique solution de $y'(x)=x\,y(x)$ vérifiant $y(0)=1$.	☐	☐

Exercices d'application du chapitre 5

●○○ **Exercice 1**

Calculer les primitives des fonctions suivantes

1. $x \mapsto e^x \cos x$,
2. $x \mapsto \sqrt{e^x - 1}$,
3. $x \mapsto x\sqrt[3]{1+x}$.

●○○ **Exercice 2**

Calculer les intégrales suivantes

1. $\int_0^{\pi/2} \sin t \cos^3 t \, \mathrm{d}t$,
2. $\int_0^{\pi/2} \cos^4 t \, \mathrm{d}t$,
3. $\int_0^{\pi/2} t^2 \cos t \, \mathrm{d}t$.

●○○ **Exercice 3**

Résoudre l'équation $y' - \arctan(x) y = 0$.

●○○ **Exercice 4**

Résoudre les équations homogènes d'ordre 1 suivantes

1. $y' - \cos(x) y = 0$,
2. $(1+x^2)y' + x y = 0$.

●○○ **Exercice 5**

Résoudre les équations d'ordre 1 suivantes

1. $y' - 3y = e^{6x}$,
2. $y' - 3y = 2e^{3x}$,
3. $y' - y = x + e^x$.

●○○ **Exercice 6**

Résoudre les équations homogènes d'ordre 2 suivantes

1. $y'' - 8y' + 15y = 0$,
2. $y'' - \sqrt{3}y' + y = 0$.

●○○ **Exercice 7**

Résoudre les équations d'ordre 2 suivantes

1. $y'' - 4y' + 4y = x e^{2x}$,
2. $y'' - 4y = x + e^{4x}$.

●○○ **Exercice 8**

Trouver les couples $(a, b) \in \mathbb{R}^2$ tels que l'équation $y'' + a y' + b y = 0$ admet une unique solution sur $\mathbb{R}$ de la forme $x \mapsto e^{\alpha x}$ avec $\alpha \in \mathbb{R}$.

●●○ **Exercice 9**

Résoudre l'équation $y''' - 2y'' + y' - 2y = 0$ en se ramenant à une équation d'ordre 2.

EXERCICES

●○○ **Exercice 10**

Considérons l'équation suivante avec a et b deux fonctions continues

$$y' + a(x)y + b(x)y^2 = 0.$$

1. Montrer que, si y est une solution de cette équation qui ne s'annule pas, alors $z = \frac{1}{y}$ est solution d'une équation linéaire.
2. Déterminer les solutions ne s'annulant pas sur $\mathbb{R}$ de $y' + y - xy^2 = 0$.

●●○ **Exercice 11**

Soit y une solution sur $\mathbb{R}_+^*$ de l'équation différentielle $x^2y'' - 2y = x$. Montrer que $z : t \mapsto y(e^t)$ est solution d'une équation différentielle puis en déduire l'expression de y.

●○○ **Exercice 12**

Trouver une équation différentielle linéaire d'ordre 1 dont les solutions sont exactement les fonctions $x \mapsto \frac{C+x}{1+x^2}$ pour $C \in \mathbb{R}$.

●●○ **Exercice 13**

Soit ω un réel strictement positif. Le mouvement d'une particule chargée dans un champ magnétique uniforme dirigé selon l'axe Oz est régi par le système différentiel (en la variable t) suivant

$$\begin{cases} x'' &= \omega y' \\ y'' &= -\omega x' \\ z'' &= 0 \end{cases}$$

En considérant la fonction $u = x' + iy'$, résoudre ce système différentiel.

●●○ **Exercice 14**

Déterminer les fonctions $f : \mathbb{R} \to \mathbb{R}$ continues telles que $x \mapsto f(x) - \int_0^x t f(t)\,\mathrm{d}t$ est constante.

●●● **Exercice 15**

Trouver les fonctions f continues vérifiant

$$\forall x \in \mathbb{R}, \qquad f(x) - \int_0^x f(x-t)\cos t\,\mathrm{d}t = 1.$$

Exercices d'approfondissement du chapitre 5

●○○ **Exercice A**

Existe-t-il une fonction bornée définie sur $\mathbb{R}_+^*$ solution de l'équation $y' + y = \ln x$?

●●○ **Exercice B**

Déterminer les couples $(a, b) \in \mathbb{R}^2$ tels que toutes les solutions de l'équation $y'' + ay' + by = 0$ sont bornées sur $\mathbb{R}_+$.

●●● **Exercice C**

Soit $a \in \mathbb{R}$. Trouver les fonctions f de classe $\mathscr{C}^2$ vérifiant $f'(x) = f(a-x)$ pour tout $x \in \mathbb{R}$.

●●● **Exercice D Lemme de Gronwall**

Soit $f:\mathbb{R}_+\to\mathbb{R}$, $g:\mathbb{R}_+\to\mathbb{R}_+$ continues et $A\geq 0$ tels que

$$\forall x\in\mathbb{R}_+^*,\qquad f(x)\leq A+\int_0^x f(t)g(t)\,\mathrm{d}t\,.$$

Montrer que

$$\forall x\in\mathbb{R}_+,\qquad f(x)\leq A\exp\left(\int_0^x g(t)\,\mathrm{d}t\right).$$

Problèmes du chapitre 5

●○○ **Problème 1**

Soit J un intervalle de $\mathbb{R}$ et $F:J\times\mathbb{R}\to\mathbb{R}$. Notons (E) l'équation différentielle $y'=F(x,y)$.

1. Donner une équation de la droite du plan de pente $F(x_0,y_0)$ et passant par le point (x_0,y_0).
2. Supposons, dans cette question, que F est de la forme $F:(x,y)\mapsto\alpha(x)y+\beta(x)$ avec α et β deux fonctions de classe $\mathscr{C}^1$ et fixons $x_0\in J$.
   a. Montrer que, si $\alpha(x_0)=0$, alors les tangentes au point d'abscisse x_0 aux courbes représentatives des solutions de (E) sont parallèles.
   b. Montrer que, si $\alpha(x_0)\neq 0$, alors les tangentes au point d'abscisse x_0 aux courbes représentatives des solutions de (E) sont concourantes.
3. Dans cette question, supposons que la fonction F vérifie
   - pour tout $(x_0,y_0)\in J\times\mathbb{R}$, il existe une solution de (E) satisfaisant la condition de Cauchy $y(x_0)=y_0$;
   - les tangentes au point d'abscisse x_0 aux courbes représentatives des solutions de (E) sont soit toutes parallèles, soit toutes concourantes.

   a. Soit $x_0\in J$; montrer que $\frac{F(x_0,y_0)-F(x_0,y_1)}{y_0-y_1}$ est une quantité qui ne dépend pas du couple (y_0,y_1) avec $y_0\neq y_1$. On note désormais $\alpha(x_0)$ cette quantité.
   b. En déduire que, si $\alpha(x_0)\neq 0$, $F(x_0,y_0)-\alpha(x_0)y_0$ ne dépend pas de y_0.
   c. Conclure que F est linéaire en la seconde variable.

●○○ **Problème 2**

Soit $f:[0,\frac{\pi}{2}]\to\mathbb{R}$ une fonction continue. Considérons l'équation différentielle (E) $y''+y=f$.

1. Dans cette question, $f:x\mapsto x+1$.
   a. Résoudre (E).
   b. Combien de solutions vérifient $y(0)=y(\frac{\pi}{2})=0$? $y(0)=y'(\frac{\pi}{2})=0$?

   Soit y une solution de (E).
2. Déterminer les fonctions de classe $\mathscr{C}^2$ A et B vérifiant les conditions suivantes
$$\begin{cases} A(x)\cos(x) & + & B(x)\sin(x) & = & y(x)\\ A'(x)\cos(x) & + & B'(x)\sin(x) & = & 0\end{cases}$$
3. En déduire qu'il existe $(\kappa_1,\kappa_2)\in\mathbb{R}^2$ tel que, pour tout $x\in[0,\frac{\pi}{2}]$,
$$y(x)=\int_0^x \sin(x-t)f(t)\,\mathrm{d}t+\kappa_1\cos(x)+\kappa_2\sin(x).$$
4. Montrer qu'il existe une unique solution de (E) satisfaisant $y(0)=y(\frac{\pi}{2})=0$, donnée par
$$\forall x\in\left[0,\tfrac{\pi}{2}\right],\qquad y(x)=-\cos(x)\int_0^x \sin(t)f(t)\,\mathrm{d}t-\sin(x)\int_x^{\pi/2}\cos(t)f(t)\,\mathrm{d}t\,.$$

PROBLÈMES

5. Déterminer une condition nécessaire et suffisante pour que (E) admette une solution satisfaisant $y(0) = y'(\frac{\pi}{2}) = 0$.

●●○ **Problème 3**

Partie A – Réduction de l'équation

Soit (E) une équation différentielle de la forme $a(x)y'' + b(x)y' + c(x)y = 0$ avec a une fonction ne s'annulant pas. Pour toute fonction u deux fois dérivable ne s'annulant pas, définissons z deux fois dérivable telle que $y = uz$.

1. Montrer que si y est solution de (E), alors z est encore solution d'une équation différentielle.
2. En déduire que si u est solution d'une équation (E'') linéaire d'ordre 1 (à préciser), alors z est solution d'une équation (E') de la forme

$$\alpha(x)z'' + \beta(x)z = 0.$$

Partie B – Entrelacement des zéros

Soit p, q deux fonctions, y_1 et y_2 des solutions non nulles des équations $y'' + p(x)y = 0$ et $y'' + q(x)y = 0$ respectivement.
On admet que les zéros de ces solutions sont isolés, c'est-à-dire que, pour chaque zéro d'une de ces solutions, il existe un intervalle ouvert ne contenant pas d'autre zéro.
Considérons la fonction, appelée wronskien, $W : x \mapsto y_1(x)y_2'(x) - y_1'(x)y_2(x)$.

1. Donner une expression simple de la dérivée de W.
2. Supposons, dans cette question, que, pour tout $x \in \mathbb{R}$, $p(x) \leq q(x)$.
   a. Soit $x_1 < x_2$ deux zéros consécutifs de y_2. On admet que $y_2'(x_1)y_2'(x_2) < 0$.
   Montrer que y_1 admet au moins un zéro dans l'intervalle $[x_1, x_2]$.
   b. En déduire que si p est minorée par une constante strictement positive ω^2, alors y_1 admet une infinité de zéros et que deux zéros consécutifs sont séparés d'au plus $\frac{\pi}{\omega}$.
   c. En déduire aussi que si q est majorée par 0, alors y_2 admet au plus un zéro.
3. Supposons que $p = q$ (c'est-à-dire que y_1 et y_2 sont solutions de la même équation différentielle d'ordre 2). Montrer que si y_1 et y_2 n'ont pas de zéro commun, alors y_1 admet exactement un zéro entre deux zéros consécutifs de y_2.

Partie C – Équation de Bessel

Soit n le nombre de zéros sur un intervalle de longueur π d'une solution sur $\mathbb{R}_+^*$ de l'équation

$$y'' + \frac{1}{x}y' + \left(1 - \frac{\lambda^2}{x^2}\right)y = 0.$$

Montrer que si $\lambda \geq \frac{1}{2}$, alors $n \leq 1$.

CORRIGÉS

Corrigés des Vrai/Faux

a) Faux, elles sont de la forme $x \mapsto C e^{-ax}$ avec $C \in \mathbb{R}$.

b) Faux, elles sont de la forme $x \mapsto C e^{A(x)}$ avec $C \in \mathbb{R}$ et A une primitive de a.

c) Faux, les racines de l'équation caractéristique sont $1 \pm i$ et non $-1 \pm i$.

d) Vrai, il n'y a que la solution nulle.

e) Vrai, la différence de deux solutions périodiques définies sur $\mathbb{R}$ est une solution de l'équation homogène et elle est bornée (comme différence de deux fonctions bornées) : c'est la fonction nulle.

f) Vrai.

g) Faux.

h) Faux.

i) Vrai par définition de l'équation caractéristique.

j) Faux, les solutions de cette équation sont $x \mapsto C e^{\arccos x}$ avec $C \in \mathbb{R}$.

k) Faux, si c'était le cas, il faudrait que i et $2i$ soient solutions d'une équation polynomiale réelle de degré 2 (dont les racines sont réelles ou conjuguées).

l) Vrai.

m) Faux, il suffit de vérifier que le membre de gauche sera de la forme $Q(x)e^{\alpha x}$ avec Q un polynôme.

n) Vrai.

Corrigés des exercices d'application du chapitre 5

Exercice 1

1. Remarquons que la fonction considérée est la partie réelle de la fonction $x \mapsto e^{(1+i)x}$. Une primitive est donc donnée par

$$\begin{aligned}\mathfrak{Re}\Big(\int_0^x e^{(1+i)t}\,\mathrm{d}t\Big) &= \mathfrak{Re}\Big(\frac{e^{(1+i)x}-1}{1+i}\Big)\\ &= \frac{1}{2}\mathfrak{Re}\big((1-i)(e^{(1+i)x}-1)\big)\\ &= \frac{1}{2}\big(e^x\cos x - 1 + e^x \sin x\big).\end{aligned}$$

Les primitives recherchées sont donc $x \mapsto \frac{1}{2}e^x(\cos x + \sin x) + C$ avec $C \in \mathbb{R}$.

2. Posons $u=\sqrt{e^t-1}$ c'est-à-dire $t=\ln(1+u^2)$ dans le calcul suivant

$$\begin{aligned}\int_0^x \sqrt{e^t-1}\,\mathrm{d}t &= \int_0^{\ln(1+x^2)} u\frac{2u}{1+u^2}\,\mathrm{d}u\\ &= 2\int_0^{\ln(1+x^2)} 1-\frac{1}{1+u^2}\,\mathrm{d}u\\ &= 2\ln(1+x^2)-2\arctan\big(\ln(1+x^2)\big).\end{aligned}$$

Les primitives recherchées sont $x\mapsto 2\ln(1+x^2)-2\arctan\big(\ln(1+x^2)\big)+C$ avec $C\in\mathbb{R}$.

3. Posons $u=\sqrt[3]{1+t}$ c'est-à-dire $t=u^3-1$ dans le calcul d'intégrale suivant

$$\begin{aligned}\int_{-1}^x t\sqrt[3]{1+t}\,\mathrm{d}t &= \int_0^{\sqrt[3]{1+x}} (u^3-1)u3u^2\,\mathrm{d}u\\ &= 3\int_0^{\sqrt[3]{1+x}} (u^6-u^3)\,\mathrm{d}u\\ &= \frac{3}{7}(1+x)^{\frac{7}{3}}-\frac{3}{4}(1+x)^{\frac{4}{3}}.\end{aligned}$$

Les primitives recherchées sont $x\mapsto \frac{3}{7}(1+x)^{\frac{7}{3}}-\frac{3}{4}(1+x)^{\frac{4}{3}}+C$ avec $C\in\mathbb{R}$.

Exercice 2

1. La fonction à intégrer est presque de la forme $4u'u^3$ avec $u=\cos t$ qui s'intègre en u^4. On se ramène à ce cas

$$\begin{aligned}\int_0^{\pi/2} \sin t\cos^3 t\,\mathrm{d}t &= -\frac{1}{4}\int_0^{\pi/2} -4\sin t\cos^3 t\,\mathrm{d}t\\ &= -\frac{1}{4}\big[\cos^4 t\big]_0^{\pi/2}\\ &= \frac{1}{4}\end{aligned}$$

2. Commençons par linéariser le $\cos^4$ avec la formule d'Euler :

$$\begin{aligned}\cos^4 x &= \Big(\frac{e^{ix}+e^{-ix}}{2}\Big)^4\\ &= \frac{e^{i4x}+4e^{i2x}+6+4e^{-i2x}+e^{i4x}}{16}\\ &= \frac{1}{8}\cos(4x)+\frac{1}{2}\cos(2x)+\frac{3}{8}.\end{aligned}$$

Intégrons désormais chacun de ces termes :

$$\begin{aligned}\int_0^{\pi/2} \cos^4 t\,\mathrm{d}t &= \int_0^{\pi/2} \frac{1}{8}\cos(4t)+\frac{1}{2}\cos(2t)+\frac{3}{8}\,\mathrm{d}t\\ &= \Big[\frac{1}{32}\sin(4t)+\frac{1}{4}\sin(2t)+\frac{3}{8}t\Big]_0^{\pi/2}\\ &= \frac{3\pi}{16}.\end{aligned}$$

3. Procédons par intégrations par parties à deux reprises afin d'« éliminer » le terme t^2 :

$$\begin{aligned}\int_0^{\pi/2} t^2\cos t\,\mathrm{d}t &= \left[t^2\sin t\right]_0^{\pi/2} - 2\int_0^{\pi/2} t\sin t\,\mathrm{d}t\\ &= \frac{\pi^2}{4} - 2\left[-t\cos t\right]_0^{\pi/2} - 2\int_0^{\pi/2}\cos t\,\mathrm{d}t\\ &= \frac{\pi^2}{4} - 2.\end{aligned}$$

Exercice 3

Cette équation est linéaire d'ordre 1, homogène. Les solutions sont de la forme $x \mapsto Ce^{A(x)}$ avec $C \in \mathbb{R}$ et A une primitive de arctan. Comme une primitive de arctan (obtenue par intégration par parties) est

$$x \mapsto x\arctan(x) - \frac{1}{2}\ln(1+x^2),$$

on en déduit que les solutions sont les fonctions

$$x \mapsto Ce^{x\arctan x - \frac{1}{2}\ln(1+x^2)} = \frac{C}{\sqrt{1+x^2}}e^{x\arctan x},$$

avec $C \in \mathbb{R}$.

Exercice 4

1. Une primitive de $-\cos$ est $-\sin$. Les solutions sont donc les fonctions $x \mapsto Ce^{\sin(x)}$ avec $C \in \mathbb{R}$.
2. Divisons l'équation par la quantité (non nulle) $1+x^2$ pour se ramener à une forme « normalisée ». Une primitive de $x \mapsto \frac{x}{1+x^2}$ est $x \mapsto \frac{1}{2}\ln(1+x^2)$ (la fonction est de la forme $\frac{u'}{u}$ avec $u(x) = 1+x^2$). Les solutions sont donc les fonctions

$$x \mapsto Ce^{-\frac{1}{2}\ln(1+x^2)} = \frac{C}{\sqrt{1+x^2}},$$

avec $C \in \mathbb{R}$.

Exercice 5

1. Les solutions de l'équation homogène sont les fonctions $x \mapsto Ce^{3x}$ avec $C \in \mathbb{R}$. Le second membre est produit d'un polynôme (constant) par une exponentielle avec le coefficient de l'exponentielle non solution de l'équation caractéristique ($6 \neq 3$). Il existe donc une solution particulière de la forme $x \mapsto Ae^{6x}$ avec $A \in \mathbb{R}$: en reportant on obtient $(6A-3A)e^{6x} = e^{6x}$ soit $A = \frac{1}{3}$. En conclusion, les solutions sont de la forme $x \mapsto \frac{1}{3}e^{6x} + Ce^{3x}$ avec $C \in \mathbb{R}$.
2. Les solutions de l'équation homogène sont les fonctions $x \mapsto Ce^{3x}$ avec $C \in \mathbb{R}$. Le second membre est produit d'un polynôme (constant) par une exponentielle avec le coefficient de l'exponentielle solution simple de l'équation caractéristique. Il existe donc une solution particulière de la forme $x \mapsto (Ax+B)e^{3x}$ avec A, $B \in \mathbb{R}$: en reportant on obtient $A = 2$ (et aucune condition sur B). En conclusion, les solutions sont de la forme $x \mapsto 2xe^{3x} + Ce^{3x}$ avec $C \in \mathbb{R}$.
3. Les solutions de l'équation homogène sont les fonctions $x \mapsto Ce^{x}$ avec $C \in \mathbb{R}$.
Pour le second membre x (polynôme de degré 1), une solution particulière est $x \mapsto -x-1$.
Pour le second membre e^x (exponentielle de coefficient racine de l'équation caractéristique), une solution particulière est $x \mapsto xe^x$.
D'après le principe de superposition, les solutions sont de la forme $x \mapsto xe^x - x - 1 + Ce^x$ avec $C \in \mathbb{R}$.

Exercice 6

1. L'équation caractéristique $r^2-8r+15=0$ admet deux racines distinctes : 3 et 5. Par conséquent, les solutions sont les fonctions $x \mapsto Ae^{3x}+Be^{5x}$ avec A et $B \in \mathbb{R}$.
2. L'équation caractéristique $r^2-\sqrt{3}r+1=0$ admet deux racines complexes conjuguées : $\frac{\sqrt{3}\pm i}{2}$. Par conséquent, les solutions sont les fonctions $x \mapsto e^{\frac{\sqrt{3}}{2}x}\left(A\cos(\frac{x}{2})+B\sin(\frac{x}{2})\right)$ avec A et $B \in \mathbb{R}$.

Exercice 7

1. L'équation caractéristique $r^2-4r+4=0$ admet une unique racine 2. Par conséquent, les solutions de l'équation homogène sont les fonctions $x \mapsto (Ax+B)e^{2x}$ avec A, $B \in \mathbb{R}$.
Le second membre est le produit d'un polynôme de degré 1 et d'une exponentielle dont le coefficient est racine double de l'équation caractéristique : il existe donc une solution de la forme $x \mapsto (ax^3+bx^2)e^{2x}$. En reportant, on obtient $a=\frac{1}{6}$ et $b=0$. On a ainsi trouvé une solution particulière et les solutions de l'équation complète sont

$$x \mapsto \frac{1}{6}x^3e^{2x}+(Ax+B)e^{2x},$$

avec A, $B \in \mathbb{R}$.
2. L'équation caractéristique $r^2-4=0$ admet deux racines distinctes ± 2. Par conséquent, les solutions de l'équation homogène sont les fonctions $x \mapsto Ae^{2x}+Be^{-2x}$ avec A, $B \in \mathbb{R}$.
Une solution de l'équation avec le second membre x est $x \mapsto -\frac{1}{4}x$.
Une solution de l'équation avec le second membre e^{4x} est $x \mapsto \frac{1}{12}e^{4x}$.
En conclusion, les solutions de l'équation complète sont les fonctions $x \mapsto -\frac{1}{4}x+\frac{1}{12}e^{4x}+Ae^{2x}+Be^{-2x}$, avec A, $B \in \mathbb{R}$.

Exercice 8

La fonction $x \mapsto e^{\alpha x}$ avec $\alpha \in \mathbb{R}$ est solution de l'équation $y''+ay'+by=0$ si, et seulement si, $\alpha^2+a\alpha+\beta=0$ c'est-à-dire α est racine de l'équation caractéristique.
Ainsi, il existe une unique solution de cette forme si, et seulement si, l'équation caractéristique admet une unique racine réelle, c'est-à-dire si son discriminant est nul : la condition recherchée sur les couples $(a,b) \in \mathbb{R}^2$ est donc $a^2=4b$.

Exercice 9

Considérons la fonction $z=y'-2y$. Alors, y vérifie l'équation de l'énoncé si, et seulement si, $z''+z=0$, c'est-à-dire si z est de la forme

$$x \mapsto A\cos x+B\sin x,$$

avec A et $B \in \mathbb{R}$. On est donc ramené à résoudre l'équation d'ordre 1 suivante

$$y'-2y=A\cos x+B\sin x.$$

Les solutions de l'équation homogène sont de la forme $x \mapsto Ce^{2x}$ avec $C \in \mathbb{R}$. Une fonction de la forme

$$x \mapsto \alpha\cos x+\beta\sin x$$

est solution de l'équation avec second membre si, et seulement si,

$$\begin{cases} -2\alpha & + & \beta & = A \\ -\alpha & - & 2\beta & = B \end{cases}$$

donc $\alpha=-\frac{2A+B}{5}$ et $\beta=\frac{A-2B}{5}$. Lorsque A et B décrivent $\mathbb{R}$, les constantes α et β décrivent $\mathbb{R}$. En conclusion, les solutions sont de la forme $x \mapsto \alpha\cos x+\beta\sin x+Ce^{2x}$, avec α, β et $C \in \mathbb{R}$.

Exercice 10

1. Comme y ne s'annule pas, la fonction z est dérivable et $z' = -\frac{y'}{y^2}$. Alors, l'équation en y équivaut à l'équation suivante en z obtenue en divisant par y^2 :

$$-z' + a(x)z + b(x) = 0.$$

2. Avec le formalisme de la première question, cette équation équivaut à $-z' + z = x$. Cette équation est linéaire d'ordre 1 à coefficients constants. Une solution particulière est $x \mapsto x+1$ (cas simple d'un second membre produit d'une exponentielle et d'un polynôme). Les solutions de l'équation homogène sont les fonctions $x \mapsto Ce^x$ avec $C \in \mathbb{R}$. Par conséquent, z est de la forme $x \mapsto x+1+Ce^x$ et y de la forme $x \mapsto \frac{1}{x+1+Ce^x}$ avec $C \in \mathbb{R}$.
Pour terminer la question, il convient de déterminer les fonctions y de cette forme définies sur $\mathbb{R}$. On étudie pour cela les variations de $x \mapsto x+1+Ce^x$ et on trouve que la fonction s'annule seulement si $C \geq -1$. En conclusion, les solutions qui ne s'annulent pas, définies sur $\mathbb{R}$ de l'équation sont les fonctions $x \mapsto \frac{1}{x+1+Ce^x}$ avec $C < -1$.

Exercice 11

Par définition, pour tout $x \in \mathbb{R}_+^*$, $y(x) = z(\ln x)$ donc

$$y'(x) = \frac{1}{x} z'(\ln x),$$
$$y''(x) = \frac{1}{x^2} z''(\ln x) - \frac{1}{x^2} z'(\ln x).$$

Ainsi, $x^2 y''(x) - 2y(x) = (z'' - z' - 2z)(\ln x)$ et nous devons résoudre $z''(t) - z'(t) - 2z(t) = e^t$. Les solutions de l'équation homogène sont les fonctions $t \mapsto Ae^{2t} + Be^{-t}$ avec $A, B \in \mathbb{R}$ et une solution particulière est $t \mapsto -\frac{1}{2}e^t$ (le second membre de cette équation linéaire est le produit d'une exponentielle et d'un polynôme). Par conséquent, z est de la forme $t \mapsto Ae^{2t} + Be^{-t} - \frac{1}{2}e^t$ et y est de la forme

$$x \mapsto Ax^2 + B\frac{1}{x} - \frac{1}{2}x.$$

La réciproque est immédiate.

Exercice 12

Cherchons l'équation sous la forme $y' + a(x)y = b(x)$.
D'après le théorème de structure, les solutions de l'équation homogène sont les $x \mapsto \frac{C}{1+x^2}$ pour $C \in \mathbb{R}$ donc $x \mapsto \frac{1}{1+x^2}$ est solution de $y' + a(x)y = 0$. En remplaçant dans cette équation, on obtient que la fonction a est définie par $a(x) = \frac{2x}{1+x^2}$.
Il reste alors à exploiter que $x \mapsto \frac{x}{1+x^2}$ est solution de l'équation avec second membre pour déterminer l'expression de b :

$$b(x) = \frac{(1+x^2) - 2x^2}{(1+x^2)^2} + \frac{2x}{1+x^2}\frac{x}{1+x^2} = \frac{1}{1+x^2}.$$

En conclusion, l'équation recherchée est $y' + \frac{2x}{1+x^2}y = \frac{1}{1+x^2}$.

Exercice 13

On remarque que les deux premières équations donnent la condition

$$u' = x'' + iy'' = \omega y' - i\omega x' = -i\omega u.$$

Par conséquent, u est de la forme $t \mapsto Ce^{-i\omega t}$ avec $C \in \mathbb{C}$. En notant $C = Ae^{i\varphi}$ avec $A \geq 0$ et $\varphi \in \mathbb{R}$, on obtient

$$\left\{\begin{array}{lcl} x' & = & A\cos(-\omega t+\varphi) \\ y' & = & A\sin(-\omega t+\varphi) \\ z'' & = & 0 \end{array}\right.$$

d'où

$$\left\{\begin{array}{lcl} x & = & -\frac{A}{\omega}\sin(-\omega t+\varphi)+\alpha \\ y & = & \frac{A}{\omega}\cos(-\omega t+\varphi)+\beta \\ z & = & \gamma t+\delta \end{array}\right.$$

avec A, φ, α, β, γ et $\delta \in \mathbb{R}$.

Exercice 14

Soit f une fonction continue telle que $\varphi : x \mapsto f(x) - \int_0^x t f(t)\,\mathrm{d}t$ est constante. Alors, φ est dérivable donc

$$f : x \mapsto \int_0^x t f(t)\,\mathrm{d}t + \varphi(x)$$

est dérivable. Ainsi, la condition φ constante équivaut à φ' nulle donc f est solution de l'équation $y' - xy = 0$.

En conclusion, les fonctions recherchées sont les $x \mapsto Ce^{\frac{x^2}{2}}$ avec $C \in \mathbb{R}$.

Exercice 15

L'équation se réécrit : pour tout $x \in \mathbb{R}$,

$$\begin{aligned} f(x) &= 1 + \int_0^x f(x-t)\cos t\,\mathrm{d}t \\ &= 1 + \int_0^x f(u)\cos(x-u)\,\mathrm{d}u && \text{en posant } u = x-t \\ &= 1 + \cos x \int_0^x f(u)\cos u\,\mathrm{d}u + \sin x \int_0^x f(u)\sin u\,\mathrm{d}u. \end{aligned}$$

Cette nouvelle formulation permet de montrer qu'une solution f est dérivable puisque le membre de droite est dérivable comme somme et produit de fonctions dérivables. Par conséquent, l'équation se réécrit encore $f(0) = 1$ et, pour tout $x \in \mathbb{R}$,

$$\begin{aligned} f'(x) &= \cos^2 x f(x) - \sin x \int_0^x f(u)\cos u\,\mathrm{d}u + \sin^2 x f(x) + \cos x \int_0^x f(u)\sin u\,\mathrm{d}u \\ &= f(x) - \sin x \int_0^x f(u)\cos u\,\mathrm{d}u + \cos x \int_0^x f(u)\sin u\,\mathrm{d}u. \end{aligned}$$

Ceci permet à nouveau de remarquer que f' est dérivable puis de donner la formulation équivalente suivante : $f(0) = f'(0) = 1$ et, pour tout $x \in \mathbb{R}$,

$$\begin{aligned} f''(x) &= f'(x) - \sin x \cos x f(x) - \cos x \int_0^x f(u)\cos u\,\mathrm{d}u \\ &\quad + \sin x \cos x f(x) - \sin x \int_0^x f(u)\sin u\,\mathrm{d}u \\ &= f'(x) - f(x) + 1. \end{aligned}$$

En résolvant l'équation, on obtient que f est de la forme

$$x \mapsto 1 + e^{\frac{x}{2}}\left(A\cos\frac{\sqrt{3}}{2}x + B\sin\frac{\sqrt{3}}{2}x\right),$$

avec A et $B \in \mathbb{R}$. En utilisant les conditions initiales, on trouve $A = 0$ et $B = \frac{2}{\sqrt{3}}$.

En conclusion, l'unique solution est la fonction $x \mapsto 1 + \frac{2}{\sqrt{3}} e^{\frac{x}{2}} \sin\frac{\sqrt{3}}{2}x$.

Corrigés des exercices d'approfondissement du chapitre 5

Exercice A

Supposons qu'il existe une telle solution y. Alors, y' diverge vers $+\infty$ comme somme d'une fonction qui diverge vers $+\infty$ en $+\infty$ (à savoir ln) et d'une fonction bornée (à savoir $-y$). Par conséquent, il existe $A > 0$ tel que, pour tout $x \geq A$, $y'(x) \geq 1$. En intégrant ceci entre A et $x \geq A$, on obtient la minoration $y(x) \geq x - A + y(A)$ donc, par comparaison, y diverge vers $+\infty$ en $+\infty$ ce qui contredit le caractère bornée de y.

Exercice B

Discutons selon le signe de $\Delta = a^2 - 4b$ du discriminant de l'équation caractéristique.

▹ Si $\Delta < 0$, les racines complexes conjuguées de l'équation caractéristique sont $\frac{-a \pm i\sqrt{-\Delta}}{2}$. Les solutions sont

$$x \mapsto e^{-\frac{a}{2}x}\left(A\cos\frac{\sqrt{-\Delta}}{2}x + B\sin\frac{\sqrt{-\Delta}}{2}x\right),$$

avec A, $B \in \mathbb{R}$. Elles sont toutes bornées sur $\mathbb{R}_+$ si, et seulement si, $a > 0$ (et donc $b > \frac{a^2}{4} > 0$).

▹ Si $\Delta = 0$, la racine double de l'équation caractéristique est $-\frac{a}{2}$. Les solutions sont

$$x \mapsto (Ax + B)e^{-\frac{a}{2}x},$$

avec A, $B \in \mathbb{R}$. Elles sont toutes bornées sur $\mathbb{R}_+$ si, et seulement si, $a \geq 0$ (et donc $b = \frac{a^2}{4} \geq 0$).

▹ Si $\Delta > 0$, les deux racines réelles x_1 et x_2 de l'équation caractéristique sont $\frac{-a \pm \sqrt{\Delta}}{2}$. Les solutions sont bornées sur $\mathbb{R}_+$ si, et seulement si, ces deux racines sont négatives.

- Si $x_1 \leq 0$ et $x_2 \leq 0$, alors $a = -(x_1 + x_2) \geq 0$ et $b = x_1 x_2 \geq 0$.
- Si $a \geq 0$ et $b \geq 0$, alors $\Delta \leq a^2$ donc $-a \pm \sqrt{\Delta} \leq 0$ et par conséquent, $x_1 \leq 0$ et $x_2 \leq 0$.

Dans ce cas encore, les solutions sont bornées sur $\mathbb{R}_+$ si, et seulement si, $a \geq 0$ et $b \geq 0$.

En conclusion, toutes les solutions de l'équation $y'' + ay' + by = 0$ sont bornées sur $\mathbb{R}_+$ si, et seulement si, $a \geq 0$ et $b \geq 0$ et $(a, b) \neq (0, 0)$.

Exercice C

▹ Dérivons cette relation (on trouve ainsi une condition nécessaire mais *a priori* non suffisante). Pour tout $x \in \mathbb{R}$,

$$f''(x) = -f'(a - x) = -f(a - (a - x)) = -f(x).$$

Ainsi, en résolvant cette équation différentielle linéaire d'ordre 2 à coefficients constants, f est de la forme $x \mapsto A\cos x + B\sin x$ avec A, $B \in \mathbb{R}$.

▹ Déterminons à quelle condition sur A et $B \in \mathbb{R}$, la fonction $x \mapsto A\cos x + B\sin x$ est solution du problème initial. On doit avoir, pour tout $x \in \mathbb{R}$,

$$\begin{aligned} -A\sin x + B\cos x &= A\cos(a-x) + B\sin(a-x) \\ &= (A\sin a - B\cos a)\sin x + (A\cos a + B\sin a)\cos x. \end{aligned}$$

En identifiant, on obtient le système suivant

$$\left\{ \begin{array}{rcrcl} (\sin a + 1) & A & - & \cos a & B & = 0 \\ \cos a & A & + & (\sin a - 1) & B & = 0 \end{array} \right.$$

- Si $a = \frac{\pi}{2}[2\pi]$, alors ce système équivaut à $A = 0$.
- Si $a = -\frac{\pi}{2}[2\pi]$, alors ce système équivaut à $B = 0$.
- Sinon, le système équivaut à $B = \frac{\sin a + 1}{\cos a} A$.

On en déduit selon les cas les solutions du problème initial.

Exercice D

Notons $\varphi : x \mapsto A + \int_0^x f(t)g(t)\,\mathrm{d}t$. La fonction φ est de classe $\mathscr{C}^1$ comme primitive d'une fonction continue et, pour tout $x \in \mathbb{R}_+^*$,

$$\varphi'(x) = f(x)g(x) \leq g(x)\varphi(x).$$

Ainsi, $(\varphi'(x) - g(x)\varphi(x))\exp\left(-\int_0^x g(t)\,\mathrm{d}t\right) \leq 0$. Le premier membre est la dérivée de la fonction

$$x \mapsto \varphi(x)\exp\left(-\int_0^x g(t)\,\mathrm{d}t\right)$$

donc en intégrant entre 0 et x, on obtient

$$\varphi(x)\exp\left(-\int_0^x g(t)\,\mathrm{d}t\right) - \varphi(0)\exp\left(-\int_0^0 g(t)\,\mathrm{d}t\right) \leq 0,$$

soit

$$\varphi(x)\exp\left(-\int_0^x g(t)\,\mathrm{d}t\right) \leq \varphi(0) = A,$$

d'où l'inégalité demandée.

Corrigés des problèmes du chapitre 5

Problème 1

1. Une équation est $y = F(x_0, y_0)(x - x_0) + y_0$.
2. a. Si $\alpha(x_0) = 0$, alors $F(x_0, y_0) = \beta(x_0)$: la pente des tangentes considérées ne dépend pas de y_0 et toutes ces droites sont donc parallèles.
b. Si $\alpha(x_0) \neq 0$, les tangentes aux courbes représentatives des solutions telles que $y(x_0) = 0$ et $y(x_0) = 1$ se coupent en $\left(x_0 - \frac{1}{\alpha(x_0)}, -\frac{\beta(x_0)}{\alpha(x_0)}\right)$. On vérifie immédiatement que ce point appartient à toutes les tangentes considérées.
3. Remarquons que si toutes les tangentes en x_0 sont parallèles, alors elles ont la même pente $F(x_0, y)$ qui ne dépend donc pas de y ; si elles sont concourantes, le point de concours des tangentes en y_0 et y_1 est

$$(x_c, y_c) = \left(\frac{F(x_0,y_0)x_0 - y_0 - F(x_0,y_1)x_0 - y_1}{F(x_0,y_0) - F(x_0,y_1)}, \frac{y_1 F(x_0,y_0) - y_0 F(x_0,y_1)}{F(x_0,y_0) - F(x_0,y_1)}\right).$$

a. D'après la remarque précédente, si les tangentes sont parallèles, cette quantité est nulle, sinon

$$\frac{F(x_0,y_0)-F(x_0,y_1)}{y_0-y_1}=\frac{1}{x_0-x_c}.$$

b. Après des calculs élémentaires, $F(x_0,y_0)-\alpha(x_0)y_0=-y_c\alpha(x_0)$.
c. Comme les résultats précédents sont valables pour tout $x_0\in J$, $F(x,y)=\alpha(x)y+\beta(x)$ en notant $\beta(x)=F(x,y)-\alpha(x)y$ (qui ne dépend pas de y d'après la question précédente).

Problème 2

1. a. Les solutions sont $x\mapsto\kappa_1\cos(x)+\kappa_2\sin(x)+x+1$ avec $(\kappa_1,\kappa_2)\in\mathbb{R}^2$. On remarque que $y(0)=\kappa_1+1$, $y(\frac{\pi}{2})=\kappa_2+\frac{\pi}{2}+1$ et $y'(\frac{\pi}{2})=-\kappa_1+1$.
b. Il existe donc une unique solution telle que $y(0)=y(\frac{\pi}{2})=0$ avec $\kappa_1=-1$ et $\kappa_2=-1-\frac{\pi}{2}$.
Il n'existe aucune solution telle que $y(0)=y'(\frac{\pi}{2})=0$ (il faudrait $\kappa_1=1$ et $\kappa_1=-1$).
2. On dérive deux fois la première relation, on utilise l'équation (E) et on trouve un système de deux équations à deux inconnues $A'(x)$ et $B'(x)$:

$$\begin{cases} -A'(x)\sin(x) & + & B'(x)\cos(x) & = & f(x) \\ A'(x)\cos(x) & + & B'(x)\sin(x) & = & 0 \end{cases}$$

On trouve $A'(x)=-f(x)\sin(x)$ et $B'(x)=f(x)\cos(x)$ et il ne reste qu'à intégrer.
3. D'après la question précédente, une solution particulière est

$$y(x)=-\int_0^x \sin(t)f(t)\,\mathrm{d}t\cos(x)+\int_0^x \cos(t)f(t)\,\mathrm{d}t\sin(x)=\int_0^x \sin(x-t)f(t)\,\mathrm{d}t\,.$$

On conclut en utilisant les solutions de l'équation homogène.
4. Pour l'existence, il suffit de vérifier que la fonction donnée est bien solution.
Pour l'unicité, considérons y et z sont deux solutions de (E) vérifiant ces conditions initiales, alors $y-z$ est solution de l'équation homogène donc de la forme $x\mapsto\kappa_1\cos(x)+\kappa_2\sin(x)$ et les conditions initiales donnent $\kappa_1=\kappa_2=0$ et par conséquent $y=z$.
5. Avec la formule établie, on trouve $y(0)=\kappa_1$ et

$$y'\left(\tfrac{\pi}{2}\right)=-\kappa_1+\int_0^{\pi/2}\sin(t)f(t)\,\mathrm{d}t\,.$$

Il n'y a donc des solutions si, et seulement si, $\int_0^{\pi/2}\sin(t)f(t)\,\mathrm{d}t=0$. On remarque qu'il y a alors une infinité de solutions (car il n'y a pas de condition sur κ_2).

Problème 3

Partie A – Réduction de l'équation

1. Commençons par calculer les dérivées de y en fonction de z

$$y'=u'z+uz'$$
$$y''=u''z+2u'z'+uz''.$$

Ainsi, en remplaçant,

$$ay''+by'+cy=uaz''+(2u'a+ub)z'+(u''a+u'b+uc)z.$$

En conclusion, y est solution de (E) si, et seulement si, z est solution de l'équation

$$uaz''+(2u'a+ub)z'+(u''a+u'b+uc)z=0.$$

2. D'après ce calcul, si u est solution de $2u'a + ub = 0$, alors z est solution d'une équation (E') de la forme $\alpha(x)z'' + \beta(x)z = 0$ (avec $\alpha = ua$ et $\beta = u''a + u'b + uc$).

Partie B – Entrelacement des zéros

1. Après calculs, $W'x) = y_1(x)y_2''(x) - y_1''(x)y_2(x) = (p(x) - q(x))y_1(x)y_2(x)$.
2. a. Supposons (sans perte de généralité quitte à remplacer y_2 par $-y_2$) que $y_2'(x_1) > 0 > y_2'(x_2)$ donc que $y_2(x) > 0$ pour tout $x \in]x_1, x_2[$. D'après la question précédente, $W'(x)$ est alors du signe de $y_1(x)$.
• Si $y_1(x) > 0$ pour tout $x \in [x_1, x_2]$, alors W est croissante et

$$W(x_1) = y_1(x_1)y_2'(x_1) > 0 > y_1(x_2)y_2'(x_2) = W(x_2),$$

ce qui contredit la croissance.
• Si $y_1(x) < 0$ pour tout $x \in [x_1, x_2]$, alors W est décroissante et

$$W(x_1) = y_1(x_1)y_2'(x_1) < 0 < y_1(x_2)y_2'(x_2) = W(x_2),$$

ce qui contredit la décroissance.
Par conséquent, y_1 change de signe sur $[x_1, x_2]$ et donc s'annule d'après la propriété des valeurs intermédiaires.
b. Supposons que $p(x) \geq \omega^2 > 0$ pour tout $x \in \mathbb{R}$ et considérons $y_2 : x \mapsto \sin(\omega x + \varphi)$. D'après la question précédente, entre deux zéros consécutifs de y_2 (séparés de $\frac{\pi}{\omega}$), il existe au moins un zéro de y_1. Or, y_2 admet une infinité de zéros d'où le résultat.
c. Soit y_1 la fonction constante égale à 1 qui est solution de $y'' = 0$. Si y_2 admet deux zéros distincts, alors y_1 admet un zéro ce qui est absurde.
3. Soit $x_1 < x_2$ deux zéros consécutifs de y_2. D'après la question précédente, y_1 admet un zéro dans l'intervalle $[x_1, x_2]$.
Si y_1 et y_2 n'ont pas de zéro commun, alors y_1 admet un zéro dans l'intervalle ouvert $]x_1, x_2[$; si y_1 admet deux zéros distincts dans l'intervalle ouvert $]x_1, x_2[$, alors y_2 y admet aussi un zéro (toujours la même question mais en échangeant les rôles de y_1 et y_2) ce qui est impossible puisque x_1 et x_2 sont des zéros consécutifs.

Partie C – Équation de Bessel

Posons, comme dans la partie A, $y = \frac{1}{\sqrt{x}}z$; la fonction z vérifie alors

$$z'' + \left(1 - \frac{4\lambda^2 - 1}{4x^2}\right)z = 0.$$

Si $\lambda \geq \frac{1}{2}$, alors $1 - \frac{4\lambda^2-1}{4x^2} \leq 1$ donc z admet au plus un zéro dans un intervalle de longueur π (en utilisant l'entrelacement avec la fonction $x \mapsto \sin(x - x_0)$ solution de $y'' + y = 0$). Comme z et y ont les mêmes zéros, on termine la question avec la même discussion.

Chapitre 6

Suites numériques

1. Calculs explicites de suites — **2.** Manipulations de suites — **3.** Suites convergentes — **4.** Relations de comparaison pour les suites — **5.** Suites réelles — **6.** Suites extraites

Objectifs et compétences du programme

- Expliciter des suites arithmético-géométriques et des suites vérifiant une relation de récurrence linéaire d'ordre 2.
- Manipuler la définition quantifiée de convergence.
- Maîtriser les relations de comparaison asymptotique.
- Utiliser encadrements et monotonies pour établir la convergence d'une suite réelle.
- Trouver dans des cas simples les valeurs d'adhérence d'une suite.

Karl Weierstrass, mathématicien allemand (1815-1897)
Il est le père de l'analyse telle qu'on la pratique actuellement. On lui doit notamment les fameuses définitions en ε et δ (ou η dans le cadre de ce livre) pour manipuler correctement les notions de limites et de continuités. Plusieurs objets portent désormais son nom :
- une fonction continue partout et dérivable nulle part ;
- une fonction $\mathscr{P}$ de la variable complexe qui admet deux périodes non proportionnelles ;
- le théorème de Bolzano-Weierstrass dans ce chapitre.

Certains appellent même théorème de Weierstrass, le théorème des bornes atteintes abordé dans le chapitre suivant.

COURS

On ne sait expliciter le terme général que de quelques suites définies par récurrence. Aussi, après avoir traité ces exemples dans la première partie, on cherche à étudier plus généralement le comportement asymptotique d'une suite (convergence ou divergence). Il y a d'une part des méthodes générales qui découlent directement de la définition à partir d'une phrase quantifiée (troisième partie) et d'autre part des méthodes plus spécifiques qui s'appuient sur les propriétés du corps des réels (existence d'une relation d'ordre total, propriété de borne supérieure) et qui fournissent des outils puissants : lemme d'encadrement, théorème de limite monotone (cinquième section). La divergence peut quelquefois être précisée soit en utilisant des estimations asymptotiques avec les relations de comparaison (quatrième section) soit en se limitant à des suites extraites (sixième section).

1. Calculs explicites de suites

Dans cette partie, on effectue quelques calculs du terme général de suites particulières.

1.1. Récurrence linéaire d'ordre 1

Définition 6.1.

▹ Une suite arithmétique est une suite (u_n) vérifiant une relation de récurrence de la forme

$$\forall n \in \mathbb{N}, \qquad u_{n+1} = u_n + b.$$

▹ Une suite géométrique est une suite (u_n) vérifiant une relation de récurrence de la forme

$$\forall n \in \mathbb{N}, \qquad u_{n+1} = a\,u_n.$$

▹ Une suite arithmético-géométrique est une suite (u_n) vérifiant une relation de récurrence de la forme

$$\forall n \in \mathbb{N}, \qquad u_{n+1} = a\,u_n + b.$$

Proposition 6.2.

Soit (u_n) une suite vérifiant la relation

$$\forall n \in \mathbb{N}, \qquad u_{n+1} = a\,u_n + b$$

▹ Si $a = 1$ (cas d'une suite arithmétique), alors, pour tout n, $u_n = u_0 + n\,b$.
▹ Si $a \neq 1$, alors, pour tout n, $u_n = \ell + a^n(u_0 - \ell)$ avec $\ell = b/(1-a)$.

Remarque

Si $b = 0$ (cas d'une suite géométrique), alors $\ell = 0$ et on retrouve bien $u_n = u_0 a^n$ pour tout n.

Démonstration

▹ Le cas $a = 1$ se démontrer par une simple récurrence.

▹ Si $a \neq 1$, on remarque que

$$\forall n \in \mathbb{N}, \qquad u_{n+1} - \ell = a(u_n - \ell).$$

La suite $(u_n - \ell)_n$ est alors géométrique de raison a et on conclut à nouveau par récurrence. ■

On peut généraliser l'exemple des suites arithmético-géométriques.

Proposition 6.3.

Soit $a \in \mathbb{C} \setminus \{0\}$ et (b_n) une suite. Considérons alors une suite (u_n) vérifiant

$$\forall n \in \mathbb{N}, \qquad u_{n+1} = a\, u_n + b_n.$$

Alors, pour tout $n \in \mathbb{N}$,

$$u_n = a^n \left(u_0 + \sum_{k=0}^{n-1} \frac{b_k}{a^{k+1}} \right).$$

Démonstration

Divisions la relation de récurrence par a^{n+1} : pour tout $n \in \mathbb{N}$,

$$\frac{u_{n+1}}{a^{n+1}} - \frac{u_n}{a^n} = \frac{b_n}{a_{n+1}}.$$

On conclut en calculant la somme télescopique

$$\begin{aligned} \frac{u_n}{a^n} - \frac{u_0}{1} &= \sum_{k=0}^{n-1} \left(\frac{u_{k+1}}{a^{k+1}} - \frac{u_k}{a^k} \right) \\ &= \sum_{k=0}^{n-1} \frac{b_k}{a^{k+1}}. \end{aligned}$$

■

Exemple

Une suite (u_n) vérifiant $u_{n+1} = 2u_n + 3^{n+1}$ pour tout n est donnée par

$$\forall n \geq 1, \qquad u_n = 2^n \left(u_0 + \sum_{k=1}^{n} \frac{3^k}{2^k} \right)$$

et l'on sait calculer cette somme géométrique de raison $\frac{3}{2}$ pour obtenir

$$\forall n \geq 1, \qquad u_n = (u_0 - 3)2^n + 3.3.$$

1.2. Récurrence linéaire d'ordre 2

Proposition 6.4. Suite récurrente linéaire d'ordre 2

Soit $(u_n)_n$ une suite vérifiant la relation de récurrence

$$\forall n \in \mathbb{N}, \qquad u_{n+2} = a\, u_{n+1} + b\, u_n.$$

Considérons l'équation (caractéristique) $\ell^2 = a\ell + b$.

▹ Si cette équation admet deux solutions distinctes r_1 et r_2, alors

$$\forall n \in \mathbb{N}, \qquad u_n = A r_1^n + B r_2^n,$$

avec A, $B \in \mathbb{C}$.

▹ Si cette équation admet une unique solution r, alors

$$\forall n \in \mathbb{N}, \qquad u_n = (An+B)r^n,$$

avec A, $B \in \mathbb{C}$.

Démonstration

Notons r_1, r_2 les racines (éventuellement égales) de l'équation caractéristique et définissons une suite $(v_n)_n$ par $v_n = u_{n+1} - r_1 u_n$. Alors

$$v_{n+1} = a\,u_{n+1} + b\,u_n - r_1 u_{n+1} = r_2(u_{n+1} - r_1 u_n) = r_2 v_n$$

car, d'après les relations coefficients-racines $a = r_1 + r_2$ et $b = -r_1 r_2$. Ainsi, pour tout $n \in \mathbb{N}$,

$$u_{n+1} - r_1 u_n = v_0 r_2^n.$$

Résolvons cette équation en $(u_n)_n$.

▹ Si $r_1 \neq r_2$, alors la suite $(u'_n)_n$ définie par $u'_n = \frac{v_0}{r_2 - r_1} r_2^n$ est une solution de l'équation et $(u_n - u'_n)_n$ est une suite géométrique de raison r_1 : en conclusion, la suite $(u_n)_n$ est combinaison linéaire de suites géométriques de raison r_1 et r_2.

▹ Si $r_1 = r_2$, alors la suite $(u'_n)_n$ définie par $u'_n = v_0 n r_1^{n-1}$ est une solution de l'équation et $(u_n - u'_n)_n$ est une suite géométrique de raison r_1 : en conclusion, la suite $(u_n)_n$ est combinaison linéaire de suites $(r_1^n)_n$ et $(n r_1^n)_n$. ■

Exemple

La suite de Fibonacci est la suite $(F_n)_n$ définie par $F_0 = 0$, $F_1 = 1$ et la relation de récurrence

$$\forall n \in \mathbb{N}, \qquad F_{n+2} = F_{n+1} + F_n.$$

Pour obtenir une formule close donnant F_n, déterminons les racines de l'équation caractéristique $x^2 - x - 1 = 0$. Après un rapide calcul, on trouve les racines $\Phi = \frac{1+\sqrt{5}}{2}$ (nombre d'or) et $\overline{\Phi} = \frac{1-\sqrt{5}}{2}$ (notation malheureuse car il ne s'agit pas du conjugué au sens des complexes). Ainsi, il existe A et $B \in \mathbb{R}$ tels que

$$\forall n \in \mathbb{N}, \qquad F_n = A\Phi^n + B\overline{\Phi}^n.$$

En évaluant en $n = 0$ et $n = 1$, on détermine les valeurs des constantes et on obtient

$$\forall n \in \mathbb{N}, \qquad F_n = \frac{1}{\sqrt{5}}\Phi^n - \frac{1}{\sqrt{5}}\overline{\Phi}^n.$$

Remarquons qu'avec cette expression, il ne semble pas évident que les valeurs F_n sont des entiers naturels (alors que c'est immédiat avec la relation de récurrence).

Remarque

On note dans le calcul précédent que les valeurs initiales n'interviennent que dans la détermination des constantes A et B. Pour une suite $(F_n)_n$ définie par $F_0 = 1$, $F_1 = 1$ et la relation de récurrence

$$\forall n \in \mathbb{N}, \qquad F_{n+2} = F_{n+1} + F_n,$$

on obtient

$$\forall n \in \mathbb{N}, \qquad F_n = \frac{\overline{\Phi} - 1}{\overline{\Phi} - \Phi}\Phi^n + \frac{\Phi - 1}{\Phi - \overline{\Phi}}\overline{\Phi}^n.$$

Exemple

Déterminons les suites $(u_n)_n$ vérifiant la relation de récurrence

$$\forall n \in \mathbb{N}, \qquad u_{n+2} = 2u_{n+1} - u_n.$$

Pour cela, remarquons que l'équation caractéristique $x^2 - 2x + 1 = 0$ admet 1 comme unique racine. D'après la proposition précédente, il existe A et B tels que

$$\forall n \in \mathbb{N}, \qquad u_n = An + B.$$

Il s'agit donc des suites « affines ».

Remarque

Si on étudie une suite récurrente linéaire d'ordre 2 réelle, telle que l'équation caractéristique admet deux racines complexes conjuguées $\rho e^{i\theta}$ et $\rho e^{-i\theta}$, il peut être plus commode d'écrire la suite sous la forme

$$\forall n \in \mathbb{N}, \qquad u_n = (\alpha \cos(n\theta) + \beta \sin(n\theta))\rho^n,$$

avec des constantes $\alpha, \beta \in \mathbb{R}$.

1.3. Récurrence homographique

Un exemple classique de récurrence que l'on peut résoudre en se ramenant à des suites géométriques et arithmétiques est le cas des suites homographiques.

Définition 6.5.

Une suite homographique est une suite $(u_n)_n$ vérifiant la relation de récurrence

$$\forall n \in \mathbb{N}, \qquad u_{n+1} = \frac{a u_n + b}{c u_n + d},$$

avec $ad - bc \neq 0$ et $c \neq 0$.

Remarque

La condition $ad - bc \neq 0$ garantit que les couples (a, b) et (c, d) ne sont pas proportionnels, c'est-à-dire que la fraction du membre de droite ne se simplifie pas en une constante.

Calculons explicitement les termes d'une telle suite. Pour cela considérons l'équation $\ell = \frac{a\ell + b}{c\ell + d}$.

Conseils méthodologiques

Ici, comme pour les suites arithmético-géométriques, on cherche les ℓ qui sont les limites possibles de la suite puis on va étudier l'écart du terme général de la suite à cette quantité.

▹ Si cette équation admet une unique solution $\ell = \frac{a-d}{2c}$, alors la suite $\left(\frac{1}{u_n-\ell}\right)_n$ est arithmétique ; en effet,

$$\begin{aligned}\frac{1}{u_{n+1}-\ell}-\frac{1}{u_n-\ell} &= \frac{c\,u_n+d}{a\,u_n+b-\ell(c\,u_n+d)}-\frac{1}{u_n-\ell}\\ &= \left(\frac{c\,u_n+d}{a-\ell c}-1\right)\frac{1}{u_n-\ell}\\ &= \frac{c\,u_n-c\ell}{a-\ell c}.\frac{1}{u_n-\ell}=\frac{c}{a-\ell c}=\frac{2c}{a+d}.\end{aligned}$$

Ainsi,

$$u_n = \ell + \frac{1}{\frac{1}{u_0-\ell}+n\frac{2c}{a+d}}.$$

▹ Si cette équation admet deux solutions distinctes ℓ_1 et ℓ_2, alors la suite $\left(\frac{u_n-\ell_1}{u_n-\ell_2}\right)_n$ est géométrique ; en effet, pour tout $n \in \mathbb{N}$,

$$\frac{u_{n+1}-\ell_1}{u_{n+1}-\ell_2}=\frac{(a-\ell_1 c)u_n+(b-\ell_1 d)}{(a-\ell_2 c)u_n+(b-\ell_2 d)}.$$

Or, $\ell_i = \frac{\ell_i d-b}{\ell_i c-a}$ pour $i = 1,2$ donc

$$\frac{u_{n+1}-\ell_1}{u_{n+1}-\ell_2}=\frac{a-\ell_1 c}{a-\ell_2 c}.\frac{u_n-\ell_1}{u_n-\ell_2}.$$

Ainsi, on obtient de même l'expression explicite de $(u_n)_n$.

2. Manipulations de suites

Rappelons rapidement quelques transformations sur les suites que l'on utilisera par la suite.

2.1. Manipulations algébriques

Définition 6.6.

Soit (u_n) et (v_n) deux suites. La somme et le produit terme à terme de ces suites sont les suites définies par (u_n+v_n) et $(u_n v_n)$.

Remarque

Le produit terme à terme de la suite (u_n) et de la constante λ est noté (λu_n) ou $\lambda(u_n)$.

Exemple

La somme des suites $((-1)^n)$ et $((-1)^{n+1})$ est la suite nulle.

Proposition 6.7.

La somme et le produit terme à terme de deux suites bornées sont bornées.

Démonstration
Il s'agit des propriétés du module. En effet, si (u_n) et (v_n) sont bornées respectivement par M et M', alors, pour tout $n \in \mathbb{N}$,

$$|u_n + v_n| \leq |u_n| + |v_n| \leq M + M',$$
$$|u_n v_n| = |u_n||v_n| \leq MM'.$$

■

2.2. Produit de Cauchy

Définition 6.8. Produit de Cauchy

Soit (u_n) et (v_n) deux suites. Le produit de Cauchy de ces suites est la suite (w_n) définie, pour tout n, par

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

Remarque
Le produit de Cauchy de deux suites bornées n'est pas forcément borné. Par exemple, en considérant pour (u_n) et (v_n) la suite constante égale à 1, le produit de Cauchy est la suite dont le $n^{\text{ième}}$ terme est

$$\sum_{k=0}^{n} 1 \times 1 = n + 1$$

et elle n'est évidemment pas bornée.

Proposition 6.9.

Le produit de Cauchy de deux suites non nulles est une suite non nulle.

Démonstration
Soit (u_n) et (v_n) des suites non nulles, (w_n) leur produit de Cauchy. Notons k le premier indice d'un terme non nul dans (u_n) et l le premier indice d'un terme non nul dans (v_n). Alors

$$w_{k+l} = \sum_{j=0}^{k+l} u_j v_{k+l-j} = \sum_{j=0}^{k-1} u_j v_{k+l-j} + u_k v_l + \sum_{j=k+1}^{k+l} u_j v_{k+l-j} = u_k v_l \neq 0.$$

■

Remarque
Par contraposition, si un produit de Cauchy était nul, l'une des suites de départ est nulle.

3. Suites convergentes

Toutes les suites considérées sont à valeurs dans un ensemble $\mathbb{K}$ qui désigne $\mathbb{R}$ ou $\mathbb{C}$.

3.1. Définitions et premières propriétés

Commençons par définir la nature (convergente ou divergente) d'une suite.

Définition 6.10.

▹ Une suite $(u_n)_n$ est convergente s'il existe $\ell \in \mathbb{K}$, appelé limite de $(u_n)_n$ et noté $\lim u_n$, tel que

$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N},\ \forall n \geq N, \qquad |u_n - \ell| \leq \varepsilon.$$

Si elle ne vérifie pas cette condition, la suite $(u_n)_n$ est divergente.
▹ Une suite réelle $(u_n)_n$ diverge vers $+\infty$ si

$$\forall A > 0,\ \exists N \in \mathbb{N},\ \forall n \geq N, \qquad u_n \geq A.$$

Conseils méthodologiques

▹ Pour utiliser cette définition : on peut choisir une valeur de ε (par exemple, on prendra $\varepsilon = 1$ un peu plus loin) et elle nous donne un rang N à partir duquel on dispose de l'inégalité.
▹ Pour démontrer une convergence : on part d'un ε strictement positif quelconque (le « Soit $\varepsilon > 0$ » au début des prochaines preuves) et on cherche un N tel que, pour tout $n \geq N$, $|u_n - \ell| \leq \varepsilon$ (ou de manière équivalente, puisque ε est quelconque, $|u_n - \ell| \leq K\varepsilon$ avec K une constante fixée).

Remarque

Remarquons tout d'abord que la condition de convergence ne dépend pas des premiers termes de la suite. Ce type de propriété dite asymptotique est conservée si l'on change les premiers termes.

Remarque

Comme dans toutes les phrases quantifiées, si l'on change l'ordre des quantificateurs, on modifie le sens de la définition. Par exemple, les suites $(u_n)_n$ satisfaisant

$$\exists \ell \in \mathbb{K},\ \exists N \in \mathbb{N},\ \forall \varepsilon > 0,\ \forall n \geq N, \qquad |u_n - \ell| \leq \varepsilon$$

sont les suites constantes à partir d'un certain rang.

Exemple

Une suite stationnaire est une suite constante à partir d'un certain rang. D'après la définition, les suites stationnaires sont convergentes.

Exemple

La suite $(n^{-\alpha})_n$ converge vers 0 pour tout $\alpha > 0$ puisque, pour tout $\varepsilon > 0$, la définition est vérifiée pour $N \geq \varepsilon^{-\frac{1}{\alpha}}$. En effet, par croissance des puissances d'exposant positif, si $n \geq N$, alors $n^{\alpha} \geq \varepsilon^{-1}$ et donc $n^{-\alpha} \leq \varepsilon$.

Exemple

Soit (u_n) une suite périodique convergente vers ℓ. Montrons qu'elle est constante égale à ℓ.
Par l'absurde, s'il existe un terme $u_k \neq \ell$, alors avec $\varepsilon < |u_k - \ell|$, on obtient que tous les termes au-delà d'un certain rang N vérifient $|u_n - \ell| < |u_k - \ell|$ ce qui contredit la périodicité (car la valeur u_k est prise par la suite après le rang N).

Proposition 6.11.

Si une suite converge, alors sa limite est unique.

L'idée de la démonstration est relativement simple. On raisonne par l'absurde en supposant qu'il existe deux limites ℓ et ℓ'. On choisit des bandes disjointes autour de chacune de ces valeurs. Par la définition de limite, on devrait avoir les termes (à partir d'un certain rang) de la suite à la fois dans l'une et dans l'autre bande.

Démonstration

Raisonnons par l'absurde. Si, pour une suite convergente $(u_n)_n$, il existe deux limites distinctes ℓ et ℓ', alors les définitions de convergence avec $\varepsilon = \frac{|\ell-\ell'|}{4}$ s'écrivent :

$$\exists N \in \mathbb{N},\ \forall n \geq N, \qquad |u_n - \ell| \leq \varepsilon,$$
$$\exists N' \in \mathbb{N},\ \forall n \geq N', \qquad |u_n - \ell'| \leq \varepsilon.$$

Ainsi, pour tout $n \geq \max\{N, N'\}$, on a, par inégalité triangulaire :

$$|\ell - \ell'| \leq |u_n - \ell| + |u_n - \ell'| \leq 2\varepsilon = \frac{|\ell - \ell'|}{2}.$$

D'où la contradiction. ■

Notation

On note de manière équivalente $\lim u_n = \ell$ ou $u_n \to \ell$.

Exemple

Le produit $(u_n v_n)_n$ d'une suite bornée $(u_n)_n$ par une suite $(v_n)_n$ tendant vers 0 est une suite tendant vers 0. En effet, fixons $\varepsilon > 0$; par définition, il existe $M \in \mathbb{R}$ tel que, pour tout $n \in \mathbb{N}$, $|u_n| \leq M$ et $N \in \mathbb{N}$ tel que

$$\forall n \geq N, \qquad |v_n| \leq \varepsilon.$$

Ainsi,

$$\forall n \geq N, \qquad |u_n v_n| \leq M\varepsilon$$

ce qui est la définition de $\lim u_n v_n = 0$.

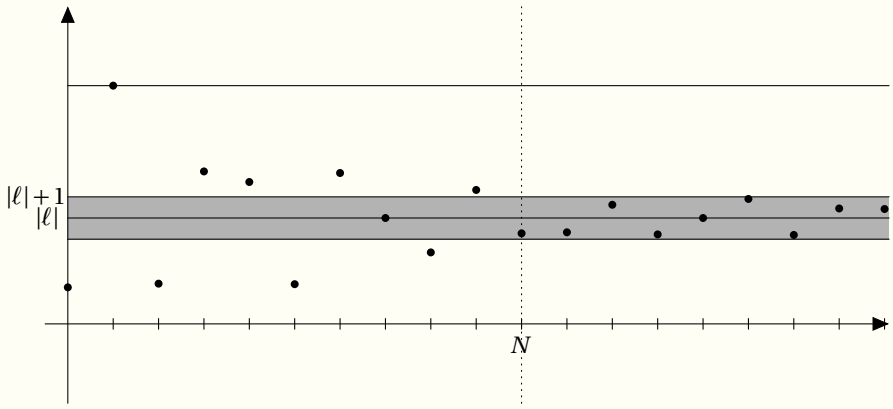
Proposition 6.12.

Toute suite convergente est bornée.

Démonstration

Appliquons la définition de convergence avec $\varepsilon = 1$.

$$\exists N \in \mathbb{N},\ \forall n \geq N, \qquad |u_n - \ell| \leq 1.$$



Ainsi, la suite $(u_n)_n$ est en valeur absolue bornée par la maximum des modules des premiers termes et de la quantité $|\ell|+1$ qui majore les termes à partir d'un certain rang, c'est-à-dire

$$\max\{|u_0|,\dots,|u_N|,|\ell|+1\}.$$

■

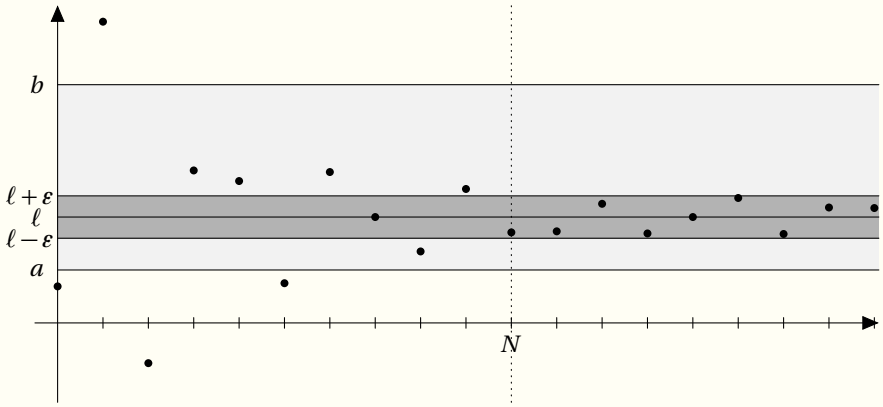
Remarque

La réciproque à cette propriété est fausse. En effet, la suite $((-1)^n)_n$ est divergente et bornée. En revanche, par contraposition, cette propriété est un critère permettant de montrer la divergence d'une suite qui n'est pas bornée.

Proposition 6.13.

Soit $(u_n)_n$ une suite réelle convergeant vers $\ell \in]a,b[$. Alors, il existe $N \in \mathbb{N}$ tel que

$$\forall n \geq N, \qquad u_n \in]a,b[.$$



Démonstration

Il suffit d'appliquer la définition avec $\varepsilon < \min\{b-\ell, \ell-a\}$. Ainsi, il existe un rang N tel que, pour tout $n \geq N$,

$$a < \ell-\varepsilon \leq u_n \leq \ell+\varepsilon < b.$$

■

On en déduit immédiatement (en considérant un intervalle ouvert contenant ℓ et inclus dans $\mathbb{R}_+^*$) le corollaire suivant.

Corollaire 6.14.

Soit $(u_n)_n$ une suite réelle convergeant vers $\ell > 0$.
Alors, il existe $N \in \mathbb{N}$ tel que pour tout $n \geq N$, $u_n > 0$.

Corollaire 6.15.

Une suite convergente à valeurs dans $\mathbb{Z}$ est stationnaire (c'est-à-dire constante à partir d'un certain rang).

Démonstration

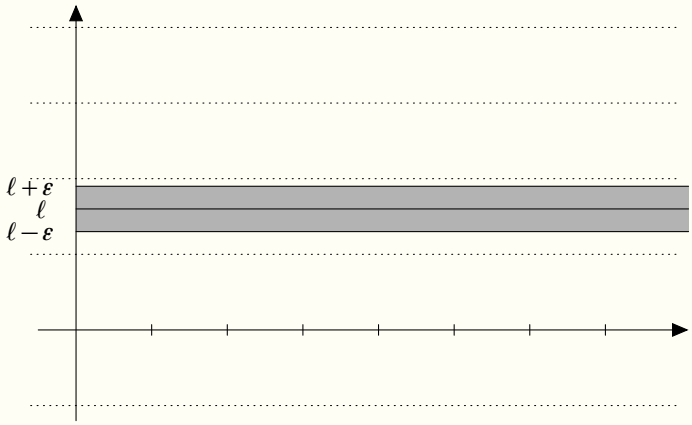
Soit $(u_n)_n$ une suite à valeurs dans $\mathbb{Z}$ de limite ℓ.

▹ Remarquons que $\ell \in \mathbb{Z}$ car sinon, il existe $\varepsilon > 0$ tel que $\mathbb{Z} \cap]\ell - \varepsilon, \ell + \varepsilon[= \varnothing$ donc

$$\forall n \in \mathbb{N}, \qquad |u_n - \ell| \geq \varepsilon,$$

ce qui contredit la convergence.

Graphiquement, les termes de la suite doivent être dans la bande autour de ℓ et sur les lignes pointillées (qui correspondent aux ordonnées entières) : c'est impossible!.

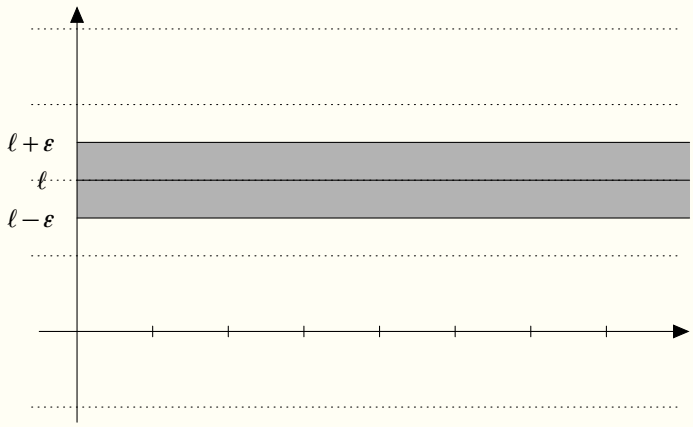


▹ Désormais $\ell \in \mathbb{Z}$. En écrivant la définition de convergence avec $\varepsilon = \frac{1}{2}$, il existe $N \in \mathbb{N}$ tel que

$$\forall n \geq N, \qquad |u_n - \ell| \leq \frac{1}{2}.$$

Comme $\mathbb{Z} \cap [\ell - \frac{1}{2}, \ell + \frac{1}{2}] = \{\ell\}$, la suite $(u_n)_n$ est constante égale à ℓ à partir du rang N.

Graphiquement, les termes de la suite doivent être dans la bande autour de ℓ et sur les lignes pointillées : il n'y a qu'une valeur possible!.



■

3.2. Propriétés opératoires

L'idée des résultats suivants est relativement simple : on cherche à estimer l'écart entre la suite considérée et sa limite supposée, que l'on majore en faisant apparaître les termes $|u_n-\ell|$ et $|v_n-\ell'|$ avec ℓ et ℓ' les limites de $(u_n)_n$ et $(v_n)_n$.

Proposition 6.16.

Soit $(u_n)_n$ une suite convergente. Alors, la suite $(|u_n|)_n$ est convergente et sa limite est $|\lim u_n|$.

Démonstration

Notons ℓ la limite de la suite $(u_n)_n$. Soit $\varepsilon>0$; d'après la définition de convergence, il existe un rang N tel que

$$\forall n\geq N,\qquad |u_n-\ell|\leq\varepsilon.$$

D'après l'inégalité triangulaire, pour tout $n\geq N$,

$$||u_n|-|\ell||\leq|u_n-\ell|\leq\varepsilon$$

Ce qui est bien le résultat attendu. ■

Exemple

Soit $(u_n)_n$ une suite convergente à valeurs dans le cercle unité $\mathbb{U}$. D'après ce résultat, la limite ℓ appartient à $\mathbb{U}$ puisque $1=|u_n|\to|\ell|$.

Proposition 6.17.

Soit $(u_n)_n$ et $(v_n)_n$ deux suites convergentes, $\lambda,\mu\in\mathbb{K}$. Alors, la suite $(\lambda u_n+\mu v_n)_n$ est convergente et sa limite est $\lambda\lim u_n+\mu\lim v_n$.

Démonstration

Notons ℓ et ℓ' les limites des suites $(u_n)_n$ et $(v_n)_n$. D'après l'inégalité triangulaire, pour tout n,

$$|\lambda u_n+\mu v_n-(\lambda\ell+\mu\ell')|\leq|\lambda|.|u_n-\ell|+|\mu|.|v_n-\ell'|,$$

d'où l'on conclut encore avec les définitions de $\lim u_n=\ell$ et $\lim v_n=\ell'$. ■

Proposition 6.18.

Soit $(u_n)_n$ et $(v_n)_n$ deux suites convergentes. Alors, la suite $(u_n v_n)_n$ est convergente et sa limite est $\lim u_n . \lim v_n$.

Démonstration

Notons ℓ et ℓ' les limites des suites $(u_n)_n$ et $(v_n)_n$. Commençons par introduire le terme ℓv_n dans les inégalités suivantes pour tout n

$$\begin{aligned}|u_n v_n - \ell\ell'| &= |u_n v_n - \ell v_n + \ell v_n - \ell\ell'| \leq |u_n v_n - \ell v_n| + |\ell v_n - \ell\ell'| \\ &\leq |v_n|.|u_n - \ell| + |\ell|.|v_n - \ell'|.\end{aligned}$$

On conclut avec les définitions de $\lim u_n = \ell$ et $\lim v_n = \ell'$ et le fait que la suite $(v_n)_n$ est bornée. ■

Proposition 6.19.

Soit $(u_n)_n$ une suite convergente telle que $\lim u_n \neq 0$. Alors, la suite $(\frac{1}{u_n})_n$ est convergente et sa limite est $\frac{1}{\lim u_n}$.

Démonstration

Notons ℓ la limite de la suite $(u_n)_n$. Comme la suite $(|u_n|)_n$ converge vers $|\ell| \neq 0$, alors, à partir d'un certain rang N, $|u_n| > \frac{1}{2}|\ell|$; ainsi, pour tout $n \geq N$,

$$\left|\frac{1}{u_n} - \frac{1}{\ell}\right| = \frac{1}{|u_n|.|\ell|}.|u_n - \ell| \leq \frac{2}{|\ell|^2}|u_n - \ell|.$$

D'où l'on conclut avec la définition de $\lim u_n = \ell$. ■

Passons désormais aux parties réelles et imaginaires des suites complexes (rappelons que l'on a établi plus haut que la suite des modules d'une suite complexe convergente est convergente).

Proposition 6.20.

Une suite $(u_n)_n$ à valeurs complexes est convergente si, et seulement si, les suites réelles $(\mathfrak{Re}(u_n))_n$ et $(\mathfrak{Im}(u_n))_n$ sont convergentes.

Démonstration

($\Rightarrow$) Soit $\ell \in \mathbb{C}$ la limite de la suite $(u_n)_n$. Soit $\varepsilon > 0$; par définition de convergence, il existe $N \in \mathbb{N}$ tel que

$$\forall n \geq N, \qquad |u_n - \ell| \leq \varepsilon.$$

Par conséquent, pour $n \geq N$

$$\begin{aligned}|\mathfrak{Re}(u_n) - \mathfrak{Re}(\ell)| &\leq |u_n - \ell| \leq \varepsilon, \\ |\mathfrak{Im}(u_n) - \mathfrak{Im}(\ell)| &\leq |u_n - \ell| \leq \varepsilon.\end{aligned}$$

Donc les suites $(\mathfrak{Re}(u_n))_n$ et $(\mathfrak{Im}(u_n))_n$ convergent vers $\mathfrak{Re}(\ell)$ et $\mathfrak{Im}(\ell)$ respectivement.

($\Leftarrow$) Soit α et β les limites (réelles) des suites $(\mathfrak{Re}(u_n))_n$ et $(\mathfrak{Im}(u_n))_n$.

Pour tout $\varepsilon > 0$, il existe $N \in \mathbb{N}$ tel que, pour $n \geq N$

$$|\mathfrak{Re}(u_n) - \alpha| \leq \varepsilon, \qquad |\mathfrak{Im}(u_n) - \beta| \leq \varepsilon.$$

Ainsi, pour $n \geq N$

$$|u_n - (\alpha + i\beta)| \leq \sqrt{|\mathfrak{Re}(u_n) - \alpha|^2 + |\mathfrak{Im}(u_n) - \beta|^2} \leq \sqrt{2}.\varepsilon.$$

En conclusion, la suite $(u_n)_n$ converge vers $\alpha + i\beta$. ■

Conseils méthodologiques

D'après cette proposition, on peut donc étudier les suites complexes en les séparant en deux suites réelles : celle des parties réelles d'une part et celle des parties imaginaires d'autre part.

3.3. Exemples

Exemple Suite arithmétique

Une suite arithmétique est convergente si, et seulement si, la raison est nulle ; une suite arithmétique réelle qui ne converge pas diverge vers $+\infty$ ou vers $-\infty$ selon le signe de la raison.

Proposition 6.21. Suite géométrique

Soit $\rho \in \mathbb{C}$. La suite géométrique $(\rho^n)_n$ est convergente si, et seulement si, la raison ρ est égale à 1 ou de module strictement inférieur à 1.

Démonstration

Les cas $\rho \in \{0, 1\}$ correspondent à une suite constante.

▹ Si $|\rho| < 1$, alors la suite converge vers 0 : pour tout $\varepsilon > 0$, il suffit de prendre $N \geq \frac{\ln \varepsilon}{\ln|\rho|}$.

▹ Si $|\rho| > 1$, alors la suite diverge en module vers $+\infty$: pour tout $A > 0$, il suffit de prendre $N \geq \frac{\ln A}{\ln|\rho|}$.

▹ Soit $|\rho| = 1$ et $\rho \neq 1$. Supposons que la suite $(\rho^n)_n$ converge vers ℓ. Comme $\rho^{n+1} = \rho\rho^n$, on obtient en passant à la limite que $\ell = \rho\ell$ d'où $\ell = 0$, ce qui est impossible car les termes de la suite sont de module 1. ■

Exemple Suite arithmético-géométrique

Une suite $(u_n)_n$ vérifiant la relation

$$\forall n \in \mathbb{N}, \qquad u_{n+1} = a\,u_n + b$$

est convergente si, et seulement si, $|a| < 1$ ou $a = 1$ et $b = 0$ ou $u_0 = a\,u_0 + b$.

▹ Le cas $a = 1$ correspond à une suite arithmétique.

▹ Si $a \neq 1$, considérons $\ell \in \mathbb{C}$ tel que $\ell = a\ell + b$. Alors, d'après les calculs déjà effectués en début de chapitre,

$$\forall n \in \mathbb{N}, \qquad u_{n+1} - \ell = a(u_n - \ell).$$

La suite $(u_n - \ell)_n$ est alors géométrique de raison a et on conclut comme précédemment.

Exemple

Pour tout α de module strictement plus petit que 1, une suite $(u_n)_n$ vérifiant la relation de récurrence

$$\forall n \in \mathbb{N}, \qquad u_{n+2} = 2\alpha\,u_{n+1} - \alpha^2\,u_n$$

est de la forme $((An + B)\alpha^n)_n$ avec $A, B \in \mathbb{C}$ (car α est racine double de l'équation caractéristique $x^2 - 2\alpha x + \alpha^2 = 0$). Elle converge vers 0 car $\lim(An + B)\alpha^n = 0$.

3.4. Théorèmes de Cesàro

La démonstration du théorème suivant est relativement formatrice : on étudie la moyenne des termes d'une suite convergente et on montre que cette moyenne converge vers la même limite.

Proposition 6.22. Cesàro

Si la suite $(u_n)_n$ converge vers ℓ, alors la suite $(v_n)_n = \left(\frac{1}{n+1}\sum_{k=0}^{n} u_k\right)_n$ converge vers ℓ.

Remarque

La réciproque est fausse : en effet, la suite $(u_n)_n = ((-1)^n)_n$ est divergente mais converge au sens de Cesàro, c'est-à-dire la suite $\left(\frac{1}{n+1}\sum_{k=0}^{n} u_k\right)_n$ converge.

Démonstration

Soit $\varepsilon > 0$. Écrivons la convergence de la suite $(u_n)_n$: il existe un rang N tel que

$$\forall n \geq N, \qquad |u_n - \ell| \leq \varepsilon.$$

Alors, toujours pour tout $n \geq N$, on sépare la somme définissant u_n à l'indice N et on cherche à contrôler chaque terme indépendamment :

$$\begin{aligned} |v_n - \ell| &\leq \frac{1}{n+1}\sum_{k=0}^{n} |u_k - \ell| = \frac{1}{n+1}\sum_{k=0}^{N} |u_k - \ell| + \frac{1}{n+1}\sum_{k=N+1}^{n} |u_k - \ell| \\ &\leq \frac{1}{n+1}\sum_{k=0}^{N} |u_k - \ell| + \varepsilon. \end{aligned}$$

Or, $\left(\frac{1}{n+1}\sum_{k=0}^{N} |u_k - \ell|\right)_n$ converge vers 0 (produit d'une suite qui tend vers 0 et d'une constante) donc il existe un rang $N' \in \mathbb{N}$ à partir duquel cette suite est inférieure à ε.

Pour tout $n \geq \max\{N, N'\}$, $|v_n - \ell| \leq 2\varepsilon$. ■

Conseils méthodologiques

Dans tous les résultats « à la Cesàro », on sépare la somme en deux :

- on gère les termes correspondant aux grandes valeurs de l'indice par l'hypothèse asymptotique sur $(u_n)_n$;
- on contrôle les premiers termes avec la « normalisation ».

Proposition 6.23. Cesàro

Si la suite $(u_n)_n$ tend vers $+\infty$, alors la suite $(v_n)_n = \left(\frac{1}{n+1}\sum_{k=0}^{n} u_k\right)_n$ tend vers $+\infty$.

Démonstration

Écrivons la divergence de la suite $(u_n)_n$. Soit $A > 0$; il existe un rang N tel que $u_n \geq A$ pour tout $n \geq N$. Alors, pour tout $n \geq N$,

$$\begin{aligned} v_n &= \frac{1}{n+1}\sum_{k=0}^{N} u_k + \frac{1}{n+1}\sum_{k=N+1}^{n} u_k \\ &\geq \frac{1}{n+1}\sum_{k=0}^{N} u_k + \frac{n-N}{n+1} A. \end{aligned}$$

Or, $(\frac{1}{n+1}\sum_{k=0}^{N}|u_k-\ell|)_n$ converge vers 0 (produit d'une suite qui tend vers 0 et d'une constante) donc il existe un rang $N'\in\mathbb{N}$ à partir duquel cette suite est supérieure à -1. De même, il existe un rang $N''\in\mathbb{N}$ à partir duquel $\frac{n-N}{n+1}\geq\frac{1}{2}$. Pour tout $n\geq\max\{N,N',N''\}$, $v_n\geq\frac{1}{2}A-1$. ■

En fait, ces démonstrations se généralisent aisément.

Proposition 6.24.

Soit $(a_n)_n$ une suite à valeurs positives telle que $\sum_{k=0}^{n}a_k\to+\infty$ quand $n\to+\infty$.
Si $(u_n)_n$ converge vers ℓ, alors $\left(\sum_{k=0}^{n}a_ku_k/\sum_{k=0}^{n}a_k\right)_n$ converge vers ℓ.

Remarque

Avec (a_n) la suite constante égale à 1, on retrouve la proposition initiale de Cesàro.

Démonstration

Il suffit de faire le même découpage que précédemment avec $v_n=\frac{\sum_{k=0}^{n}a_ku_k}{\sum_{k=0}^{n}a_k}$

$$\begin{aligned}|v_n-\ell|&\leq\frac{1}{\sum_{k=0}^{n}a_k}\sum_{k=0}^{n}a_k|u_k-\ell|=\frac{1}{\sum_{k=0}^{n}a_k}\sum_{k=0}^{N}a_k|u_k-\ell|+\frac{1}{\sum_{k=0}^{n}a_k}\sum_{k=N+1}^{n}a_k|u_k-\ell|\\&\leq\frac{1}{\sum_{k=0}^{n}a_k}\sum_{k=0}^{N}a_k|u_k-\ell|+\varepsilon.\end{aligned}$$

On conclut à nouveau de la même façon en remarquant que le premier terme est de limite nulle donc est plus petit que ε à partir d'un certain rang. ■

4. Relations de comparaison pour les suites

4.1. Définitions

Définition 6.25.

Soit $(u_n)_n$ et $(v_n)_n$ deux suites ne s'annulant pas à partir d'un certain rang.
▹ La suite $(u_n)_n$ est dominée par $(v_n)_n$ si la suite $\left(\frac{u_n}{v_n}\right)_n$ est bornée. On note $u_n=O(v_n)$.
▹ La suite $(u_n)_n$ est négligeable devant $(v_n)_n$ si la suite $\left(\frac{u_n}{v_n}\right)_n$ est de limite nulle. On note $u_n=o(v_n)$.
▹ La suite $(u_n)_n$ est équivalente à $(v_n)_n$ si la suite $\left(\frac{u_n}{v_n}\right)_n$ est de limite 1. On note $u_n\sim v_n$.

On peut s'affranchir de l'hypothèse les suites ne s'annulent pas à partir d'un certain rang en revenant à la définition de limite et en multipliant par le dénominateur. Cela permet d'avoir des définitions plus générales. Toutefois comme l'objectif ici est d'apprendre à manipuler ces relations de comparaison, nous nous limiterons au cas de suites ne s'annulant pas à partir d'un certain rang dans les démonstrations des propriétés.

Définition 6.26.

Soit $(u_n)_n$ et $(v_n)_n$ deux suites.
▹ La suite $(u_n)_n$ est dominée par $(v_n)_n$ si

$$\exists M > 0,\ \exists N \in \mathbb{N},\ \forall n \geq N, \qquad |u_n| \leq M|v_n|.$$

▹ La suite $(u_n)_n$ est négligeable devant $(v_n)_n$ si

$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N},\ \forall n \geq N, \qquad |u_n| \leq \varepsilon|v_n|.$$

▹ La suite $(u_n)_n$ est équivalente à $(v_n)_n$ si

$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N},\ \forall n \geq N, \qquad |u_n - v_n| \leq \varepsilon|v_n|.$$

On vérifie immédiatement que l'équivalence entre suites est une relation d'équivalence (réflexive, symétrique et transitive).

Exemple

Traduisons les relations de comparaison d'une suite $(u_n)_n$ avec la suite constante égale à $c \neq 0$:

- $u_n = O(c)$ si, et seulement si, $(u_n)_n$ est bornée.
- $u_n = o(c)$ si, et seulement si, $\lim u_n = 0$.
- $u_n \sim c$ si, et seulement si, $\lim u_n = c$.

On a pris soin d'éliminer le cas $c = 0$ qui ne rentre pas dans le cadre de notre définition intuitive. Avec la définition complète, on obtient qu'une suite est équivalente à 0 si, et seulement si, elle est nulle à partir d'un certain rang.

Généralisons cet exemple.

Exemple

Si une suite $(u_n)_n$ tend vers ℓ, alors toutes les suites $(v_n)_n$ qui lui sont équivalentes tendent aussi vers ℓ. Il suffit pour cela de remarquer que $v_n = \frac{v_n}{u_n}.u_n \to 1.\ell = \ell$.

Remarque

Une suite $(u_n)_n$ négligeable devant ou équivalente à $(v_n)_n$ est dominée par $(v_n)_n$. En effet, si la suite $\left(\frac{u_n}{v_n}\right)_n$ converge, alors elle est bornée.

Remarque

Si $(u_n)_n, (v_n)_n$ sont deux suites équivalentes, alors elles sont de même signe à partir d'un certain rang ; en effet, la suite $\left(\frac{u_n}{v_n}\right)_n$ converge vers 1 donc, à partir d'un certain rang, elle est positive.

Proposition 6.27.

Soit $(u_n)_n$, $(v_n)_n$ deux suites. $u_n \sim v_n$ si, et seulement si, $u_n - v_n = o(v_n)$.

Démonstration

Soit $(u_n)_n$ et $(v_n)_n$ deux suites ne s'annulant pas à partir d'un certain rang. Alors,

$$\frac{u_n}{v_n} \to 1 \qquad \Leftrightarrow \qquad \frac{u_n}{v_n} - 1 \to 0 \qquad \Leftrightarrow \qquad \frac{u_n - v_n}{v_n} \to 0.$$

■

Remarque

L'égalité dans les notations o et O a un statut très particulier; en effet, les égalités $u_n = o(w_n)$ et $v_n = o(w_n)$ n'entraînent pas $u_n = v_n$. En toute logique, il faudrait comprendre ce signe = comme un symbole d'appartenance : $u_n = o(w_n)$ signifie que la suite $(u_n)_n$ appartient à l'ensemble des suites négligeables devant la suite $(w_n)_n$.

Toutefois, on utilise quelques propriétés de l'égalité habituelle par commodité d'écriture : ainsi, on écrit $u_n = v_n + o(w_n)$ lorsque l'on veut signifier $u_n - v_n = o(w_n)$.

4.2. Exemples

Exemple

Pour tout $k < l$, $n^k = o(n^l)$ car $\frac{n^k}{n^l} = n^{k-l} \to 0$. Ceci permet d'ordonner les puissances pour la relation de négligeabilité.

Remarque

Ne pas se méprendre lorsque les exposants sont négatifs : par exemple, $\frac{1}{n^2} = o(\frac{1}{n})$ car $-2 < -1$.

Conseils méthodologiques

Pour obtenir un équivalent d'une suite définie comme une somme, il suffit souvent de mettre en facteur le terme *a priori* prépondérant.

Appliquons cette idée aux suites polynomiales ou rationnelles.

Exemple

▹ Considérons par exemple une suite définie comme un polynôme en n : le terme prépondérant, et donc l'équivalent, est alors le monôme de plus haut degré

$$3n^3 + 5n + 2 = 3n^3(1 + \frac{5}{3n^2} + \frac{2}{3n^3}) \sim 3n^3,$$

car $1 + \frac{5}{3n^2} + \frac{2}{3n^3} \to 1$. De même,

$$\frac{5}{n^2} + \frac{1}{n} = \frac{1}{n}\left(\frac{5}{n} + 1\right) \sim \frac{1}{n}.$$

▹ La même idée permet de gérer toutes les suites définies par $P(n)/Q(n)$ où P et Q sont des polynômes de degré p et q est de coefficients dominants c_P et c_Q pour trouver l'équivalent $\frac{c_P}{c_Q} n^{p-q}$.

Voici un autre exemple où l'on exploite cette idée de mettre en facteur le « terme prépondérant ».

Exemple

▹ Montrons que $\ln(n+1) \sim \ln n$. En effet,

$$\ln(n+1) = \ln n + \ln(1 + \frac{1}{n}) = \ln n\left(1 + \frac{\ln(1+\frac{1}{n})}{\ln n}\right) \text{et} 1 + \frac{\ln(1+\frac{1}{n})}{\ln n} \to 1.$$

▹ Montrons que $\ln(\ln(n+1)) \sim \ln(\ln n)$. En effet,

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

et $\ln\left(1+\frac{\ln(1+\frac{1}{n})}{\ln n}\right) \to 0$ alors que $\ln(\ln n) \to +\infty$ donc $\ln\left(1+\frac{\ln(1+\frac{1}{n})}{\ln n}\right) = o(\ln(\ln n))$.

Généralisons cet exemple.

Proposition 6.28.

Soit $(u_n)_n$ et $(v_n)_n$ deux suites équivalentes de limite $+\infty$. Alors $\ln u_n \sim \ln v_n$.

Démonstration

À partir d'un certain rang N, les suites considérées sont à termes strictement positifs. Alors, pour tout $n \geq N$, $\ln u_n - \ln v_n = \ln \frac{u_n}{v_n}$ donc $\ln u_n - \ln v_n \to \ln 1 = 0$ soit encore $\ln u_n = \ln v_n + o(1)$. Comme $(l\,n\,u_n)_n$ diverge vers $+\infty$, on obtient bien $\ln u_n \sim \ln v_n$. ■

Exemple

Montrons que $\sum_{k=0}^{n} k! \sim n!$. En effet,

$$\begin{aligned}\sum_{k=0}^{n} k! - n! = \sum_{k=0}^{n-1} k! &= n! \sum_{k=0}^{n-1} \frac{1}{n(n-1)\dots(k+1)} \\ &\leq n!\left(\frac{1}{n} + \sum_{k=0}^{n-2} \frac{1}{n(n-1)}\right) \\ &\leq n!\left(\frac{1}{n} + \frac{1}{n}\right).\end{aligned}$$

La connaissance des fonctions usuelles permet d'identifier le « terme prépondérant ».

Proposition 6.29. Croissances comparées

Pour tout $r > 0$,

$$\ln n = o(n^r), \qquad n^r = o(e^n).$$

Exemple

$$e^n + n^5 \ln n + 2n^{13} = e^n\left(1 + \frac{n^5 \ln n}{e^n} + \frac{2n^{13}}{e^n}\right) \sim e^n,$$

car $\frac{n^5 \ln n}{e^n} + \frac{2n^{13}}{e^n} \to 0$ d'après les croissances comparées.

Admettons temporairement la formule de Stirling qui permet de positionner la factorielle par rapport aux suites usuelles exponentielles, puissances et logarithmes.

Proposition 6.30. Formule de Stirling

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

D'autres équivalents peuvent aussi être obtenus avec des moyens relativement simples.

Exemple

Si $(u_n)_n$ est une suite de limite nulle et $p \in \mathbb{N}$, alors

$$\frac{1}{1-u_n} = 1 + u_n + \ldots + u_n^p + o(u_n^p).$$

En effet, d'après les calculs sur les sommes géométriques

$$\frac{1}{1-u_n} - (1 + u_n + \ldots + u_n^p) = \frac{1}{1-u_n} - \frac{1-u_n^{p+1}}{1-u_n} = \frac{u_n^{p+1}}{1-u_n} = u_n^p \frac{u_n}{1-u_n}$$

d'où le résultat en remarquant que $\frac{u_n}{1-u_n} \to 0$.

4.3. Propriétés opératoires

Proposition 6.31.

Soit $(u_n)_n$, $(v_n)_n$ et $(w_n)_n$ trois suites telles que $u_n = o(w_n)$ et $v_n = o(w_n)$ et $\lambda \in \mathbb{R}$. Alors, $u_n + \lambda v_n = o(w_n)$.

Démonstration

Soit $(u_n)_n, (v_n)_n$ et $(w_n)_n$ des suites ne s'annulant pas à partir d'un certain rang telles que $\frac{u_n}{w_n} \to 0$ et $\frac{v_n}{w_n} \to 0$. Alors,

$$\frac{u_n + \lambda v_n}{w_n} = \frac{u_n}{w_n} + \lambda \frac{v_n}{w_n} \to 0.$$

■

Proposition 6.32.

Soit $(u_n)_n$, $(u'_n)_n$, $(v_n)_n$ des suites telles que $u_n = o(u'_n)$. Alors, $u_n v_n = o(u'_n v_n)$.

Démonstration

Soit $(u_n)_n$, $(v_n)_n$ et $(u'_n)_n$ des suites ne s'annulant pas à partir d'un certain rang telles que $\frac{u_n}{u'_n} \to 0$. Alors,

$$\frac{u_n v_n}{u'_n v_n} = \frac{u_n}{u'_n} \to 0.$$

■

Exemple

Soit $(u_n)_n$ une suite telle que $u_n = o(\frac{1}{n})$. Alors, $2n u_n = o(\frac{2n}{n})$ donc $2n u_n = o(2)$ et, plus simplement, $2n u_n = o(1)$.

Proposition 6.33.

Soit $(u_n)_n$, $(u'_n)_n$, $(v_n)_n$ et $(v'_n)_n$ des suites telles que $u_n \sim u'_n$ et $v_n \sim v'_n$. Alors, $u_n v_n \sim u'_n v'_n$ et $\frac{u_n}{v_n} \sim \frac{u'_n}{v'_n}$.

Démonstration

Soit $(u_n)_n, (v_n)_n, (u'_n)_n$ et $(v'_n)_n$ des suites ne s'annulant pas à partir d'un certain rang telles que $\frac{u_n}{u'_n} \to 1$ et $\frac{v_n}{v'_n} \to 1$. Alors,

$$\frac{u_n v_n}{u'_n v'_n} = \frac{u_n}{u'_n}\frac{v_n}{v'_n} \to 1, \qquad \frac{u_n v'_n}{u'_n v_n} = \frac{u_n}{u'_n}\frac{v'_n}{v_n} \to 1.$$

■

Remarque

Cette proposition indique que les équivalents peuvent se multiplier (ou se diviser). En revanche, on ne peut *a priori* additionner les équivalents $n \sim n+1$ et $-n \sim -n$ mais $0 \not\sim 1$, ni les composer $n \sim n+1$ et pourtant $\exp(n) \not\sim \exp(n+1)$.

Exemple

Cherchons un équivalent simple au coefficient binomial $\binom{2n}{n}$ à l'aide de la formule de Stirling.

$$\binom{2n}{n} = \frac{(2n)!}{n!^2} \sim \frac{\sqrt{4\pi n}2^{2n} n^{2n} e^{-2n}}{2\pi n n^{2n} e^{-2n}} \sim \frac{2^{2n}}{\sqrt{\pi n}}.$$

Exemple

Étudions la suite $(u_n)_n$ avec

$$u_n = \frac{2n^2+1}{n-1}.$$

Alors, le numérateur est équivalent à $2n^2$ et le dénominateur est équivalent à n donc $u_n \sim 2n$.
Poursuivons l'étude avec les termes suivants,

$$u_n - 2n = \frac{2n+1}{n-1} \sim \frac{2n}{n} = 2$$

puis

$$u_n - 2n - 2 = \frac{3}{n-1} \sim \frac{3}{n}.$$

Ainsi, $u_n - 2n - 2 = \frac{3}{n} + o(\frac{3}{n})$ donc $u_n - 2n - 2 = \frac{3}{n} + o(\frac{1}{n})$.
En conclusion,

$$u_n = 2n + 2 + \frac{3}{n} + o(\frac{1}{n}).$$

Poursuivons cet exemple avec la suite $(v_n)_n$ définie par $v_n = \frac{(2n^2+1)(n^2+1)}{(n-1)^2}$

$$v_n = \frac{2n^2+1}{n-1} \times \frac{n^2+1}{n-1} = \left(2n+2+\frac{3}{n}+o\left(\frac{1}{n}\right)\right)\left(n+1+\frac{2}{n}+o\left(\frac{1}{n}\right)\right) = 2n^2+4n+9+o(1).$$

Tous les termes n'étant pas écrits sont négligeables par rapport à 1 donc incorporés au terme $o(1)$.

5. Suites réelles

Pour les suites réelles, on peut exploiter l'existence de la relation d'ordre $\leq$: encadrements, variations et monotonies...

5.1. Inégalités

Proposition 6.34. Lemme d'encadrement

Soit $(u_n)_n$, $(v_n)_n$ et $(w_n)_n$ trois suites réelles telles que
- à partir d'un certain rang N, $u_n \leq v_n \leq w_n$;
- les suites $(u_n)_n$ et $(w_n)_n$ convergent vers la même limite ℓ.

Alors, la suite $(v_n)_n$ converge et sa limite est ℓ.

Démonstration

Soit $\varepsilon > 0$. D'après la définition des convergences des suites $(u_n)_n$ et $(w_n)_n$, il existe des rangs N' et N'' tels que

$$\forall n \geq N', \qquad u_n \geq \ell - \varepsilon,$$
$$\forall n \geq N'', \qquad w_n \leq \ell + \varepsilon.$$

Ainsi, pour $n \geq \max\{N, N', N''\}$,

$$\ell - \varepsilon \leq u_n \leq v_n \leq w_n \leq \ell + \varepsilon$$

donc $|v_n - \ell| \leq \varepsilon$. En conclusion, $\lim v_n = \ell$. ■

Exemple

Étudions la convergence de la suite $(v_n)_n$ définie par $v_n = \sum_{k=1}^{n} \frac{n}{n^2+k}$. Chacun des termes de la somme définissant v_n est compris entre $\frac{n}{n^2+n}$ et $\frac{n}{n^2+1}$, d'où

$$n\frac{n}{n^2+n} \leq v_n \leq n\frac{n}{n^2+1}.$$

Or, les deux termes extrêmes de cet encadrement convergent vers 1 donc $v_n \to 1$.

Remarque

Soit $(u_n)_n$ et $(v_n)_n$ deux suites réelles telles qu'à partir d'un certain rang $|u_n| \leq v_n$ et $\lim v_n = 0$. Alors, $\lim u_n = 0$ car on peut appliquer le lemme d'encadrement aux inégalités $-v_n \leq u_n \leq v_n$.

Proposition 6.35.

Soit $(u_n)_n$ et $(v_n)_n$ deux suites réelles telles qu'à partir d'un certain rang $u_n \leq v_n$.
- ▹ Si les suites $(u_n)_n$ et $(v_n)_n$ convergent respectivement vers ℓ et ℓ', alors $\ell \leq \ell'$.
- ▹ Si la suite $(u_n)_n$ diverge vers $+\infty$, alors $(v_n)_n$ diverge aussi vers $+\infty$.
- ▹ Si la suite $(v_n)_n$ diverge vers $-\infty$, alors $(u_n)_n$ diverge aussi vers $-\infty$.

Remarque

Pour deux suites réelles $(u_n)_n$ et $(v_n)_n$ qui convergent respectivement vers ℓ et ℓ' et telles qu'à partir d'un certain rang $u_n < v_n$, l'inégalité entre les limites n'est pas nécessairement stricte : par exemple, avec $u_n = 0$ et $v_n = \frac{1}{n}$, les limites sont égales alors que l'on dispose de l'inégalité stricte $u_n < v_n$ pour tout indice $n \in \mathbb{N}$.

Démonstration

▹ Pour tout $\varepsilon > 0$, il existe un rang $N \in \mathbb{N}$ tel que pour $n \geq N$

$$\ell - \varepsilon \leq u_n \leq v_n \leq \ell' + \varepsilon.$$

Par conséquent, $\ell \leq \ell' + 2\varepsilon$ pour tout $\varepsilon > 0$ donc $\ell \leq \ell'$.

▹ Pour tout $A > 0$, il existe un rang $N \in \mathbb{N}$ tel que pour $n \geq N$, $v_n \geq u_n \geq A$, d'où le résultat.

▹ Il suffit d'appliquer le point précédent avec $-v_n \leq -u_n$. ■

Exemple

Étudions la suite $(v_n)_n$ définie par $v_n = \sum_{k=1}^{n} \frac{1}{k}$.

▹ Pour tout $t \in [k, k+1]$, $\frac{1}{t} \leq \frac{1}{k}$ donc $\int_k^{k+1} \frac{1}{t}\,\mathrm{d}t \leq \frac{1}{k}$ et, avec la relation de Chasles pour les intégrales,

$$v_n \geq \sum_{k=1}^{n} \int_k^{k+1} \frac{1}{t}\,\mathrm{d}t = \int_1^{n+1} \frac{1}{t}\,\mathrm{d}t = \ln(n+1).$$

Ainsi, $v_n \to +\infty$.

▹ Pour tout $t \in [k, k+1]$, $\frac{1}{t} \leq \frac{1}{k}$ donc $\int_k^{k+1} \frac{1}{t}\,\mathrm{d}t \leq \frac{1}{k} \leq \int_{k-1}^{k} \frac{1}{t}\,\mathrm{d}t$ et, avec la relation de Chasles pour les intégrales,

$$\ln(n+1) \leq v_n \leq 1 + \ln(n).$$

Ainsi, par encadrement, $\lim \frac{v_n}{\ln(n)} = 1$, c'est-à-dire $v_n \sim \ln n$.

Exemple Critère de D'Alembert

Soit une suite $(u_n)_n \in \mathbb{C}^{\mathbb{N}}$ ne s'annulant pas à partir d'un certain rang telle que $\left|\frac{u_{n+1}}{u_n}\right| \to \ell$.

▹ Si $\ell < 1$, alors la suite $(u_n)_n$ converge vers 0. En effet, il existe $N \in \mathbb{N}$ tel que

$$\forall n \geq N, \qquad \left|\frac{u_{n+1}}{u_n}\right| \leq \frac{1+\ell}{2}.$$

Ainsi, pour tout $n \geq N$, $|u_n| \leq \left(\frac{1+\ell}{2}\right)^{n-N} |u_N|$ donc par encadrement et comparaison à une suite géométrique de raison $\frac{1+\ell}{2} \in]0, 1[$, $u_n \to 0$.

▹ Si $\ell > 1$, alors la suite $(|u_n|)_n$ diverge vers $+\infty$. En effet, il existe $N \in \mathbb{N}$ tel que

$$\forall n \geq N, \qquad \left|\frac{u_{n+1}}{u_n}\right| \geq \frac{1+\ell}{2}.$$

Ainsi, pour tout $n \geq N$, $|u_n| \geq \left(\frac{1+\ell}{2}\right)^{n-N} |u_N|$ donc par comparaison à une suite géométrique de raison $\frac{1+\ell}{2} > 1$, $u_n \to +\infty$.

5.2. Écriture décimale

Proposition 6.36.

Pour tout $x \in \mathbb{R}$, il existe un unique entier $n \in \mathbb{Z}$ tel que $n \leq x < n+1$. Cet entier n est noté $\lfloor x \rfloor$ et appelé la partie entière de x.

Démonstration

La partie $A = \{m \in \mathbb{Z},\ m \leq x\}$ est non vide, majorée donc admet une borne supérieure : c'est l'entier n de l'énoncé. ■

Exemple

On obtient immédiatement $\lfloor 3,14 \rfloor = 3$ et $\lfloor -3,14 \rfloor = -4$.

Remarque

Attention à ne pas confondre la partie entière avec la troncature de la partie décimale.

Remarque

Cette propriété permet, par exemple, de prouver le caractère archimédien de $\mathbb{R}$: pour tout $y \in \mathbb{R}$ et tout $x > 0$, il existe $n \in \mathbb{N}$ tel que $nx > y$.

Proposition 6.37. Écriture décimale

Pour tout $x \in \mathbb{R}_+$, il existe une unique suite d'entiers naturels $(x_n)_n$ telle que

$$\forall n \in \mathbb{N}, \qquad \sum_{k=0}^{n} x_k 10^{-k} \leq x < \sum_{k=0}^{n} x_k 10^{-k} + 10^{-n}.$$

De plus, la suite $(x_n)_n$ est à valeurs dans $[\![0,9]\!]$ pour $n \geq 1$ et n'est pas constante égale à 9 à partir d'un certain rang.

Démonstration

▹ Montrons l'existence et l'unicité de la suite par récurrence.

- Le terme x_0 doit vérifier $x_0 \leq x < x_0 + 1$ donc $x_0 = \lfloor x \rfloor$.
- Soit $n \in \mathbb{N}$. Supposons $x_0, \ldots, x_n$ uniquement déterminés. Alors, le terme x_{n+1} doit vérifier

$$\sum_{k=0}^{n} x_k 10^{n+1-k} + x_{n+1} \leq 10^{n+1} x < \sum_{k=0}^{n} x_k 10^{n+1-k} + x_{n+1} + 1,$$

d'où $x_{n+1} = \left\lfloor 10^{n+1} x - \sum_{k=0}^{n} x_k 10^{n+1-k} \right\rfloor$ et il existe une unique valeur pour x_{n+1}.

▹ Pour tout $n \geq 0$, $x_{n+1} = \left\lfloor 10^{n+1} x - \sum_{k=0}^{n} x_k 10^{n+1-k} \right\rfloor$ et $0 \leq 10^n x - \sum_{k=0}^{n} x_k 10^{n-k} < 1$ donc

$$0 \leq 10^{n+1} x - \sum_{k=0}^{n} x_k 10^{n+1-k} < 10.$$

Ainsi $x_{n+1} \in [\![0,9]\!]$.

▹ Si la suite $(x_n)_n$ est constante égale à 9 à partir du rang N, alors, pour tout $n \geq N$,

$$\sum_{k=0}^{N} x_k 10^{-k} + 10^{-N} - 10^{-n} = \sum_{k=0}^{N} x_k 10^{-k} + \sum_{k=N+1}^{n} 9.10^{-k} \leq x,$$

et

$$x < \sum_{k=0}^{N} x_k 10^{-k} + \sum_{k=N+1}^{n} 9.10^{-k} + 10^{-n} = \sum_{k=0}^{N} x_k 10^{-k} + 10^{-N}.$$

En passant à la limite dans la première inégalité, on obtient la contradiction

$$\sum_{k=0}^{N} x_k 10^{-k} + 10^{-N} \leq x < \sum_{k=0}^{N} x_k 10^{-k} + 10^{-N}.$$

■

Remarque

La suite $(x_n)_n$ est l'écriture décimale propre de x :

$$x = x_0, x_1 x_2 x_3 \dots$$

Les valeurs $\sum_{k=0}^{n} x_k 10^{-k}$ et $\sum_{k=0}^{n} x_k 10^{-k} + 10^{-n}$ sont les approximations décimales à l'ordre n de x par défaut et par excès.

Pour obtenir l'écriture décimale de $x \in \mathbb{R}_-$, on se ramène à l'écriture décimale de $-x$.

Définition 6.38.

Les nombres décimaux sont les réels dont la suite des décimales est constante égale à 0 à partir d'un certain rang.

Remarque

En informatique, la place mémoire est finie (bornée par les spécificités de la machine) : il n'y a donc pas de réels. On les remplace en général par des décimaux, appelés flottants, et ils sont codés par deux entiers. Plus précisément, un décimal x s'écrit $\pm a 10^n$ avec $a \in \mathbb{N}$ (la mantisse) et $n \in \mathbb{Z}$ (l'exposant).

5.3. Suites monotones

Définition 6.39.

Soit $(u_n)_n$ une suite réelle.
▹ La suite $(u_n)_n$ est croissante si, pour tout $n \in \mathbb{N}$, $u_{n+1} \geq u_n$.
▹ La suite $(u_n)_n$ est décroissante si, pour tout $n \in \mathbb{N}$, $u_{n+1} \leq u_n$.
▹ La suite $(u_n)_n$ est monotone si elle est croissante ou décroissante. Si les inégalités sont strictes, la suite est strictement croissante, strictement décroissante et strictement monotone respectivement.

Rappelons qu'il existe des suites non monotones comme $((-1)^n)_n$.

Exemple

Une suite réelle $(u_n)_n$ vérifiant la relation de récurrence

$$\forall n \in \mathbb{N}, \qquad u_{n+1} = u_n + u_n^4$$

est croissante puisque, pour tout $n \in \mathbb{N}$, $u_{n+1} - u_n = u_n^4 \geq 0$.

Exemple

Considérons une suite $(u_n)_n$ vérifiant $u_{n+1} = f(u_n)$ pour tout $n \in \mathbb{N}$.
▹ Si $f(x) \geq x$ pour tout $x \in \mathbb{R}$, la suite $(u_n)_n$ est croissante car pour tout $n \in \mathbb{N}$, $u_{n+1} = f(u_n) \geq u_n$.
▹ Si $f(x) \leq x$ pour tout $x \in \mathbb{R}$, la suite $(u_n)_n$ est décroissante.
▹ Si f est croissante et $u_1 \geq u_0$ (respectivement $u_1 \leq u_0$), la suite $(u_n)_n$ est croissante (respectivement décroissante) : en effet, on montre rapidement que $u_{n+1} \geq u_n$ par récurrence sur n.
▹ Si f est décroissante et $u_2 \geq u_0$ (respectivement $u_2 \leq u_0$), les suites $(u_{2n})_n$ et $(u_{2n+1})_n$ sont croissante et décroissante (respectivement décroissante et croissante).

Rappelons la propriété de la borne supérieure du corps des réels qui provient de la construction du corps des réels (admise dans cet ouvrage).

Proposition 6.40.

Toute partie non vide majorée de $\mathbb{R}$ admet une borne supérieure.

Le résultat suivant s'avère une conséquence immédiate de cette propriété.

Proposition 6.41. Limite monotone

▹ Une suite croissante majorée converge.
▹ Une suite croissante non majorée diverge vers $+\infty$.

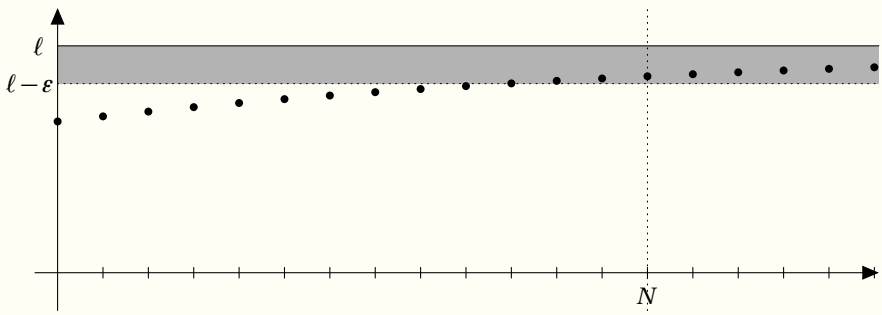
Démonstration

Soit $(u_n)_n$ une suite croissante.

▹ Si $(u_n)_n$ est majorée, considérons $\ell = \sup\{u_n,\ n \in \mathbb{N}\}$. Par définition de la borne supérieure, pour tout $\varepsilon > 0$, il existe $N \in \mathbb{N}$ tel que $\ell - \varepsilon < u_N \leq \ell$ (car $\ell - \varepsilon$ est strictement plus petit que le plus petit des majorants). Comme la suite est croissante,

$$\forall n \geq N, \qquad \ell - \varepsilon < u_N \leq u_n \leq \ell.$$

Ainsi, la suite $(u_n)_n$ converge vers ℓ.



▹ Si $(u_n)_n$ n'est pas majorée, alors, pour tout $A > 0$, il existe $N \in \mathbb{N}$ tel que $u_N \geq A$. Comme la suite est croissante,

$$\forall n \geq N, \qquad u_n \geq u_N \geq A.$$

Ainsi, la suite $(u_n)_n$ diverge vers $+\infty$. ■

Remarque

En adaptant cette démonstration, on voit que la borne supérieure d'une partie X est limite d'une suite d'éléments de X. Ceci est relativement utile et mérite d'être retenu.

On a bien évidemment un résultat analogue pour les suites minorées décroissantes, pour les suites bornées et monotones à partir d'un certain rang.

Proposition 6.42. Suites adjacentes

Soit deux suites réelles $(a_n)_n$ et $(b_n)_n$ telles que
- la suite $(a_n)_n$ est croissante ;
- la suite $(b_n)_n$ est décroissante ;
- $b_n - a_n \to 0$.

Alors, les suites $(a_n)_n$ et $(b_n)_n$ convergent vers la même limite.

Démonstration

Remarquons tout d'abord que la suite $(b_n - a_n)_n$ est décroissante car $(b_n)_n$ et $(-a_n)_n$ sont toutes deux décroissantes. Par ailleurs, cette suite est de limite nulle donc est à termes positifs ; ainsi, pour tout n, $a_n \leq b_n \leq b_0$ et donc la suite $(a_n)_n$ est majorée par b_0. De même, la suite $(b_n)_n$ est minorée par a_0. D'après les propriétés de monotonie, ces deux suites sont convergentes.

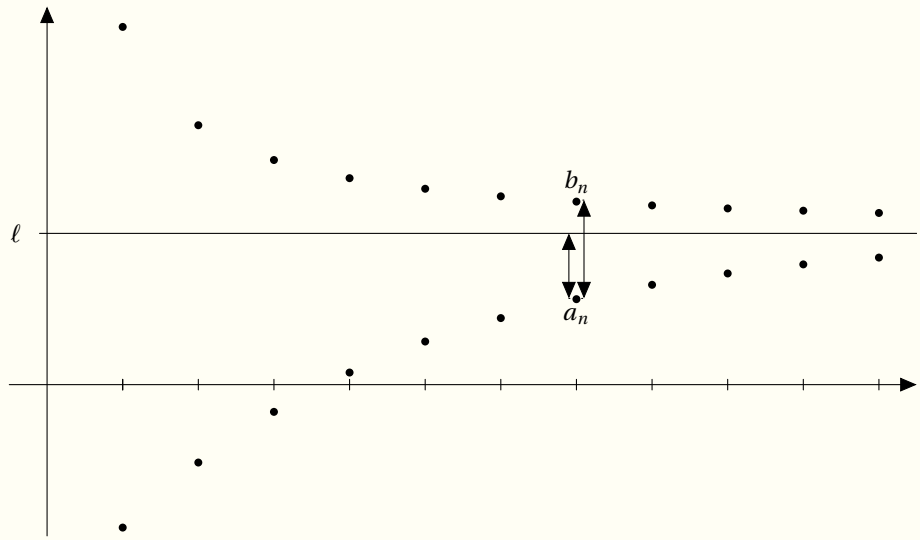
De plus,

$$\lim b_n - \lim a_n = \lim(b_n - a_n) = 0.$$

D'où $\lim b_n = \lim a_n$. ■

Remarque

Pour deux suites réelles $(a_n)_n$ et $(b_n)_n$ vérifiant les hypothèses de la proposition précédente, dites adjacentes, de limites ℓ, on a toujours $0 \leq \ell - a_n \leq b_n - a_n$ et $0 \leq b_n - \ell \leq b_n - a_n$.



Exemple Moyenne arithmético-géométrique

Soit $a, b > 0$. Définissons les suites $(a_n)_n$ et $(b_n)_n$ par $a_0 = a$ et $b_0 = b$ et les relations de récurrence

$$\forall n \in \mathbb{N}, \qquad \begin{cases} a_{n+1} & = & \sqrt{a_n b_n} \\ b_{n+1} & = & \frac{a_n + b_n}{2}. \end{cases}$$

▹ Commençons par remarquer que, pour tout $n \in \mathbb{N}$,

$$\begin{aligned} b_{n+1} - a_{n+1} &= \frac{1}{2}\left(a_n - 2\sqrt{a_n b_n} + b_n\right) \\ &= \frac{1}{2}\left(\sqrt{a_n}^2 - 2\sqrt{a_n}\sqrt{b_n} + \sqrt{b_n}^2\right) \\ &= \frac{1}{2}\left(\sqrt{b_n} - \sqrt{a_n}\right)^2 \geq 0. \end{aligned}$$

Ainsi, pour tout $n \geq 0$, $b_{n+1} \geq a_{n+1}$, c'est-à-dire pour tout $n \geq 1$, $b_n \geq a_n$.

▹ Pour tout $n \geq 1$, $b_{n+1} - b_n = \frac{a_n - b_n}{2} \leq 0$. Ainsi, la suite $(b_n)_n$ est décroissante à partir du rang 1.

▹ Pour tout $n \geq 1$, $\frac{a_{n+1}}{a_n} = \sqrt{\frac{b_n}{a_n}} \geq 1$. Ainsi, la suite $(a_n)_n$ est croissante à partir du rang 1.

▹ Les suites $(a_n)_n$ et $(b_n)_n$ convergent (monotones bornées) vers ℓ et ℓ'. En passant à la limite dans l'une des relations, on trouve $\ell = \ell'$.

Exemple Développement en série de e

Soit les suites $(a_n)_n$ et $(b_n)_n$ définies, pour tout $n \in \mathbb{N}$, par

$$a_n = \sum_{k=0}^{n} \frac{1}{k!}, \qquad b_n = \sum_{k=0}^{n} \frac{1}{k!} + \frac{1}{n.n!}.$$

▹ La suite $(a_n)_n$ est clairement strictement croissante.
▹ La suite $(b_n)_n$ est strictement décroissante car, pour tout $n \in \mathbb{N}$,

$$\begin{aligned} b_{n+1} - b_n &= \sum_{k=0}^{n+1} \frac{1}{k!} + \frac{1}{(n+1).(n+1)!} - \sum_{k=0}^{n} \frac{1}{k!} - \frac{1}{n.n!} \\ &= \frac{1}{(n+1)!} + \frac{1}{(n+1).(n+1)!} - \frac{1}{n.n!} = -\frac{1}{n(n+1).(n+1)!} < 0. \end{aligned}$$

▹ La suite différence vérifie $b_n - a_n = \frac{1}{n.n!} \to 0$.
Ainsi, d'après la proposition précédente, les suites $(a_n)_n$ et $(b_n)_n$ convergent vers une même limite (qui est égale à $e = \exp(1)$).

Remarque

Remarquons que ce dernier exemple permet en particulier de montrer que cette limite e n'est pas rationnelle. En effet, si $e = \frac{p}{q}$ avec $(p,q) \in \mathbb{Z} \times \mathbb{N} \setminus \{0\}$, alors $a_q < e < a_q + \frac{1}{q.q!}$. Ainsi,

$$q!a_q < q!\frac{p}{q} < q!a_q + \frac{1}{q} < q!a_q + 1$$

et l'entier $q!\frac{p}{q}$ est strictement inclus entre deux entiers consécutifs $q!a_q$ et $q!a_q + 1$: contradiction.

5.4. Exemples de parties denses de $\mathbb{R}$

Définition 6.43.

Une partie $A \subset \mathbb{R}$ est dense dans $\mathbb{R}$ si elle vérifie l'une des deux propriétés suivantes
- tout réel est limite d'une suite à valeurs dans A ;
- $\forall x \in \mathbb{R},\ \forall \varepsilon > 0,\ A \cap]x-\varepsilon, x+\varepsilon[\neq \varnothing$.

Montrons que les deux affirmations de la définition sont équivalentes.

Démonstration

▹ Si x est limite d'une suite d'éléments de A, alors l'ensemble $A \cap]x-\varepsilon, x+\varepsilon[$ contient les éléments de la suite à partir d'un certain rang.
▹ Réciproquement, pour construire le $n^{\text{ième}}$ terme d'une suite à valeur dans A convergeant vers x, il suffit de prendre un élément u_n de l'ensemble (supposé non vide) $A \cap]x-2^{-n}, x+2^{-n}[$. Par construction, pour tout n, $|u_n - x| \leq 2^{-n}$ donc, par le lemme d'encadrement, $\lim u_n = x$. ■

Proposition 6.44.

Les ensembles $\mathbb{Q}$ et $\mathbb{R} \setminus \mathbb{Q}$ sont denses dans $\mathbb{R}$.

Démonstration

▹ Pour $\mathbb{Q}$, il s'agit de la proposition concernant l'écriture décimale des réels.

▹ Si $(u_n)_n$ est une suite de rationnels convergeant vers x, alors $(u_n + \frac{\sqrt{2}}{n})_n$ est une suite d'irrationnels convergeant vers x. ■

En fait, la même proposition sur l'écriture décimale donne

Proposition 6.45.

L'ensemble des nombres décimaux est dense dans $\mathbb{R}$.

Citons d'autres parties denses de $\mathbb{R}$ bien utiles.

Proposition 6.46. Sous-groupes de $\mathbb{R}$

Une partie non vide de $\mathbb{R}$ stable par addition et différence (sous-groupe de $(\mathbb{R},+)$) est de la forme $a\mathbb{Z} = \{ka,\ k \in \mathbb{Z}\}$ avec $a \in \mathbb{R}$ ou est dense.

Démonstration

Soit G une partie non vide de $\mathbb{R}$ stable par addition et différence. Si $G = \{0\}$, le résultat est évident. Sinon, G contient des éléments strictement positifs (car il y a au moins un élément non nul et son opposé). Notons $a = \inf G \cap \mathbb{R}_+^*$.

▹ Si $a \neq 0$, alors a est le plus petit élément $G \cap \mathbb{R}_+^*$: sinon, il existe un élément $x \in G$ entre $]a, 2a[$ (car $2a$ n'est pas un minorant de $G \cap \mathbb{R}_+^*$), puis un élément $y \in G$ dans $]a, x[$ (car x n'est pas un minorant de $G \cap \mathbb{R}_+^*$). Alors, $x - y \in G \cap]0, a[$ ce qui contredit la définition de a.

Ainsi, $a\mathbb{Z} \subset G$. Réciproquement, si $x \in G$, alors il existe $n = \left\lfloor \frac{x}{a} \right\rfloor \in \mathbb{Z}$ et $0 \leq y < a$ tel que $x = na + y$; ainsi, $y = x - na \in G \cap [0, a[$ donc $y = 0$ et $x \in a\mathbb{Z}$. Par double-inclusion, $G = a\mathbb{Z}$.

▹ Si $a = 0$, alors pour tout $\varepsilon > 0$, il existe $x \in G \cap [0, \varepsilon]$. Ensuite, pour tout $y \in \mathbb{R}$, il existe $n \in \mathbb{Z}$ tel que $0 \leq y - nx < x \leq \varepsilon$ (n est la partie entière de $\frac{y}{x}$) donc $G \cap]y - \varepsilon, y + \varepsilon[\neq \varnothing$: G est dense dans $\mathbb{R}$. ■

On retrouve que $\mathbb{Q}$ est dense dans $\mathbb{R}$ car c'est une partie non vide stable par addition et différence qui n'est pas de la forme $a\mathbb{Z}$.

Exemple

L'ensemble $\mathbb{Z} + \sqrt{2}\mathbb{Z}$ est stable par addition et différence mais n'est pas de la forme $a\mathbb{Z}$ donc est dense ; en effet, s'il existait $a \in \mathbb{R}$ tel que $\mathbb{Z} + \sqrt{2}\mathbb{Z} = a\mathbb{Z}$, alors il existerait $p, q \in \mathbb{Z}^*$ tels que $\sqrt{2} = ap$ et $1 = aq$ d'où $\sqrt{2} = \frac{p}{q}$ ce qui est impossible car $\sqrt{2} \notin \mathbb{Q}$.

Plus généralement, si on considère $x,\ y \in \mathbb{R}$ tels que $\frac{x}{y} \notin \mathbb{Q}$, alors $x\mathbb{Z} + y\mathbb{Z}$ est dense dans $\mathbb{R}$.

6. Suites extraites

6.1. Définition

Définition 6.47.

Soit $(u_n)_n$ une suite à valeurs dans $\mathbb{K}$. Une suite extraite de $(u_n)_n$ est une suite $(u_{\varphi(n)})_n$ où l'application $\varphi : \mathbb{N} \to \mathbb{N}$ est strictement croissante.

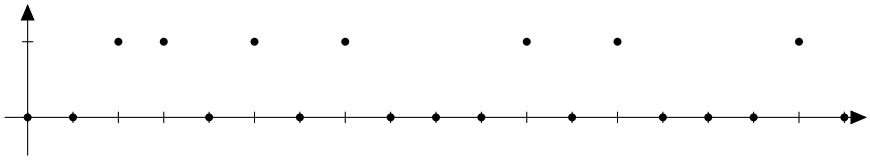
Une suite extraite est donc obtenue en ne conservant qu'une partie des termes de la suite de départ (ceux dont les indices sont atteints par l'application φ).

Exemple

Parmi les exemples les plus simples, remarquons que $(u_{n+1})_n$, $(u_{2n})_n$ et $(u_{2n+1})_n$ sont des suites extraites de $(u_n)_n$.

Exemple

Considérons la suite (u_n) qui vaut 1 lorsque l'indice est premier, 0 sinon. Il est élémentaire que cette suite diverge.



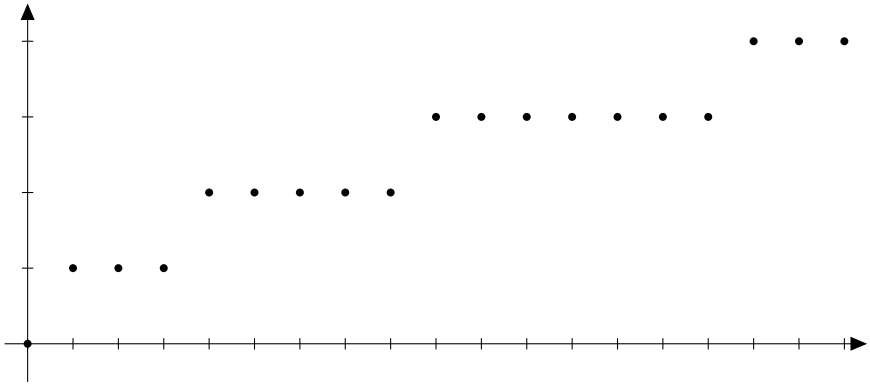
Pour tout $a \geq 2$, la suite extraite (u_{an}) est nulle à partir du rang 2 donc converge.

Remarque

On a quelquefois besoin d'extraire une suite d'une suite déjà extraite. Il est important de savoir que $(u_{\varphi\circ\psi(n)})_n$ est une suite extraite de $(u_{\varphi(n)})_n$ (alors que la plupart du temps, un terme de $(u_{\psi\circ\varphi(n)})_n$ n'est pas de la forme $u_{\varphi(n)}$).

Exemple

Considérons la suite (u_n) définie, pour tout n, par $u_n = \lfloor\sqrt{n}\rfloor$.



La suite extraite (u_{n^2+2n}) est strictement croissante, puisque, pour tout n, $u_{n^2+2n} = n$.

6.2. Propriétés

Proposition 6.48.

Une suite extraite d'une suite convergente est convergente vers la même limite.

Conseils méthodologiques

Cette proposition sert essentiellement à montrer qu'une suite n'est pas convergente en exhibant deux suites extraites de limites différentes.

Démonstration

Soit $(u_n)_n$ une suite de limite ℓ et $(u_{\varphi(n)})_n$ une suite extraite.

Soit $\varepsilon > 0$; il existe $N \in \mathbb{N}$ tel que pour tout $n \geq N$, $|u_n - \ell| \leq \varepsilon$; comme $\varphi(n) \geq n$ (par une récurrence immédiate), pour tout $n \geq N$, $|u_{\varphi(n)} - \ell| \leq \varepsilon$.

■

En fait, la réciproque est élémentaire : une suite converge si, et seulement si, toutes ses suites extraites convergent (car la suite est l'une de ses propres suites extraites).

Exemple

Soit $(u_n)_n$ une suite telle que les suites $(u_{2n})_n$ et $(u_{2n+1})_n$ convergent vers la même limite ; montrons que la suite $(u_n)_n$ converge.
Soit $\varepsilon > 0$; d'après les définitions de convergence, il existe des rangs N et N'

$$\forall n \geq N, \qquad |u_{2n} - \ell| \leq \varepsilon,$$

$$\forall n \geq N', \qquad |u_{2n+1} - \ell| \leq \varepsilon.$$

Ainsi, pour tout $n \geq \max\{2N, 2N'+1\}$, $|u_n - \ell| \leq \varepsilon$, donc $u_n \to \ell$.

Exemple

Soit $(u_n)_n$ une suite telle que les suites $(u_{2n})_n$, $(u_{2n+1})_n$ et $(u_{3n})_n$ convergent ; montrons que la suite $(u_n)_n$ converge en se ramenant au cas précédent : la suite $(u_{6n})_n$ est extraite à la fois de $(u_{2n})_n$ et $(u_{3n})_n$ donc ces deux suites ont la même limite ; de même, la suite $(u_{6n+3})_n$ est extraite à la fois de $(u_{2n+1})_n$ et $(u_{3n})_n$ donc ces deux suites ont la même limite. En conclusion, les suites $(u_{2n})_n$ et $(u_{2n+1})_n$ ont la même limite.

On peut obtenir une caractérisation des limites des suites extraites.

Définition 6.49.

Soit $(u_n)_n \in \mathbb{K}^{\mathbb{N}}$ et $\ell \in \mathbb{K}$. Les trois propriétés suivantes sont équivalentes.

1. Il existe une suite extraite de $(u_n)_n$ qui converge vers ℓ.
2. L'ensemble $\{n \in \mathbb{N},\ |u_n - \ell| \leq \varepsilon\}$ est infini pour tout $\varepsilon > 0$.
3. $\forall \varepsilon > 0,\ \forall N \in \mathbb{N},\ \exists n > N, \qquad |u_n - \ell| \leq \varepsilon.$

Démonstration

$(1 \Rightarrow 2)$ Soit ℓ la limite d'une suite extraite $(u_{\varphi(n)})_n$ et $\varepsilon > 0$. Par définition de la convergence, il existe $N \in \mathbb{N}$ tel que

$$\forall n \geq N, \qquad |u_{\varphi(n)} - \ell| \leq \varepsilon.$$

Ainsi,

$$\varphi([\![N, +\infty[\![) \subset \{n \in \mathbb{N},\ |u_n - \ell| \leq \varepsilon\},$$

donc ce dernier ensemble est infini.

$(2 \Rightarrow 3)$ Soit $\varepsilon > 0$ et $N \in \mathbb{N}$. Comme $\{n \in \mathbb{N},\ |u_n - \ell| \leq \varepsilon\}$ est infini, il n'est pas inclus dans $[\![0, N]\!]$ donc contient un entier $n > N$.

$(3 \Rightarrow 1)$ Construisons une suite extraite $(u_{\varphi(n)})_n$ telle que $|u_{\varphi(n)} - \ell| \leq 2^{-n}$ pour tout $n \in \mathbb{N}$ par récurrence.

▹ Appliquons la propriété 3 avec $\varepsilon = 2^{-0} = 1$ et $N = 0$. Il existe un entier $\varphi(0) > 0$ tel que $|u_{\varphi(0)} - \ell| \leq 1$.

▹ Soit $n \in \mathbb{N}$. Supposons une telle suite extraite construite jusqu'au rang n et appliquons la propriété 3 avec $\varepsilon = 2^{-(n+1)}$ et $N = \varphi(n)$: il existe un entier $\varphi(n+1) > \varphi(n)$ tel que

$$|u_{\varphi(n+1)} - \ell| \leq 2^{-(n+1)}.$$

La suite $(u_{\varphi(n)})_n$ est par construction (*via* le lemme d'encadrement) de limite ℓ.

■

Définition 6.50.

Une valeur d'adhérence de la suite $(u_n)_n \in \mathbb{K}^{\mathbb{N}}$ est un élément $\ell \in \mathbb{K}$ vérifiant l'une des propriétés de la proposition précédente.

Une suite convergente admet donc une unique valeur d'adhérence. La situation peut être bien plus complexe comme le montrent les quelques exemples suivants.

Exemple

La suite $((-1)^n n)_n$ n'admet pas de valeur d'adhérence réelle car $\lim |u_{\varphi(n)}| = +\infty$ pour toute fonction strictement croissante $\varphi : \mathbb{N} \to \mathbb{N}$.

Exemple

La suite $(((-1)^n + 1)n)_n$ admet 0 comme unique valeur d'adhérence réelle mais ne converge pas : la suite des termes d'indice impair diverge en valeur absolue vers $+\infty$ donc une suite extraite convergente ne contient que des termes d'indice pair qui sont tous nuls ; ainsi, la seule valeur d'adhérence est 0.

Exemple

La suite $((-1)^n)_n$ admet deux valeurs d'adhérence réelles : les termes d'indice pair (respectivement impair) forment une suite extraite égale à 1 (respectivement à -1).

Exemple

La suite $(\cos \frac{n\pi}{3})_n$ est 6-périodique et prend quatre valeurs : en étudiant les termes selon le reste modulo 6, on trouve que les valeurs d'adhérence sont $\{\pm\frac{1}{2}, \pm 1\}$.

Exemple

La suite $(\cos(n))_n$ admet une infinité de valeurs d'adhérence réelles; plus précisément, l'ensemble des valeurs d'adhérence est $[-1,1]$. En effet, considérons $\ell \in [-1,1]$, $\theta \in \mathbb{R}$ tel que $\ell = \cos\theta$, et $\varepsilon \in]0,\pi[$. Le sous-groupe $\mathbb{Z}+2\pi\mathbb{Z}$ est dense dans $\mathbb{R}$ donc l'ensemble $A = (\mathbb{Z}+2\pi\mathbb{Z}) \cap]\theta-\varepsilon, \theta+\varepsilon[$ est infini. Pour tout $p+2q\pi \in A$,

$$|\cos|p| - \ell| = |\cos(p+2q\pi) - \cos(\theta)| \leq |p+2q\pi-\theta| \leq \varepsilon.$$

donc $\cos|p| \in]\ell-\varepsilon, \ell+\varepsilon[$. De plus, comme $\varepsilon < \pi$, il y a au plus deux éléments de l'ensemble A qui ont le même cosinus. La suite $(\cos(n))_n$ prend donc une infinité de valeurs dans $]\ell-\varepsilon, \ell+\varepsilon[$.

Proposition 6.51.

Soit $(u_n)_n$ une suite telle que $u_{n+1} - u_n \to 0$. Alors, l'ensemble des valeurs d'adhérence est un intervalle (éventuellement vide).

Démonstration

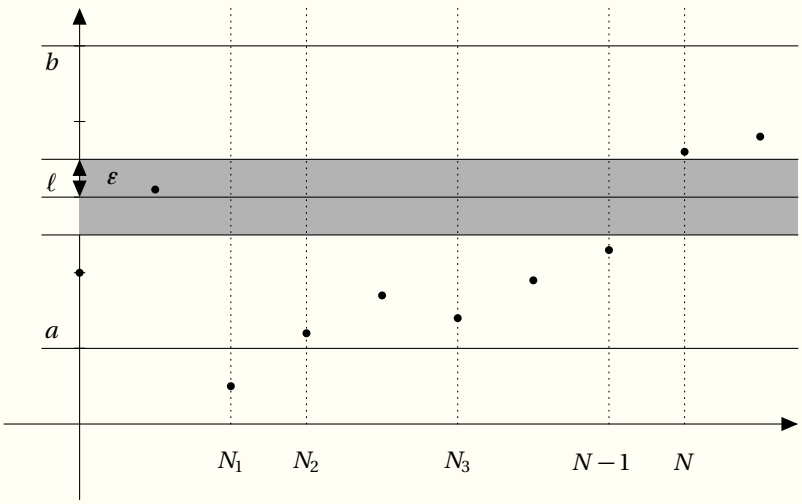
Si $(u_n)_n$ admet 0 ou 1 valeur d'adhérence, le résultat est évident. Si $a < b$ sont deux valeurs d'adhérence de $(u_n)_n$, montrons par l'absurde que tout $\ell \in]a,b[$ est une valeur d'adhérence. Supposons ainsi que

$$\exists \varepsilon > 0,\ \exists N_1 \in \mathbb{N},\ \forall n \geq N_1, \qquad |u_n - \ell| \geq \varepsilon.$$

Comme $u_{n+1} - u_n \to 0$, il existe $N_2 \in \mathbb{N}$ tel que

$$\forall n \geq N_2, \qquad |u_{n+1} - u_n| \leq \varepsilon.$$

Comme a est une valeur d'adhérence, il existe $N_3 \geq \max\{N_1, N_2\}$ tel que $u_{N_3} \leq \frac{\ell+a}{2} < \ell$.
Comme b est une valeur d'adhérence, l'ensemble $\{n > N_3, u_n \geq \ell\}$ est non vide; notons N son plus petit élément. Alors, $u_N \geq \ell > u_{N-1}$.
Or, d'après l'hypothèse sur ℓ, on a $u_N \geq \ell + \varepsilon$ et $u_{N-1} \leq \ell - \varepsilon$ donc $u_N - u_{N-1} \geq 2\varepsilon$ ce qui contredit l'hypothèse $|u_{n+1} - u_n| \leq \varepsilon$.



■

Remarque

En fait, il n'est pas difficile de montrer que cet intervalle est nécessairement fermé.

6.3. Théorème de Bolzano-Weierstrass

Proposition 6.52. Bolzano-Weierstrass

Toute suite réelle bornée admet une suite extraite convergente.

On obtient une formulation équivalente avec la définition de valeur d'adhérence.

Proposition 6.53. Bolzano-Weierstrass

Toute suite réelle bornée admet une valeur d'adhérence.

Démonstration

Notons X l'ensemble des réels x tel qu'il existe une infinité d'indices n tels que $u_n \geq x$. Comme la suite $(u_n)_n$ est bornée, l'ensemble X est majoré et non vide donc admet une borne supérieure ℓ. Pour tout $\varepsilon > 0$, $\ell + \varepsilon \notin X$ donc il y a, au plus, un nombre fini d'indices n tels que $u_n \geq \ell + \varepsilon$: par conséquent, il y a une infinité d'indices n tels que $\ell + \varepsilon \geq u_n \geq \ell - \varepsilon$.

Construisons alors une suite extraite $(u_{\varphi(n)})_n$ qui converge vers ℓ.

Soit $\varphi(0)$ un indice k tel que $\ell + 1 \geq u_k \geq \ell - 1$.

Si $\varphi(n)$ est défini pour un entier n fixé, on prend pour $\varphi(n+1)$ l'un des indices $k > \varphi(n)$ vérifiant l'inégalité $\ell + 2^{-n-1} \geq u_k \geq \ell - 2^{-n-1}$ (ce qui est possible car il y a une infinité de tels indices). Ainsi, pour tout $n \in \mathbb{N}$, $|u_{\varphi(n)} - \ell| \leq 2^{-n}$ donc $u_{\varphi(n)} \to \ell$. ■

Donnons une seconde preuve de ce théorème s'appuyant sur le principe dichotomique.

Démonstration

Notons $[a_0, b_0]$ un intervalle contenant tous les termes de la suite $(u_n)_n$. Construisons, par récurrence, des suites $(a_n)_n$ et $(b_n)_n$ telles que, pour tout $n \in \mathbb{N}$, l'intervalle $[a_n, b_n]$ contient les termes de la suite $(u_n)_n$ correspondant à un ensemble infini d'indices et $b_n - a_n = 2^{-n}(b_0 - a_0)$.

Supposons les termes a_n et b_n définis. L'un (au moins) des deux intervalles $[a_n, \frac{a_n+b_n}{2}]$ et $[\frac{a_n+b_n}{2}, b_n]$ contient les termes de la suite $(u_n)_n$ correspondant à un ensemble infini d'indices ; notons cet intervalle $[a_{n+1}, b_{n+1}]$.

Les suites $(a_n)_n$ et $(b_n)_n$ convergent vers la même limite ℓ en remarquant qu'elles sont adjacentes. Enfin, il existe une suite extraite $(u_{\varphi(n)})_n$ telle que, pour tout $n \in \mathbb{N}$, $u_{\varphi(n)} \in [a_n, b_n]$: cette suite extraite converge vers ℓ par le lemme d'encadrement. ■

Le résultat s'étend naturellement aux suites complexes.

Proposition 6.54. Bolzano-Weierstrass

Toute suite complexe bornée admet une suite extraite convergente.

Démonstration

Soit $(u_n)_n$ une suite complexe bornée ; les suites $(\mathfrak{Re}(u_n))_n$ et $(\mathfrak{Im}(u_n))_n$ sont des suites réelles bornées. D'après le théorème de Bolzano-Weierstrass (cas réel), il existe une suite extraite $(\mathfrak{Re}(u_{\varphi(n)}))_n$ qui converge vers un réel noté α. Comme la suite $(\mathfrak{Im}(u_{\varphi(n)}))_n$ est une suite réelle bornée, on peut aussi en extraire une suite $(\mathfrak{Im}(u_{\varphi\circ\psi(n)}))_n$ qui converge vers un réel noté β. De plus, la suite $(\mathfrak{Re}(u_{\varphi\circ\psi(n)}))_n$ converge vers α en tant que suite extraite d'une suite convergente. Alors, la suite $(u_{\varphi\circ\psi(n)})_n$, extraite de $(u_n)_n$, converge vers le complexe $\alpha + i\beta$. ■

Proposition 6.55.

Soit $(u_n)_n$ une suite réelle bornée et $\ell \in \mathbb{R}$ telle que toute suite extraite convergente de $(u_n)_n$ converge vers ℓ. Alors, la suite $(u_n)_n$ converge vers ℓ.

Démonstration

Si $(u_n)_n$ ne converge pas vers ℓ, alors il existe $\varepsilon > 0$ tel que, pour tout $N \in \mathbb{N}$ tel qu'il existe un indice $n \geq N$, $|u_n - \ell| > \varepsilon$. On peut donc (quitte à ne considérer que les indices n tels que $|u_n - \ell| > \varepsilon$) obtenir une suite extraite $(u_{\varphi(n)})_n$ qui ne prend pas ses valeurs dans $[\ell - \varepsilon, \ell + \varepsilon]$. Cette suite est bornée (car $(u_n)_n$ est bornée) donc admet à son tour une suite extraite qui converge, d'après le théorème de Bolzano-Weierstrass, vers un réel différent de ℓ par construction. Ainsi, $(u_n)_n$ admet deux suites extraites qui convergent vers des valeurs distinctes : contradiction. ■

FICHE

SYNTHÈSE

Précisons la définition de convergence d'une suite numérique ($\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$).

Convergence, divergence

▹ Une suite $(u_n)_n \in \mathbb{K}^{\mathbb{N}}$ est convergente s'il existe un élément $\ell \in \mathbb{K}$, appelé limite de $(u_n)_n$ et noté $\lim u_n$, tel que

$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N},\ \forall n \geq N, \qquad |u_n - \ell| \leq \varepsilon.$$

Si elle ne vérifie pas cette condition, la suite $(u_n)_n$ est divergente.
▹ Une suite réelle $(u_n)_n$ diverge vers $+\infty$ si

$$\forall A > 0,\ \exists N \in \mathbb{N},\ \forall n \geq N, \qquad u_n \geq A.$$

▹ Une suite complexe converge si, et seulement si, la suite de ses parties réelles et la suite de ses parties imaginaires convergent.

Pour les suites réelles, les encadrements peuvent permettre d'établir la convergence ou de donner la valeur de la limite.

Encadrement

Soit $(u_n)_n$, $(v_n)_n$ et $(w_n)_n$ trois suites réelles telles que
- à partir d'un certain rang, $u_n \leq v_n \leq w_n$;
- les suites $(u_n)_n$ et $(w_n)_n$ convergent vers la même limite ℓ.

Alors, la suite $(v_n)_n$ converge et sa limite est ℓ.

Avec la propriété de la borne supérieure, on peut déduire le comportement asymptotique de l'étude des variations.

Monotonie

▹ Une suite réelle croissante majorée converge vers sa borne supérieure.
▹ Une suite réelle croissante non majorée diverge vers $+\infty$.
▹ Soit deux suites réelles $(a_n)_n$ et $(b_n)_n$ telles que
- la suite $(a_n)_n$ est croissante ;
- la suite $(b_n)_n$ est décroissante ;
- $\lim(b_n - a_n) = 0$.

Alors, les suites $(a_n)_n$ et $(b_n)_n$ convergent vers la même limite.

En cas de divergence, on peut étudier les suites extraites (et leurs limites) pour obtenir d'autres comportements asymptotiques. Dans le cas des suites bornées, on est assurés de l'existence d'une telle limite.

Bolzano-Weierstrass

De toute suite (réelle ou complexe) bornée, on peut extraire une suite convergente.

Les suites permettent de caractériser la notion de densité.

Parties denses

Une partie $A \subset \mathbb{R}$ est dense dans $\mathbb{R}$ si elle vérifie l'une des deux propriétés équivalentes suivantes :

- tout réel est limite d'une suite à valeurs dans A ;
- $\forall x \in \mathbb{R},\ \forall \varepsilon > 0,\ A \cap]x-\varepsilon, x+\varepsilon[\neq \varnothing$.

Les ensembles des décimaux, des rationnels ou des irrationnels sont denses dans $\mathbb{R}$.
Une partie de $\mathbb{R}$ non-vide, stable par addition et soustraction est de la forme $a\mathbb{Z}$ ou est dense dans $\mathbb{R}$.

EXERCICES

Vrai ou faux ?

		Vrai	Faux
a)	$\lim\left(1+\frac{3}{n}\right)^n = 1.$	☐	☐
b)	Une suite $(u_n)_n$ telle que $u_{n+2} = 2u_{n+1} - u_n$ pour tout $n \in \mathbb{N}$ converge si, et seulement si, $u_0 = u_1 = 0$.	☐	☐
c)	Une suite $(u_n)_n$ telle que $5u_{n+2} = 5u_{n+1} - u_n$ pour tout $n \in \mathbb{N}$ converge vers 0.	☐	☐
d)	Une suite monotone converge.	☐	☐
e)	Une suite bornée converge.	☐	☐
f)	Une suite géométrique bornée converge.	☐	☐
g)	Pour deux suites $(u_n)_n$ et $(v_n)_n$ de limite 1, $u_n^n = v_n^n$.	☐	☐
h)	Le quotient de deux suites convergentes est convergent.	☐	☐
i)	La partie réelle d'une suite complexe convergente est convergente.	☐	☐
j)	La partie entière d'une suite réelle convergente est convergente.	☐	☐
k)	Le maximum d'une suite réelle convergente et de la suite nulle définit une suite convergente.	☐	☐
l)	Le maximum de deux suites réelles convergentes définit une suite convergente.	☐	☐
m)	Une suite (u_n) croissante telle que (u_{2n}) est convergente est convergente.	☐	☐
n)	Si les suites extraites $(u_{2n})_n$ et $(u_{2n+1})_n$ convergent vers la même limite, alors $(u_n)_n$ converge.	☐	☐

Exercices d'application du chapitre 6

●○○ **Exercice 1**

Préciser les intersections suivantes $\bigcap_{n\in\mathbb{N}\setminus\{0\}}\left[-\frac{1}{n},\frac{1}{n}\right]$ et $\bigcap_{n\in\mathbb{N}\setminus\{0\}}\left]-\frac{1}{n},\frac{1}{n}\right[$.

●○○ **Exercice 2**

Le problème des tours de Hanoï consiste à déplacer des disques de tailles différentes entre trois pics de sorte à ce qu'au départ, tous les disques sont sur le premier pic (la taille étant décroissante avec la hauteur) et qu'à l'arrivée, tous les disques sont sur le troisième pic. Les contraintes sont les suivantes

- on ne peut déplacer qu'un disque du sommet d'un pic vers le sommet d'un autre pic ;
- on ne peut poser un disque que sur un pic vide ou un pic dont le disque au sommet est de rayon plus grand que celui que l'on pose.

Notons u_n le nombre minimal de déplacements pour n disques.

1. Déterminer une relation de récurrence pour la suite $(u_n)_n$.
2. En déduire la valeur de u_n.

●●○ **Exercice 3**

Expliciter la suite $(u_n)_n$ définie par $u_0 = 0$ et, pour tout n, $u_{n+1} = 2u_n + 2^n - 1$.

●○○ **Exercice 4**

Soit $\alpha \in \mathbb{R}$ et $(u_n)_n$ une suite réelle vérifiant, pour tout $n \in \mathbb{N}$,

$$u_{n+2} - 2\cos(\alpha)u_{n+1} + u_n = 0.$$

Donner l'expression de u_n en fonction de n.

●○○ **Exercice 5**

Soit $a, b \in \mathbb{R}$, (u_n) une suite vérifiant la relation de récurrence linéaire d'ordre 2

$$\forall n \in \mathbb{N},\ u_{n+2} = a\,u_{n+1} + b\,u_n.$$

1. Montrer que la suite (u_{2n}) vérifie aussi une relation de récurrence linéaire d'ordre 2.
2. Justifier que les carrés des racines de l'équation caractéristique de (u_n) sont les racines de l'équation caractéristique de (u_{2n}).

●○○ **Exercice 6**

Soit $(u_n)_n$ une suite réelle convergente de limite ℓ.

1. Montrer que si $\ell \notin \mathbb{Z}$, la suite $(\lfloor u_n \rfloor)_n$ converge.
2. Dans le cas général, est-ce que $(\lfloor u_n \rfloor)_n$ est convergente ?

●○○ **Exercice 7**

La suite $(\sqrt[n]{1789 + (-1)^n})_n$ est-elle convergente ?

●○○ **Exercice 8**

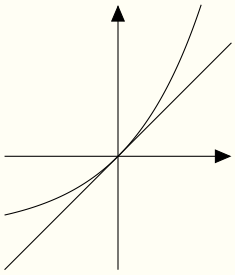
Étudier, par inégalités, la convergence des suites définies par

1. $u_n = \frac{1}{n^2}\sum_{k=1}^{n}\frac{1}{k+2}$
2. $v_n = n^2\sum_{k=1}^{n}\frac{1}{n^3+k}$

EXERCICES

●○○ **Exercice 9**

Étudier la convergence de la suite $(u_n)_n$ définie par $u_0 \in \mathbb{R}$ et, pour tout $n \in \mathbb{N}$, $u_{n+1} = e^{u_n} - 1$.



●○○ **Exercice 10**

Étudier la convergence de la suite $(u_n)_n$ définie par $u_0 = 1$ et, pour tout $n \in \mathbb{N}$,

$$u_{n+1} = u_n + \frac{1}{\sqrt{u_n}}.$$

●●○ **Exercice 11**

Soit $(u_n)_n$ une suite à valeurs strictement positives telles que $\frac{u_{n+1}}{u_n} \to \ell \in \mathbb{R}_+^*$. Montrer que $u_n^{\frac{1}{n}} \to \ell$.

●●○ **Exercice 12 Lemme de Fekete**

Soit $(u_n)_{n\geq 1}$ une suite positive telle que

$$\forall m, n \in \mathbb{N}, \qquad u_{m+n} \leq u_m + u_n.$$

Montrer que la suite $\left(\frac{u_n}{n}\right)_{n\geq 1}$ converge vers

$$\ell = \inf\left\{\frac{u_n}{n},\ n \in \mathbb{N} \setminus \{0\}\right\}.$$

●●○ **Exercice 13**

Montrer que $D = \left\{2^a 5^b \text{ avec } (a,b) \in \mathbb{Z}^2\right\}$ est dense dans $\mathbb{R}_+$, c'est-à-dire que tout réel positif est la limite d'une suite à valeurs dans D.

●●● **Exercice 14**

Soit $x \in \mathbb{R} \setminus \mathbb{Q}$, $(p_n)_{n\in\mathbb{N}}$, $(q_n)_{n\in\mathbb{N}}$ deux suites d'entiers telles que $q_n \neq 0$ pour tout entier n et $\left(\frac{p_n}{q_n}\right)_n$ converge vers x. Montrer que les suites $(|p_n|)_n$ et $(|q_n|)_n$ divergent vers $+\infty$.

●○○ **Exercice 15**

Notons, pour tout $n \in \mathbb{N} \setminus \{0\}$, d_n le nombre de diviseurs de n. Déterminer les valeurs d'adhérence de la suite $(d_n)_n$.

●●● **Exercice 16**

Soit $\lambda > 1$. Associons à toute suite $(u_n)_n$ la suite $(v_n)_n$ définie par $v_n = \lambda u_{n+1} + u_n$. Montrer que $(u_n)_n$ converge si, et seulement si, $(v_n)_n$ converge.

●●○ **Exercice 17**

Considérons la suite définie, pour tout $n \in \mathbb{N}$, par

$$u_n = \sqrt{n} - \lfloor \sqrt{n} \rfloor.$$

1. Soit $p \in \mathbb{N}$. Montrer que $(u_n)_n$ est croissante pour $p^2 < n < (p+1)^2$.
2. Exhiber des suites extraites de $(u_n)_n$ de limite 0 et 1.
3. Montrer que tout réel $x \in]0,1[$ est une valeur d'adhérence de $(u_n)_n$.

Exercices d'approfondissement du chapitre 6

●●○ **Exercice A**

Soit $(u_n)_n$ telle que, pour tous $n, p > 0$, $0 \leq u_{n+p} \leq \frac{n+p}{np}$.
Montrer que $(u_n)_n$ converge.

●●● **Exercice B**

Soit $z_1, \ldots, z_p$ des complexes de module 1 tels que la suite $(z_1^n + \ldots + z_p^n)_n$ converge et ℓ sa limite. Montrer que $\ell \in [\![0, p]\!]$.
Indication : on pourra utiliser le théorème de Cesàro.

●●○ **Exercice C**

Soit $(u_n) \in (\mathbb{R}_+^*)^{\mathbb{N}}$ de limite nulle. Montrer que la suite $\left(\sum_{k=0}^{n} \frac{u_{n-k}}{2^k}\right)$ est aussi de limite nulle.

●●○ **Exercice D**

Notons, pour tout $n \in \mathbb{N} \setminus \{0\}$, d_n le plus grand diviseur impair de n et considérons $u_n = \sum_{k=1}^{n} \frac{d_k}{k}$.

1. Montrer que
$$u_n = \begin{cases} u_{n-1} + 1 & \text{si } n \text{ est impair,} \\ \frac{1}{2}n + \frac{1}{2}u_{\frac{n}{2}} & \text{sinon.} \end{cases}$$
2. Montrer que, pour tout n, $\frac{2}{3}n < u_n < \frac{2}{3}(n+1)$.
3. En déduire que $\lim \frac{3u_n}{2n} = 1$.

●●● **Exercice E**

Soit $(\lambda_n)_n$ une suite à valeurs dans $]0,1[$, convergente dans $]0,1[$. Considérons deux suites $(u_n)_n$ et $(v_n)_n$ vérifiant $u_0 \leq v_0$ et

$$\forall n \in \mathbb{N}, \quad \begin{cases} u_{n+1} &= \lambda_n u_n + (1-\lambda_n) v_n \\ v_{n+1} &= (1-\lambda_n) u_n + \lambda_n v_n \end{cases}$$

Montrer que les suites $(u_n)_n$ et $(v_n)_n$ convergent vers la même limite.

●●○ **Exercice F**

Soit $(u_n)_n$ une suite bornée telle que $u_n + \frac{1}{2}u_{2n} \to 1$. Montrer que $(u_n)_n$ converge vers une limite que l'on précisera.

●●● **Exercice G**

Soit $(u_n)_n \in \mathbb{R}^{\mathbb{N}}$ telle que $u_{n+1} - u_n \to 0$ et $u_n \to +\infty$. Montrer qu'il existe une suite extraite $(u_{\varphi(n)})_n$ telle que $u_{\varphi(n)} - n \to 0$.

PROBLÈMES

Problèmes du chapitre 6

●○○ **Problème 1**

L'objet de ce problème est de montrer qu'il existe une fonction $f : \mathbb{R} \to \mathbb{R}$ telle que la restriction de f à tout intervalle non trivial soit surjective.

À cette fin, posons, pour tout $x \in \mathbb{R}$, $f(x)$ la limite finie de la suite $(\tan(n!\pi x))_n$ si celle-ci existe, 0 sinon.

1. a. Montrer que, pour tout $r \in \mathbb{Q}$, la fonction f est r-périodique.
   b. En déduire la restriction de f à $\mathbb{Q}$.
2. Le but de cette question est de montrer que f est périodique. Soit $y \in \mathbb{R}$.
   a. Montrer qu'il existe $x \in [0,1[$ tel que $y = \tan(\pi x)$.
   Notons, pour tout n,
$$u_n = \sum_{k=0}^{n} \frac{\lfloor kx \rfloor}{k!}.$$
   b. Justifier que la suite (u_n) converge.
   Soit ℓ la limite de cette suite.
   c. Établir que $n!(\ell - u_n) \to x$.
   d. En déduire que $f(\ell) = y$.
3. Montrer le résultat annoncé.

●●● **Problème 2**

Partie A – Définition et premiers exemples

On rappelle que le cardinal d'une partie finie E, noté card E, est son nombre d'éléments. La densité d'une partie $A \subset \mathbb{N} \setminus \{0\}$ est, quand elle existe, la limite
$$\theta(A) = \lim_{n\to\infty} \frac{1}{n} \operatorname{card}(A \cap [\![1, n]\!]).$$

1. a. Justifier que la densité d'une partie est un réel de l'intervalle $[0,1]$.
   b. Quelle est la densité d'une partie finie ?
   c. Supposons qu'une partie A soit de densité p. Montrer que le complémentaire de A dans $\mathbb{N}\setminus\{0\}$ admet une densité que l'on précisera.
   d. En déduire que si une partie est de densité nulle, alors son complémentaire dans $\mathbb{N} \setminus \{0\}$ est infini dénombrable (c'est-à-dire en bijection avec $\mathbb{N}$).
2. Soit $p \in \mathbb{N} \setminus \{0\}$. Déterminer la densité de l'ensemble $p\mathbb{N} \setminus \{0\} = \{kp,\ k \in \mathbb{N} \setminus \{0\}\}$.
3. Déterminer la densité de l'ensemble des entiers dont l'écriture décimale ne comporte aucun 9.
4. Dans toute cette question, la variable p désigne un nombre premier et $\pi(n)$ le nombre de premiers inférieurs ou égaux à n.
   a. Montrer que, pour tout $n \in \mathbb{N} \setminus \{0\}$, $\prod_{2n+1 \geq p > n+1} p$ divise $\binom{2n+1}{n}$.
   b. En déduire, par récurrence forte, que $\prod_{p \leq n} p \leq 4^n$ pour tout $n \in \mathbb{N} \setminus \{0\}$.
   Indication : on pourra montrer et utiliser que $2\binom{2l+1}{l} \leq 2^{2l+1}$.
   c. Montrer que pour tout $m \in [\![2, n]\!]$, $m^{\pi(n)-\pi(m)} \leq 4^n$ et donc
$$\pi(n) \leq \frac{n \ln 4}{\ln m} + \pi(m).$$
   d. Déterminer la densité de l'ensemble des nombres premiers.
   Indication : on pourra prendre $m = \lfloor \frac{n}{\ln(n)} \rfloor$ dans l'inégalité précédente.

Partie B – Une réciproque du théorème de Cesàro

1. Considérons une suite $(u_n)_n$ telle que $\frac{1}{n+1}\sum_{k=0}^{n}|u_k| \to 0$.
   Pour tout $\varepsilon > 0$, on définit la partie $A_\varepsilon = \{n \in \mathbb{N}^*,\ |u_n| \geq \varepsilon\}$.
   a. Montrer que $\theta(A_\varepsilon) = 0$ pour tout $\varepsilon > 0$.
   b. Pour tout $k \in \mathbb{N} \setminus \{0\}$, on pose $p_k(n) = \mathrm{card}(A_{\frac{1}{k}} \cap [\![1, n]\!])$.
   Montrer qu'il existe un entier n_k tel que, pour tout $n \geq n_k$,
   $$\frac{1}{n}p_{k+1}(n) \leq \frac{1}{k}.$$
   Supposons dorénavant que la suite $(n_k)_k$ est strictement croissante.
   c. Calculer la densité de la partie $A = \bigcup_{k\in\mathbb{N}}(A_{\frac{1}{k}} \cap [\![n_{k-1}, n_k]\!])$.
   d. Notons $\{\varphi(k),\ k \in \mathbb{N}^*\}$ avec φ une fonction croissante, le complémentaire de A dans $\mathbb{N} \setminus \{0\}$.
   Montrer que $u_{\varphi(k)} \to 0$.
2. Considérons une suite $(u_n)_n$ telle que $\frac{1}{n+1}\sum_{k=0}^{n}|u_k - \ell| \to 0$.
   Montrer qu'il existe une suite extraite de $(u_n)_n$ qui converge vers ℓ et dont l'ensemble des indices est de densité 1.

Partie C – Application aux séries

1. Montrer que la suite $\left(\sum_{k=1}^{n}\frac{1}{k}\right)_n$ diverge vers $+\infty$.
   Supposons dorénavant que la suite positive $(u_n)_n$ est telle que $\left(\sum_{k=1}^{n}\frac{u_k}{k}\right)_n$ converge.
2. Notons pour tout $n \in \mathbb{N} \setminus \{0\}$, $S_n = \sum_{k=1}^{n}\frac{u_k}{k}$ et $S_0 = 0$.
   a. En remarquant que $u_n = n(S_n - S_{n-1})$, montrer que
   $$\frac{1}{n}\sum_{k=1}^{n}u_k = S_n - \frac{1}{n}\sum_{k=1}^{n-1}S_k$$
   b. En déduire que la suite $\left(\frac{1}{n+1}\sum_{k=0}^{n}u_k\right)_n$ converge vers 0.
3. Conclure qu'il existe une suite extraite de $(u_n)_n$ qui converge vers 0 et dont l'ensemble des indices est de densité 1.

PROBLÈMES

CORRIGÉS

Corrigés des Vrai/Faux

a) Faux, e^3.

b) Faux, $(1)_n$

c) Vrai.

d) Faux, $(n)_n$.

e) Faux, $((-1)^n)_n$.

f) Faux, $((-1)^n)_n$.

g) Faux avec, par exemple $u_n = 1$ et $v_n = 1 + \frac{1}{n}$: $u_n^n = 1$ et $v_n^n \to e$.

h) Faux et si le dénominateur tend vers 0.

i) Vrai.

j) Faux, $(1 + (-1)^n/n)_n$.

k) Vrai.

l) Vrai.

m) Vrai.

n) Vrai.

Corrigés des exercices d'application du chapitre 6

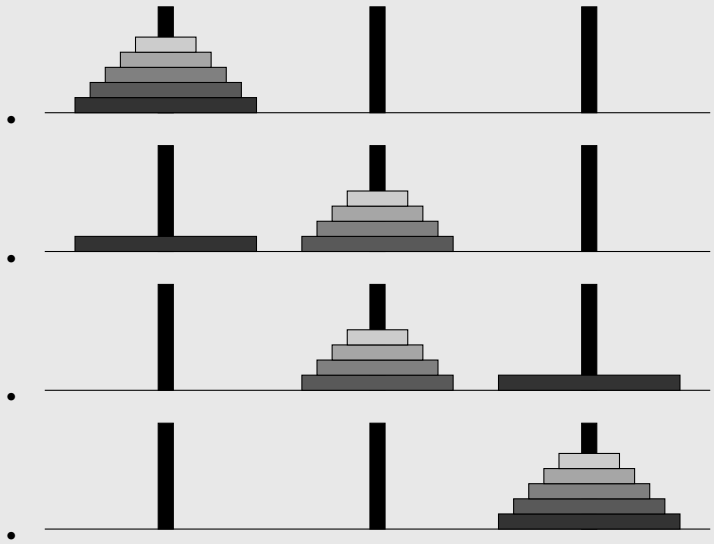
Exercice 1

Remarquons tout d'abord que 0 appartient à ces deux ensembles.
Par ailleurs, pour tout $x \in \mathbb{R}^*$, il existe $n \in \mathbb{N} \setminus \{0\}$ tel que $|x| > \frac{1}{n}$ donc $x \notin \left[-\frac{1}{n}, \frac{1}{n}\right]$ et, a fortiori, $x \notin \left]-\frac{1}{n}, \frac{1}{n}\right[$.
En conclusion,

$$\bigcap_{n \in \mathbb{N} \setminus \{0\}} \left[-\frac{1}{n}, \frac{1}{n}\right] = \bigcap_{n \in \mathbb{N} \setminus \{0\}} \left]-\frac{1}{n}, \frac{1}{n}\right[= \{0\}.$$

Exercice 2

1. Remarquons que pour déplacer une tour de $n+1$ disques, il faut à un moment déplacer le plus grand disque. Ceci impose d'avoir au préalable déplacer les n plus petits disques sur un même pic puis de les replacer ensuite sur le grand disque déplacé.



Ainsi, pour tout n, $u_{n+1} = u_n + 1 + u_n = 2u_n + 1$.
Remarquons enfin que $u_1 = 1$.
2. La suite considérée est arithmético-géométrique : on trouve, avec les notations du cours, $\ell = -1$ et donc $u_n = 2^n - 1$ pour tout $n \in \mathbb{N}$.

Exercice 3

On remarque tout d'abord que la solution de l'équation « sans second membre » $u_{n+1} = 2u_n$ est la suite géométrique de raison 2. Revenons à l'équation de départ est divisons par 2^{n+1} ; alors, pour tout n,

$$\frac{u_{n+1}}{2^{n+1}} - \frac{u_n}{2^n} = \frac{1}{2} - 2^{-n-1}.$$

Sommons ces relations en utilisant le caractère télescopique

$$\begin{aligned}\frac{u_n}{2^n} - u_0 &= \sum_{k=0}^{n-1} \frac{u_{k+1}}{2^{k+1}} - \frac{u_k}{2^k} \\ &= \sum_{k=0}^{n-1} \frac{1}{2} - 2^{-k-1} \\ &= \frac{n}{2} - 1 + 2^{-n}\end{aligned}$$

en utilisant l'expression d'une somme géométrique. En conclusion, pour tout $n \in \mathbb{N}$,

$$u_n = n2^{n-1} - 2^n + 1.$$

Exercice 4

Il suffit essentiellement de chercher les racines de l'équation caractéristique $x^2 - 2\cos\alpha x + 1 = 0$. Il s'agit de $r_1 = e^{i\alpha}$ et $r_2 = e^{-i\alpha}$. Il faut alors discuter si ces racines sont distinctes ou non.
▹ Si α est un multiple pair de π, alors $r_1 = r_2 = 1$. Alors, il existe A, $B \in \mathbb{R}$ tels que, pour tout $n \in \mathbb{N}$, $u_n = An + B$.
▹ Si α est un multiple impair de π, alors $r_1 = r_2 = -1$. Alors, il existe A, $B \in \mathbb{R}$ tels que, pour tout $n \in \mathbb{N}$, $u_n = (An + B)(-1)^n$.
▹ Sinon, $r_1 \neq r_2$ et il existe A, $B \in \mathbb{R}$ tels que, pour tout $n \in \mathbb{N}$, $u_n = A\cos(n\alpha) + B\sin(n\alpha)$.

Exercice 5

1. Calculons, pour tout $n \in \mathbb{N}$,

$$\begin{aligned} u_{2(n+2)} &= a\,u_{2n+3} + b\,u_{2n+2} \\ &= (a^2+b)u_{2n+2} + a\,b\,u_{2n+1} \\ &= (a^2+b)u_{2n+2} + b(u_{2n+2} - b\,u_{2n}) \\ &= (a^2+2b)u_{2(n+1)} - b^2 u_{2n}. \end{aligned}$$

2. Un complexe z est une racine de l'équation caractéristique de (u_n) si, et seulement si, (z^n) vérifie l'équation de récurrence linéaire de (u_n); mais alors, (z^{2n}) vérifie l'équation de récurrence linéaire de (u_{2n}) et z^2 est racine de l'équation caractéristique de (u_{2n}).

Exercice 6

1. Notons $k = \lfloor \ell \rfloor$. Comme $\ell \in]k, k+1[$, à partir d'un certain rang N, $u_n \in]k, k+1[$ donc $\lfloor u_n \rfloor = k$. Ainsi, la suite $(\lfloor u_n \rfloor)_n$ est stationnaire donc converge.
2. Pas forcément! L'idée consiste à trouver une suite qui converge vers un entier k en étant tantôt au-dessus (donc de partie entière k), tantôt en dessous (donc de partie entière $k-1$).
Par exemple, considérons, pour tout $n \in \mathbb{N}$, $u_n = \frac{(-1)^n}{n+2}$.
La suite $(u_n)_n$ converge vers 0. Or, $(\lfloor u_{2n} \rfloor)_n$ est constante égale à 0 et $(\lfloor u_{2n+1} \rfloor)_n$ est constante égale à -1 : la suite $(\lfloor u_n \rfloor)_n$ ne converge pas.

Exercice 7

Remarquons que le terme dans la racine vérifie, pour tout n, $1788 \leq 1789 + (-1)^n \leq 1790$. Ainsi, le terme d'indice n de la suite considérée est compris entre $1788^{\frac{1}{n}}$ et $1790^{\frac{1}{n}}$. Comme ces deux termes sont de limite 1, le lemme d'encadrement assure que la suite converge vers 1.

Exercice 8

Pour chacune des suites, on obtient un encadrement en majorant les termes de la somme par le plus petit et le plus grand d'entre eux. On conclut avec le lemme d'encadrement.

1. Pour tout $n > 0$,

$$\frac{1}{n^2}\frac{n}{n+2} = \frac{1}{n^2}\sum_{k=1}^{n}\frac{1}{n+2} \leq u_n \leq \frac{1}{n^2}\sum_{k=1}^{n}\frac{1}{1+2} = \frac{1}{n^2}\frac{n}{3}.$$

Comme chacun des termes extrémaux tend vers 0, on obtient par le lemme d'encadrement que la suite $(u_n)_n$ converge et $\lim u_n = 0$.

2. Pour tout $n > 0$,

$$\frac{n^3}{n^3+n} = n^2\sum_{k=1}^{n}\frac{1}{n^3+n} \leq v_n \leq n^2\sum_{k=1}^{n}\frac{1}{n^3+1} = \frac{n^3}{n^3+1}.$$

Comme chacun des termes extrémaux tend vers 1, on obtient par le lemme d'encadrement que la suite $(v_n)_n$ converge et $\lim v_n = 1$.

Exercice 9

Il convient de remarquer tout d'abord que, pour tout $x \in \mathbb{R}$, $e^x \geq x+1$. Ainsi, pour tout $n \in \mathbb{N}$, $u_{n+1} \geq u_n$: la suite $(u_n)_n$ est croissante. D'après le théorème de limite monotone, soit cette suite converge, soit elle diverge vers $+\infty$. Une limite potentielle ℓ vérifie, en passant à la limite dans la relation de récurrence $\ell = e^\ell - 1$ donc $\ell = 0$.

▹ Si $u_0 > 0$, la suite est croissante depuis un terme strictement positif donc ne peut pas tendre vers 0 : elle diverge vers $+\infty$.
▹ Si $u_0 = 0$, la suite est constante égale à 0 donc converge vers 0.
▹ Si $u_0 < 0$, une récurrence immédiate assure que $u_n \leq 0$ pour tout $n \in \mathbb{N}$. La suite $(u_n)_n$ est croissante majorée donc converge : comme la seule limite possible est 0, la suite converge vers 0.

Exercice 10

Remarquons tout d'abord que la suite est bien définie et croissante car $u_{n+1} > u_n$ pour tout $n \in \mathbb{N}$.
La suite $(u_n)_n$ est croissante réelle donc soit elle converge, soit elle diverge vers $+\infty$. Supposons par l'absurde que $\lim u_n = \ell \in \mathbb{R}$. Alors, en passant à la limite dans la relation de récurrence, $\ell = \ell + \frac{1}{\sqrt{\ell}}$ ce qui est impossible. En conclusion, $(u_n)_n$ diverge vers $+\infty$.

Exercice 11

L'hypothèse entraîne immédiatement que $\ln(u_{n+1}) - \ln u_n \to \ln \ell$. Appliquons le théorème de Cesàro à cette suite :

$$\begin{aligned}\frac{1}{n}(\ln u_n - \ln u_0) &= \frac{1}{n}\sum_{k=0}^{n-1}(\ln u_{k+1} - \ln u_k)\\ &\to \ln \ell.\end{aligned}$$

Ainsi $\frac{1}{n}\ln u_n \to \ln \ell$ et, en composant avec l'exponentielle (qui est continue), $u_n^{\frac{1}{n}} \to \ell$.

Exercice 12

Soit $\varepsilon > 0$ et $N \in \mathbb{N} \setminus \{0\}$ tel que $\frac{u_N}{N} \leq \ell + \varepsilon$. Pour tout $n \geq N$, effectuons la division euclidienne de n par N : $n = qN + r$ avec $r \in [\![0, N-1]\!]$. Alors,

$$\begin{aligned}\frac{u_n}{n} &\leq \frac{q\,u_N + u_r}{n}\\ &\leq \frac{q\,u_N}{qN} + \frac{u_r}{n}\\ &\leq \frac{u_N}{N} + \frac{M}{n},\end{aligned}$$

où $M = \max\{u_0, u_1, \ldots, u_{N-1}\}$. Soit $N' > 0$ tel que, pour tout $n \geq N'$, $\frac{M}{n} \leq \varepsilon$.
En combinant les majorations, on obtient, pour $n \geq \max(N, N')$,

$$\ell \leq \frac{u_n}{n} \leq \ell + 2\varepsilon.$$

Par conséquent, $\left(\frac{u_n}{n}\right)_{n\geq 1}$ converge vers ℓ.

Exercice 13

Remarquons tout d'abord que comme $\frac{\ln 2}{\ln 5} \notin \mathbb{Q}$ (sinon, une puissance entière de 2 serait égale à une puissance entière de 5), l'ensemble $G = \ln(2)\mathbb{Z} + \ln(5)\mathbb{Z}$, non vide, stable par addition et différence, est dense dans $\mathbb{R}$. Ainsi, tout réel est limite d'une suite à valeur dans G.
Soit $x \in \mathbb{R}_+^*$ et $t \in \mathbb{R}$ tel que $x = e^t$. Il existe une suite $(u_n)_n \in G^{\mathbb{N}}$ tel que $\lim u_n = t$. Alors, la suite $(e^{u_n})_n$ est à valeurs dans D et converge vers x.
De plus, la suite $(5^{-n})_n \in D^{\mathbb{N}}$ converge vers 0.
En conclusion, tout réel positif est la limite d'une suite à valeurs dans D.

Exercice 14

▹ Supposons que la suite $(|p_n|)_n$ ne diverge pas vers $+\infty$; alors, il existe une suite extraite $(|p_{\varphi(n)}|)_n$ bornée. D'après le théorème de Bolzano-Weierstrass, on peut en extraire une suite $(|p_{\varphi\circ\psi(n)}|)_n$ qui converge. Comme $\left(\frac{p_n}{q_n}\right)_n$ converge vers x, la suite $(|q_{\varphi\circ\psi(n)}|)_n$ converge aussi. Ces deux suites sont des suites d'entiers donc elles sont stationnaires ; par conséquent,

$$x = \lim \frac{p_n}{q_n} = \lim \frac{p_{\varphi\circ\psi(n)}}{q_{\varphi\circ\psi(n)}} \in \mathbb{Q},$$

d'où la contradiction : la suite $(|p_n|)_n$ diverge vers $+\infty$.

▹ Si la suite $\left(\frac{p_n}{q_n}\right)_n$ converge vers x, alors la suite $\left(\frac{q_n}{p_n}\right)_n$ converge vers $\frac{1}{x} \in \mathbb{R}\setminus\mathbb{Q}$. Donc, d'après le point précédent, la suite $(|q_n|)_n$ diverge vers $+\infty$.

Exercice 15

▹ Remarquons tout d'abord que toute valeur d'adhérence de $(d_n)_n$ est limite d'une suite convergente d'entiers strictement supérieurs à 1 donc est un entier strictement supérieur à 1.

▹ Réciproquement, considérons $k \in \mathbb{N}\setminus\{0,1\}$ et notons $(p_n)_n$ la suite croissante des nombres premiers. Alors, la suite $(d_{p_n^{k-1}})_n$ est extraite de $(d_n)_n$ et constante égale à k (car p_n^{k-1} admet exactement k diviseurs positifs : $1, p_n, p_n^2, \dots, p_n^{k-1}$) donc k est bien valeur d'adhérence de $(d_n)_n$.

Les valeurs d'adhérence de $(d_n)_n$ sont exactement les entiers strictement supérieurs à 1.

Exercice 16

Si la suite $(u_n)_n$ converge vers ℓ, alors par combinaison linéaire, la suite $(v_n)_n$ converge vers $(\lambda+1)\ell$.

Réciproquement, on remarque que pour tout $n \in \mathbb{N}$, $\lambda^n v_n = \lambda^{n+1} u_{n+1} - \lambda^n u_n$. Ainsi, par télescopage,

$$\forall n \in \mathbb{N}, \qquad u_n = \lambda^{-n}\Big(u_0 + \sum_{k=0}^{n-1} \lambda^k v_k\Big).$$

Si la suite $(v_n)_n$ converge vers ℓ, alors

$$\forall n \in \mathbb{N}, \qquad u_n = \lambda^{-n}\Big(u_0 + \sum_{k=0}^{n-1} \lambda^k (v_k - \ell) + \ell\,\frac{1-\lambda^n}{1-\lambda}\Big).$$

On remarque que $\lambda^{-n} u_0 \to 0$ et $\lambda^{-n}\ell\,\frac{1-\lambda^n}{1-\lambda} \to \frac{\ell}{\lambda-1}$. Montrons que le deuxième terme tend vers 0.

Soit $\varepsilon > 0$ puis $N \in \mathbb{N}$ tel que pour tout $n \geq N$, $|v_n - \ell| \leq \varepsilon$. Alors, pour tout $n \geq N$,

$$\begin{aligned}
\Big|\lambda^{-n}\sum_{k=0}^{n-1}\lambda^k(v_k-\ell)\Big| &\leq \lambda^{-n}\sum_{k=0}^{n-1}\lambda^k|v_k-\ell| \\
&= \lambda^{-n}\sum_{k=0}^{N-1}\lambda^k|v_k-\ell| + \lambda^{-n}\sum_{k=N}^{n-1}\lambda^k|v_k-\ell| \\
&\leq \lambda^{-n}\sum_{k=0}^{N-1}\lambda^k|v_k-\ell| + \varepsilon\lambda^{-n}\frac{\lambda^N-\lambda^n}{1-\lambda} \\
&\leq C\varepsilon
\end{aligned}$$

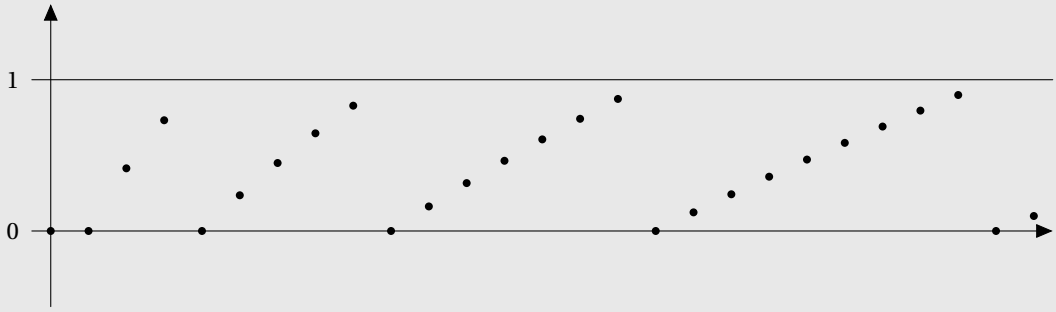
avec C constante, pour n suffisamment grand. Ainsi,

$$\lambda^{-n}\sum_{k=0}^{n-1}\lambda^k(v_k-\ell) \to 0.$$

En conclusion, $(u_n)_n$ converge vers $\frac{\ell}{\lambda-1}$.

Exercice 17

1. Pour les indices n considérés, $\lfloor\sqrt{n}\rfloor = p$ donc $u_n = \sqrt{n} - p$. D'après cette expression, la suite $(u_n)_n$ est croissante pour $p^2 < n < (p+1)^2$.
Voici l'allure de la suite d'après ce résultat.



Les suites extraites que l'on construit aux questions suivantes sont intuitées à partir de cette représentations.
2. La suite extraite $(u_{n^2})_n$ est nulle donc converge vers 0. Montrons que la suite $(u_{n^2+2n})_n$ converge vers 1 ; en effet, pour tout $n \in \mathbb{N}$,

$$u_{n^2+2n} = \sqrt{n^2+2n} - n = \frac{\sqrt{1+\frac{2}{n}}-1}{\frac{1}{n}},$$

et en reconnaissant le taux d'accroissement de la fonction $x \mapsto \sqrt{1+x}$ en 0, on en déduit $u_{n^2+2n} \to 1$.
3. Soit $x \in]0,1[$. Et définissons, pour tout $p \in \mathbb{N}$, $\varphi(p) = \max\{k < (p+1)^2,\ u_k \leq x\}$ (ce qui est bien défini d'après l'étude des variations à la première question).

$$\begin{aligned}|x - u_{\varphi(p)}| &\leq |u_{\varphi(p)+1} - u_{\varphi(p)}| \\ &= \sqrt{\varphi(p)+1} - \sqrt{\varphi(p)} \\ &= \frac{1}{\sqrt{\varphi(p)+1}+\sqrt{\varphi(p)}} \leq \frac{1}{\sqrt{p}},\end{aligned}$$

en utilisant que $\varphi(p) \geq p$. Par conséquent, $\lim u_{\varphi(p)} = x$.

Corrigés des exercices d'approfondissement du chapitre 6

Exercice A

Pour tout $n > 0$, écrivons $n = \lfloor\frac{n}{2}\rfloor + (n - \lfloor\frac{n}{2}\rfloor)$ et utilisons l'inégalité de l'énoncé :

$$0 \leq u_n \leq \frac{\lfloor\frac{n}{2}\rfloor + (n-\lfloor\frac{n}{2}\rfloor)}{\lfloor\frac{n}{2}\rfloor(n-\lfloor\frac{n}{2}\rfloor)} = \frac{n}{\lfloor\frac{n}{2}\rfloor(n-\lfloor\frac{n}{2}\rfloor)}.$$

Comme $\lfloor\frac{n}{2}\rfloor \sim \frac{n}{2}$, on obtient que

$$\frac{n}{\lfloor\frac{n}{2}\rfloor(n-\lfloor\frac{n}{2}\rfloor)} \sim \frac{4n}{n^2} = \frac{4}{n}.$$

Ce majorant tend donc vers 0 et, par lemme d'encadrement, $u_n \to 0$.

Exercice B

Remarquons que si z est un complexe de module 1 différent de 1, alors (comme le numérateur de la dernière fraction est borné par 2)

$$\frac{1}{n+1}\sum_{k=0}^{n} z^k = \frac{1}{n+1}\frac{1-z^{n+1}}{1-z} \to 0.$$

Par ailleurs, si $z=1$,

$$\frac{1}{n+1}\sum_{k=0}^{n} z^k = 1 \to 1.$$

D'après le théorème de Cesàro, comme la suite $(z_1^n+\ldots+z_p^n)_n$ converge vers ℓ,

$$\frac{1}{n+1}\sum_{k=0}^{n}(z_1^k+\ldots+z_p^k) \to \ell.$$

En combinant, ces deux remarques, on obtient que

$$\ell = \operatorname{card}\{j \in [\![0,p]\!],\ z_j = 1\} \in [\![0,p]\!].$$

Exercice C

Soit $\varepsilon > 0$. Soit $N \in \mathbb{N}$ tel que $2^{-N} \leq \varepsilon$ et $N' \in \mathbb{N}$ tel que

$$\forall n \geq N', \qquad u_n \leq \varepsilon.$$

Pour tout $n \geq N+N'$,

$$\begin{aligned}
\sum_{k=0}^{n}\frac{u_{n-k}}{2^k} &= \sum_{k=0}^{n-N'}\frac{u_{n-k}}{2^k} + \sum_{k=n-N'+1}^{n}\frac{u_{n-k}}{2^k} \\
&\leq \varepsilon\sum_{k=0}^{n-N'}\frac{1}{2^k} + \frac{1}{2^{n-N'+1}}\sum_{k=n-N'+1}^{n} u_{n-k} \\
&\leq 2\varepsilon + \frac{1}{2^{n-N'+1}}\sum_{k=0}^{N'-1} u_k \\
&\leq C\varepsilon,
\end{aligned}$$

avec C est une constante indépendante de n ; d'où le résultat annoncé.

Exercice D

1. Soit $n \geq 2$. Si n est impair, $u_n = u_{n-1} + \frac{d_n}{n} = u_{n-1}+1$; si n est pair,

$$\begin{aligned}
u_n = \sum_{k=1}^{n}\frac{d_k}{k} &= \sum_{\substack{k=1\\ k \text{ pair}}}^{n}\frac{d_k}{k} + \sum_{\substack{k=1\\ k \text{ impair}}}^{n}\frac{d_k}{k} \\
&= \sum_{k=1}^{\frac{n}{2}}\frac{d_{2k}}{2k} + \sum_{\substack{k=1\\ k \text{ impair}}}^{n} 1 \\
&= \sum_{k=1}^{\frac{n}{2}}\frac{d_k}{2k} + \frac{n}{2} = \frac{1}{2}u_{\frac{n}{2}} + \frac{n}{2}.
\end{aligned}$$

2. Montrons l'encadrement par récurrence.
- $\frac{2}{3} < u_1 = 1 < \frac{4}{3}$ et $\frac{4}{3} < u_2 = \frac{3}{2} < 2$.
- Soit $n \geq 2$ tel que l'encadrement soit vérifié jusqu'au rang n.

– Si $n+1$ est pair, alors $u_{n+1} = \frac{n+1}{2} + \frac{1}{2} u_{\frac{n+1}{2}}$. D'après l'hypothèse de récurrence,

$$\frac{n+1}{2} + \frac{1}{2} \times \frac{2}{3} \times \frac{n+1}{2} < u_{n+1} < \frac{n+1}{2} + \frac{1}{2} \times \frac{2}{3}\Big(\frac{n+1}{2} + 1\Big)$$

soit

$$\frac{2}{3}(n+1) < u_{n+1} < \frac{2}{3}(n+1) + \frac{1}{3} < \frac{2}{3}(n+2).$$

– Si $n+1$ est impair, alors $u_{n+1} = 1 + u_n = 1 + \frac{n}{2} + \frac{1}{2} u_{\frac{n}{2}}$. D'après l'hypothèse de récurrence,

$$\frac{2}{3}n + 1 < u_{n+1} < 1 + \frac{n}{2} + \frac{1}{2}.\frac{2}{3}\Big(\frac{n}{2} + 1\Big)$$

soit

$$\frac{2}{3}(n+1) < u_{n+1} < \frac{2}{3}n + \frac{4}{3} = \frac{2}{3}(n+2).$$

3. On conclut avec le lemme d'encadrement.

Exercice E

On vérifie immédiatement par récurrence que $(u_n)_n$ est croissante, $(v_n)_n$ est décroissante et pour tout n, $u_n \leq v_n$.
Remarquons que, pour tout $n \in \mathbb{N}$,

$$|u_{n+1} - v_{n+1}| = |2\lambda_n - 1|.|u_n - v_n|.$$

La suite $(|u_n - v_n|)_n$ converge vers 0. Il existe ℓ tel que $\lim u_n = \lim v_n = \ell$. Par ailleurs, comme

$$\forall n \in \mathbb{N}, \qquad \frac{1}{2}(u_{n+1} + v_{n+1}) = \frac{1}{2}(u_n + v_n),$$

on trouve $\ell = \frac{1}{2}(u_0 + v_0)$.

Exercice F

Comme $(u_n)_n$ est bornée, elle admet, d'après le théorème de Bolzano et Weierstrass, une suite extraite $(u_{\varphi(n)})_n$ convergente vers $\ell \in \mathbb{C}$. Alors, $(u_{2\varphi(n)})_n$ converge vers $f(\ell) = 2(1-\ell)$. Par récurrence, on trouve que $f^n(\ell)$ est encore une valeur d'adhérence de $(u_n)_n$. Étudions la suite arithmético-géométrique définie par $\ell_{n+1} = 2(1-\ell_n)$ et $\ell_0 = \ell$. On obtient $\ell_n = \frac{2}{3} + (-2)^n\left(\ell - \frac{2}{3}\right)$. Or, cette suite est bornée car composée de valeurs d'adhérence de la suite bornée $(u_n)_n$. Par conséquent, $\ell = \frac{2}{3}$. En conclusion, la suite $(u_n)_n$ est bornée et admet une unique valeur d'adhérence $\frac{2}{3}$ donc converge vers $\frac{2}{3}$.

Exercice G

▹ Commençons par extraire une sous-suite $(u_{\psi(n)})_n$ telle que $u_{\psi(n)} \geq n$ pour tout n. Pour cela posons $\psi(0) = 0$ et, pour tout $n \in \mathbb{N} \setminus \{0\}$, $\psi(n) = \min\{k > \psi(n-1),\ u_k \geq n\}$.
Par construction, si $\psi(n) - 1 > \psi(n-1)$, alors $u_{\psi(n)-1} < n$ et donc $0 \leq u_{\psi(n)} - n \leq u_{\psi(n)} - u_{\psi(n)-1}$.
▹ Montrons, par l'absurde qu'il existe une suite strictement croissante d'indices n tels que $\psi(n) - 1 > \psi(n-1)$. Sinon, il existe N tel que pour tout $n \geq N$, $\psi(n) - 1 = \psi(n-1)$.
Par ailleurs, il existe N' tel que, pour tout $n \geq N'$, $|u_{n+1} - u_n| \leq \frac{1}{2}$.
Notons $m = \max(N, N')$. Alors, pour tout $n \in \mathbb{N}$, $m + n \leq u_{\varphi(m+n)} = u_{\varphi(m)+n} \leq u_{\varphi(m)} + n\frac{1}{2}$, d'où $n \leq 2(u_{\varphi(m)} - m)$: contradiction.

Corrigés des problèmes du chapitre 6

Problème 1

1. a. Soit $x \in \mathbb{R}$ et $r \in \mathbb{Q}$. Notons N le dénominateur de r en écriture irréductible. Pour tout $n \geq N$, $n!\pi r \in \pi\mathbb{Z}$ donc $\tan(n!\pi(x+r)) = \tan(n!\pi x)$ donc la suite $(\tan(n!\pi(x+r)))_n$ converge vers ℓ si, et seulement si, la suite $(\tan(n!\pi x))_n$ converge vers ℓ. Par conséquent, $f(x+r) = f(x)$. En conclusion, la fonction f est r-périodique.
b. Comme $f(0) = 0$ (la suite est constante égale à 0), la restriction de f à $\mathbb{Q}$ est la fonction nulle.
2. a. D'après ses variations, la fonction $x \mapsto \tan(\pi x)$ est une bijection de $[0, \frac{1}{2}[$ dans $\mathbb{R}_+$ et de $]\frac{1}{2}, 1[$ dans $\mathbb{R}_-^*$. Elle réalise donc une bijection de $[0, \frac{1}{2}[\cup]\frac{1}{2}, 1[$ dans $\mathbb{R}$: elle est en particulier surjective et il existe donc $x \in [0, 1[\setminus\{\frac{1}{2}\}$ tel que $y = \tan(\pi x)$.
b. La suite (u_n) est croissante. Par ailleurs, pour tout n,

$$u_n \leq \sum_{k=0}^{n} \frac{1}{k!} \leq e.$$

La suite est majorée donc converge.
c. Pour tout $p > n$,

$$n! \sum_{k=n+1}^{p} \frac{\lfloor kx \rfloor}{k!} = \frac{\lfloor (n+1)x \rfloor}{n+1} + n! \sum_{k=n+2}^{p} \frac{\lfloor kx \rfloor}{k!}$$

et, par télescopage,

$$n! \sum_{k=n+2}^{p} \frac{\lfloor kx \rfloor}{k!} \leq \sum_{k=n+2}^{p} \frac{1}{n(n+1)} = \frac{1}{n+2} - \frac{1}{p+1} \leq \frac{1}{n+2}.$$

En passant à la limite en p,

$$\frac{\lfloor (n+1)x \rfloor}{n+1} \leq n!(\ell - u_n) \leq \frac{\lfloor (n+1)x \rfloor}{n+1} + \frac{1}{n+2}.$$

Désormais, le lemme d'encadrement entraîne que $n!(\ell - u_n) \to x$.
d. Par continuité de la fonction tan en x, $\tan(n!\pi(\ell - u_n)) \to \tan \pi x = y$. Or, comme, pour tout n, $n!u_n \in \mathbb{Z}$, $\tan(n!\pi\ell) \to y$ c'est-à-dire $f(\ell) = y$.
3. Soit I un intervalle non trivial. D'après la question précédente, pour tout $y \in \mathbb{R}$, il existe ℓ tel que $f(\ell) = y$. Par densité de $\ell + \mathbb{Q}$, il existe $r \in \mathbb{Q}$ tel que $\ell + r \in I$. Par la propriété de périodicité de la première question, $f(\ell + r) = f(\ell) = y$. En conclusion, $f : I \to \mathbb{R}$ est surjective.

Problème 2

Partie A – Définition et premiers exemples

1. a. Comme $A \cap [\![1, n]\!] \subset [\![1, n]\!]$, $0 \leq \operatorname{card}(A \cap [\![1, n]\!]) \leq n$. Donc, la limite $\lim\limits_{n\to\infty} \frac{1}{n} \operatorname{card}(A \cap [\![1, n]\!])$, lorsqu'elle existe, appartient à $[0, 1]$.
b. Si A est une partie finie, alors $0 \leq \frac{1}{n} \operatorname{card}(A \cap [\![1, n]\!]) \leq \frac{\operatorname{card} A}{n}$ et donc par encadrement $\theta(A) = 0$.
c. Soit une partie A soit de densité p. Alors, pour tout $n \in \mathbb{N} \setminus \{0\}$,

$$\operatorname{card}(A^{\complement} \cap [\![1, n]\!]) = n - \operatorname{card}(A \cap [\![1, n]\!]).$$

Par conséquent, en divisant par n et en passant à la limite, on trouve que $\theta(A^{\complement}) = 1 - p$.
d. Si A une partie est de densité nulle, alors $A^{\complement}$ est une partie de $\mathbb{N} \setminus \{0\}$ de densité 1 donc ne peut être une partie finie : d'où $A^{\complement}$ est dénombrable.

2. Un rapide calcul donne que :

$$\frac{1}{n}\operatorname{card}(p\mathbb{N}\setminus\{0\}\cap[\![1,n]\!])=\frac{1}{n}\left\lfloor\frac{n}{p}\right\rfloor$$

Or, $\frac{1}{n}(\frac{n}{p}-1)\leq\frac{1}{n}\left\lfloor\frac{n}{p}\right\rfloor\leq\frac{1}{p}$ donc par encadrement $\theta(p\mathbb{N}\setminus\{0\})=\frac{1}{p}$.

3. Soit A l'ensemble des entiers dont l'écriture décimale ne comporte aucun 9. Alors

$$\frac{1}{10^p}\operatorname{card}(A\cap[\![1,10^p]\!])=\left(\frac{9}{10}\right)^p\underset{p\to\infty}{\longrightarrow}0$$

De plus, pour tout $10^{p-1}\leq n<10^p$,

$$\frac{1}{n}\operatorname{card}(A\cap[\![1,n]\!])\leq\frac{1}{10^{p-1}}\operatorname{card}(A\cap[\![1,10^p]\!])=\frac{9^p}{10^{p-1}}$$

Donc $\theta(A)=0$.

4. a. Le produit $\prod\limits_{2n+1\geq p>n+1}p$ divise $(2n+1)!$ et est premier avec $n!$ et $(n+1)!$; d'après le théorème de Gauss, $\prod\limits_{2n+1\geq p>n+1}p$ divise $\binom{2n+1}{n}$.

b. Le résultat est immédiat pour $n\leq 4$.

Soit $n\geq 3$. Supposons que $\prod\limits_{p\leq k}p\leq 4^k$ pour tout $k<n$ et montrons le résultat analogue au rang n.

- Si n est pair (ou plus généralement si n n'est pas un nombre premier),

$$\prod_{p\leq n}p=\prod_{p\leq n-1}p\leq 4^{n-1}\leq 4^n$$

- Si n est impair de la forme $2l+1$, alors on décompose

$$\begin{aligned}\prod_{p\leq n}p &= \prod_{p\leq l+1}p.\prod_{l+1<p\leq n}p\\ &\leq 4^{l+1}\binom{2l+1}{l}\end{aligned}$$

en utilisant l'hypothèse de récurrence au rang $l+1$ et la majoration du a). Remarquons maintenant que

$$2\binom{2l+1}{l}=\binom{2l+1}{l}+\binom{2l+1}{l+1}\leq\sum_{k=0}^{2l+1}\binom{2l+1}{k}=2^{2l+1}$$

D'où l'on déduit : $\prod\limits_{p\leq n}p\leq 4^{l+1}.4^l=4^n$.

Ceci clôt l'étape d'hérédité de la récurrence.

c. Pour tout $m<n$,

$$4^n\geq\prod_{p\leq n}p\geq\prod_{m<p\leq n}p\geq m^{\pi(n)-\pi(m)}$$

En passant au logarithme, on obtient $\pi(n)\leq\frac{n\ln 4}{\ln m}+\pi(m)$.

d. Conformément à l'indication, prenons $m=\left\lfloor\frac{n}{\ln(n)}\right\rfloor$ dans l'inégalité précédente. On obtient alors :

$$\pi(n)\leq\pi\left(\left\lfloor\frac{n}{\ln(n)}\right\rfloor\right)+\frac{n\ln 4}{\ln\left\lfloor\frac{n}{\ln(n)}\right\rfloor}\leq\left\lfloor\frac{n}{\ln(n)}\right\rfloor+\frac{n\ln 4}{\ln\left\lfloor\frac{n}{\ln(n)}\right\rfloor}$$

Par conséquent, $\frac{\pi(n)}{n}\to 0$ et l'ensemble des nombres premiers est de densité 0.

CORRIGÉS

Partie B – Une réciproque du théorème de Cesàro

1. a. Soit $\varepsilon > 0$. Alors,
$$\frac{1}{n+1}\sum_{k=0}^{n}|u_k| \geq \varepsilon \frac{1}{n+1}\operatorname{card}(A_\varepsilon \cap [\![1,n]\!]).$$
Comme la limite du membre de gauche est nulle, le terme de droite l'est aussi et $\theta(A_\varepsilon) = 0$.
b. D'après la question précédente, $\frac{1}{n}p_{k+1}(n) \to 0$. Par conséquent, il existe un rang n_k tel que, pour tout $n \geq n_k$, $\frac{1}{n}p_{k+1}(n) \leq \frac{1}{k}$.
c. Comme $t \mapsto A_t$ est décroissante pour l'inclusion (c'est-à-dire si $s \leq t$, alors $A_t \subset A_s$), on en déduit que, pour tout k,
$$A \cap [\![1,n_k]\!] = \bigcup_{j\leq k}(A_{\frac{1}{j}} \cap [\![n_{j-1},n_j]\!]) \subset A_{\frac{1}{k}} \cap [\![1,n_k]\!].$$
Par conséquent, $\operatorname{card}(A \cap [\![1,n_k]\!]) \leq p_k(n_k) \leq \frac{n_k}{k-1}$ et de manière immédiate, $\lim\limits_{k\to\infty}\frac{1}{n_k}\operatorname{card}(A \cap [\![1,n_k]\!]) = 0$.
Pour $n_{k-1} < n \leq n_k$, on écrit
$$\frac{1}{n}\operatorname{card}(A \cap [\![1,n]\!]) \leq \frac{1}{n}\operatorname{card}(A_{\frac{1}{k}} \cap [\![1,n]\!]) \leq \frac{1}{n}p_k(n) \leq \frac{1}{k-1}.$$
Comme cette inégalité est vraie pour tout k, A est de densité nulle.
d. Soit $m \notin A$. Il existe k tel que $n_{k-1} < m \leq n_k$. Comme $m \notin A$, $m \notin A_{\frac{1}{k}}$ et donc $|u_m| < \frac{1}{k}$. On en déduit que $u_{\varphi(k)} \to 0$.
2. Appliquons la question précédente en remplaçant u_n par $u_n - a$. On en déduit qu'il existe une suite extraite de $(u_n - a)_n$ qui converge vers 0 et dont l'ensemble des indices est de densité 1. D'où le résultat.

Partie C – Application aux séries

1. Comme la fonction $t \mapsto \frac{1}{t}$ est décroissante,
$$\int_k^{k+1}\frac{\mathrm{d}t}{t} \leq \frac{1}{k}$$
En sommant ces inégalités et en utilisant la relation de Chasles, on obtient :
$$\sum_{k=1}^{n}\frac{1}{k} \geq \int_1^{n+1}\frac{\mathrm{d}t}{t} = \ln(n+1)$$
Donc $\left(\sum_{k=1}^{n}\frac{1}{k}\right)_n$ diverge vers $+\infty$.
2. a. En écrivant $u_n = n(S_n - S_{n-1})$, on obtient, pour tout $n \in \mathbb{N}$
$$\begin{aligned}\frac{1}{n}\sum_{k=1}^{n}u_k &= \frac{1}{n}\sum_{k=1}^{n}k(S_k - S_{k-1}) = \frac{1}{n}\left(\sum_{k=1}^{n}kS_k - \sum_{k=0}^{n-1}(k+1)S_k\right)\\ &= S_n - \frac{1}{n}\sum_{k=1}^{n-1}S_k\end{aligned}$$
b. Or $\lim\limits_{n\to\infty}\frac{1}{n}\sum\limits_{k=1}^{n-1}S_k = \lim\limits_{n\to\infty}S_n$ donc $\frac{1}{n}\sum\limits_{k=1}^{n-1}u_k \to 0$.
3. D'après la partie précédente, il existe une suite extraite de $(u_n)_n$ qui converge vers 0 et dont l'ensemble des indices est de densité 1.

Chapitre 7

Fonctions continues

1. Études locales — **2.** Fonctions continues sur un intervalle — **3.** Fonctions continues sur un segment

Objectifs et compétences du programme

- Savoir manipuler les définitions de limites, de limites à droite et à gauche, de continuité.
- Comprendre le caractère local de la continuité.
- Utiliser les suites pour gérer des questions de limites/de continuité.
- Connaître les propriétés des fonctions continues sur un intervalle (notamment le théorème des valeurs intermédiaires).
- Connaître les propriétés des fonctions continues sur un segment (notamment le théorème des bornes atteintes).

Bernard Bolzano, mathématicien tchèque (1781-1848)
Il est l'un de ces oubliés dont on a retrouvé les résultats scientifiques de manière posthume. Prêtre dans l'empire autrichien (plus précisément en Bohème), son activité mathématique peut être essentiellement séparée en deux pans distincts :
• d'une part, son travail sur les « fonctions » qui l'amène à définir la continuité et à adopter une approche rigoureuse, par exemple pour démontrer la propriété des valeurs intermédiaires ;
• d'autre part, un important travail de logicien pour fournir des bases à tous les domaines scientifiques, qui influencera la génération suivante, notamment Georg Cantor ou Richard Dedekind.

COURS

Ce chapitre se concentre sur la notion de fonction continue. Afin de l'aborder correctement, on reprend la notion de limite d'une fonction et on examine les propriétés associées (notamment la caractérisation séquentielle de limite qui permet d'établir un lien fort avec le chapitre précédent). Ensuite, on établit différentes propriétés des fonctions continues (théorème des valeurs intermédiaires, théorème de la bijection, théorème des bornes atteintes, théorème de Heine) en prenant soin de trier ces résultats selon l'hypothèse « topologique » sur l'ensemble de définition : I intervalle ou I segment.

1. Études locales

1.1. Limites

Dans la suite de ce chapitre, I désigne un intervalle réel.

1.1.1. Définitions

Définition 7.1.

Soit $f : I \to \mathbb{R}$ et a une borne ou un point de l'intervalle I.

▹ La fonction f admet la limite $\ell \in \mathbb{R}$ en a si

$$\forall \varepsilon > 0,\ \exists \eta > 0,\ \forall x \in]a-\eta, a+\eta[\cap I, \qquad |f(x)-\ell| \leq \varepsilon.$$

On note alors $f(x) \underset{x\to a}{\to} \ell$.

▹ La fonction f admet la limite $+\infty$ en a si

$$\forall M > 0,\ \exists \eta > 0,\ \forall x \in]a-\eta, a+\eta[\cap I, \qquad f(x) \geq M.$$

On note alors $f(x) \underset{x\to a}{\to} +\infty$.

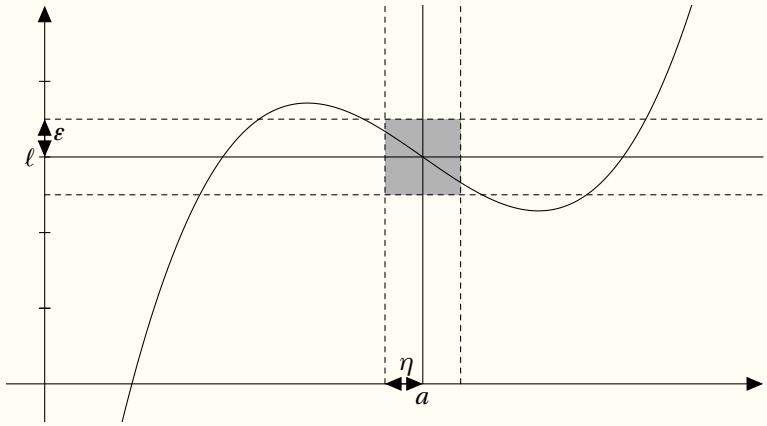
▹ La fonction f admet la limite $-\infty$ en a si

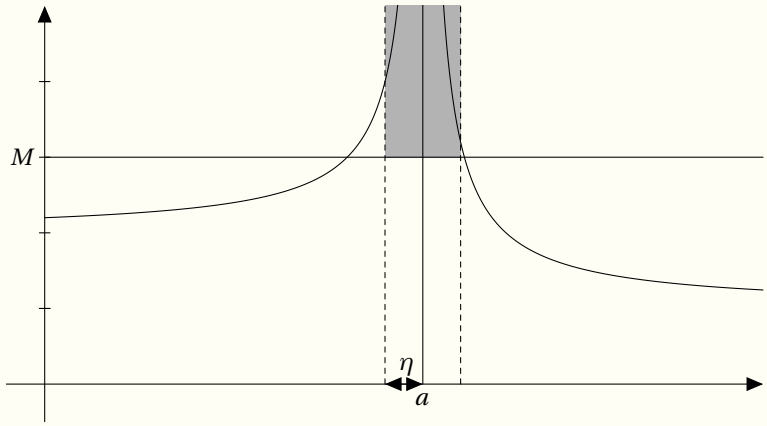
$$\forall m < 0,\ \exists \eta > 0,\ \forall x \in]a-\eta, a+\eta[\cap I, \qquad f(x) \leq m.$$

On note alors $f(x) \underset{x\to a}{\to} -\infty$.

Regardons de plus près ces phrases quantifiées, il y a d'une part la partie qui concerne le voisinage de a (la partie centrale de la phrase quantifiée) et d'autre part la partie qui concerne le voisinage de ℓ, $+\infty$ ou $-\infty$, les termes du début et de la fin de la phrase. On retrouvera cette structure dans les définitions suivantes.

Illustrons les deux premières occurrences de ces définitions. Pour une « zone verticale » fixée (voisinage de ℓ ou de $+\infty$), il existe un voisinage (« horizontal ») de a sur lequel la fonction prend des valeurs dans la zone.





Définition 7.2.

Soit $f : I \to \mathbb{R}$. Supposons que $+\infty$ soit une borne de I.

▹ La fonction f admet la limite $\ell \in \mathbb{R}$ en $+\infty$ si

$$\forall \varepsilon > 0,\ \exists A > 0,\ \forall x \geq A, \qquad |f(x) - \ell| \leq \varepsilon.$$

▹ La fonction f admet la limite $+\infty$ en $+\infty$ si

$$\forall M > 0,\ \exists A > 0,\ \forall x \geq A, \qquad f(x) \geq M.$$

▹ La fonction f admet la limite $-\infty$ en $+\infty$ si

$$\forall m < 0,\ \exists A > 0,\ \forall x \geq A, \qquad f(x) \leq m.$$

Remarque

Les définitions de limite finie se généralisent immédiatement aux fonctions à valeurs complexes $f : I \to \mathbb{C}$ en remplaçant la valeur absolue par le module.
En revanche, il n'est plus possible de définir « admettre la limite $+\infty$ ou $-\infty$ » pour de telles fonctions puisque $\mathbb{C}$ n'est pas muni d'une relation d'ordre naturelle.

Remarquons tout de suite l'unicité de la limite.

Proposition 7.3.

La limite d'une fonction est, quand elle existe, unique. La limite de f en a sera notée $\lim\limits_{x\to a} f(x)$ ou $\lim\limits_{a} f$.

Démonstration

Raisonnons par l'absurde en supposant qu'il existe deux limites en a distinctes ℓ et ℓ' et appliquons les définitions de limites avec $\varepsilon = \frac{|\ell-\ell'|}{4}$. Il existe alors $\eta > 0$ et $\nu > 0$ tels que

$$\forall x \in]a-\eta, a+\eta[\cap I \qquad |f(x)-\ell| \leq \varepsilon,$$
$$\forall x \in]a-\nu, a+\nu[\cap I \qquad |f(x)-\ell'| \leq \varepsilon.$$

Ainsi, pour tout $x \in]a-\min(\eta,\nu), a+\min(\eta,\nu)[\cap I$, on obtient la contradiction suivante

$$|\ell-\ell'| \leq |f(x)-\ell| + |f(x)-\ell'| \leq 2\varepsilon = \frac{|\ell-\ell'|}{2}.$$

■

1.1.2. Limites à gauche et à droite

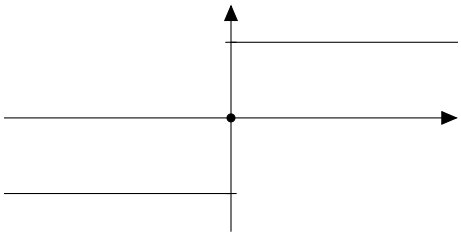
Les limites n'existent pas toujours, aussi on utilise quelquefois les définitions suivantes plus faibles.

Définition 7.4.

Soit $f : I \to \mathbb{R}$ et a une borne ou un point de l'intervalle I. La fonction f admet une limite à gauche en a (respectivement à droite) si la fonction $f_{|]-\infty,a[\cap I}$ (respectivement $f_{|]a,+\infty[\cap I}$) admet une limite en a.
Lorsqu'elles existent, ces limites sont notées $\lim\limits_{x\to a^-} f(x) = \lim\limits_{a^-} f$ et $\lim\limits_{x\to a^+} f(x) = \lim\limits_{a^+} f$.

Exemple

Considérons la fonction « signe » définie par $f : x \mapsto \begin{cases} 1 & \text{si } x > 0, \\ -1 & \text{si } x < 0, \\ 0 & \text{si } x = 0. \end{cases}$



Cette fonction n'admet pas de limite en 0 ; en revanche, $\lim\limits_{0^-} f = -1$ et $\lim\limits_{0^+} f = 1$.

Remarque

Si une fonction f admet une limite en a, alors elle y admet des limites à droite et à gauche égales. La réciproque est fausse (il suffit de considérer la fonction constante égale à 1 sauf en un point).

Les limites à gauche et à droite permettent de donner le résultat suivant de régularité.

Proposition 7.5. Limite monotone

Soit $f :]a,b[\to \mathbb{R}$ croissante. Alors, f admet des limites finies à gauche et à droite en tout point $c \in]a,b[$. De plus,

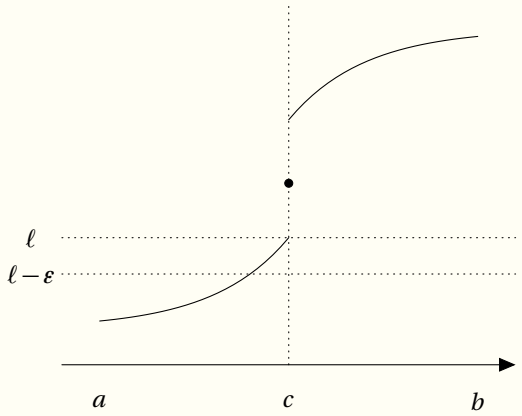
$$\lim_{c^-} f \leq f(c) \leq \lim_{c^+} f.$$

Il existe évidemment un analogue pour les fonctions décroissantes (en changeant le sens des inégalités).

Démonstration

Soit $c \in]a,b[$.

▹ Notons $\ell = \sup\{f(x),\ x < c\}$. Pour tout $\varepsilon > 0$, $\ell - \varepsilon$ n'est pas un majorant de $\{f(x),\ x < c\}$ donc il existe $y < c$ tel que $\ell - \varepsilon \leq f(y)$.



Par croissance de f, on obtient que, pour tout $z \in [y,c[$, $\ell - \varepsilon \leq f(z) \leq \ell$. Par conséquent,

$$\lim_{c^-} f = \sup\{f(x),\ x < c\}.$$

▹ On montre de même que $\lim_{c^+} f = \inf\{f(x),\ x > c\}$. ■

Remarque

Ce résultat peut se redire moins formellement : une fonction croissante n'admet des discontinuités que sous forme de « sauts » vers le haut.

Les résultats aux bornes de l'intervalle se démontrent de manière analogue quitte à considérer les bornes supérieure et inférieure dans $\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$.

Proposition 7.6.

Soit $f :]a,b[\to \mathbb{R}$ croissante. Alors, f admet une limite à gauche en b (finie ou $+\infty$) et une limite à droite en a (finie ou $-\infty$).

1.1.3. Caractère localement borné

Proposition 7.7.

Si $f : I \to \mathbb{R}$ admet une limite finie en $a \in I$, alors il existe un intervalle ouvert J contenant a tel que f est bornée sur $I \cap J$.

Démonstration
Appliquons la définition de limite avec $\varepsilon = 1$: il existe $\eta > 0$ tel que

$$\forall x \in]a-\eta, a+\eta[\cap I, \qquad \lim_a f - 1 \leq f(x) \leq \lim_a f + 1.$$

■

En fait, on peut être un peu plus fin dans l'utilisation de la définition de limite finie non nulle.

Proposition 7.8.

Si $f : I \to \mathbb{R}$ admet une limite finie non nulle en a, alors il existe un intervalle ouvert J contenant a tel que f est de signe constant sur $I \cap J$.

Démonstration
Étudions le cas où $\lim_a f > 0$. Il suffit d'appliquer la définition avec $\varepsilon = \frac{1}{2} \lim_a f$; en effet, il existe alors $\eta > 0$ tel que

$$\forall x \in]a-\eta, a+\eta[\cap I, \qquad \lim_a f - \frac{1}{2} \lim_a f \leq f(x) \leq \lim_a f + \frac{1}{2} \lim_a f.$$

En particulier, pour tout $x \in]a-\eta, a+\eta[\cap I$, $f(x) \geq \frac{1}{2} \lim_a f > 0$. ■

Comme pour les suites, il existe une généralisation avec un intervalle ouvert.

Proposition 7.9.

Si $f : I \to \mathbb{R}$ admet en a une limite ℓ dans un intervalle ouvert $]\alpha, \beta[$, alors il existe un intervalle ouvert J contenant a sur lequel f est à valeurs dans $]\alpha, \beta[$.

1.1.4. Aspects séquentiels

Passons maintenant aux aspects séquentiels de la définition de limite.

Proposition 7.10.

Soit $f : I \to \mathbb{R}$, a une borne ou un point de l'intervalle I et $\ell \in \overline{\mathbb{R}}$. f admet la limite ℓ en a si, et seulement si, pour toute une suite $(u_n)_n$ d'éléments de I convergeant vers a, la suite $(f(u_n))_n$ converge vers ℓ.

Démonstration
On effectue la démonstration dans le cas $a \in \mathbb{R}$.
($\Rightarrow$) Soit $\varepsilon > 0$. Il existe $\eta > 0$ tel que, pour tout $x \in]a-\eta, a+\eta[\cap I, |f(x) - \ell| \leq \varepsilon$.
Comme $(u_n)_n$ converge vers a, il existe $N \in \mathbb{N}$ tel que, pour tout $n \geq N, |u_n - a| \leq \eta$.

En combinant ces deux résultats, on obtient, pour tout $n \geq N, |f(u_n) - \ell| \leq \varepsilon$
Donc $(f(u_n))_n$ converge vers ℓ.
($\Leftarrow$) Raisonnons par l'absurde en supposant que f n'admet pas ℓ comme limite en a :

$$\exists \varepsilon > 0,\ \forall \eta > 0,\ \exists x \in I \cap]a-\eta, a+\eta[\qquad |f(x) - \ell| > \varepsilon.$$

Ainsi, en prenant, $\eta = \frac{1}{n}$, on définit un élément $x_n \in I \cap]a - \frac{1}{n}, a + \frac{1}{n}[$. La suite $(x_n)_n$ d'éléments de I converge vers a, et pourtant la suite $(f(x_n))_n$ ne converge pas vers ℓ : contradiction.

■

Conseils méthodologiques

On utilise cette propriété pour montrer qu'une fonction f n'admet pas de limite en a en exhibant deux suites $(u_n)_n$ et $(v_n)_n$ de même limite a telles que les suites images $(f(u_n))_n$ et $(f(v_n))_n$ admettent des limites différentes (ou même que l'une n'admet pas de limite).

Exemple

La fonction $f : x \mapsto \cos \frac{1}{x}$ n'admet pas de limite en 0. En effet, les suites $u = \left(\frac{1}{2n\pi}\right)_n$ et $v = \left(\frac{1}{2n\pi + \frac{\pi}{2}}\right)_n$ tendent toutes les deux vers 0 et $f(u_n) \to 1$ alors que $f(v_n) \to 0$.

1.2. Continuité

Définition 7.11. Continuité en un point

Soit $f : I \to \mathbb{R}$ et $a \in I$.
▹ La fonction f est continue en a si $\lim\limits_{a} f = f(a)$.
▹ La fonction f est continue à gauche en a si $\lim\limits_{a-} f = f(a)$.
▹ La fonction f est continue à droite en a si $\lim\limits_{a+} f = f(a)$.

La définition pour les fonctions à valeur dans $\mathbb{C}$ est exactement la même.

Remarque

En fait, il y a redondance dans la définition de la continuité en un point a : il suffit en fait de dire f est continue en a si f admet une limite finie en a et cette limite sera, par définition égale à $f(a)$.

Définition 7.12. Continuité

Une fonction $f : I \to \mathbb{R}$ est continue sur I si f est continue en x pour tout $x \in I$.

Les fonctions cos, sin, polynômes et puissances sont continues sur leurs domaines de définition.

Exemple Fonctions lipschitziennes

Soit $k \in \mathbb{R}_+$. Une fonction $f : I \to \mathbb{R}$ est k-lipschitzienne si

$$\forall x, x' \in I, \qquad |f(x) - f(x')| \leq k|x - x'|.$$

▹ Montrons que toute fonction lipschitzienne sur I est continue en $a \in I$ en prenant $x' = a$ dans la définition.

Pour tout $\varepsilon > 0$, posons $\eta = \frac{\varepsilon}{k}$. Alors, pour tout $x \in I$ tel que $|x-a| \leq \eta$,

$$|f(x)-f(a)| \leq k|x-a| \leq k\eta = \varepsilon.$$

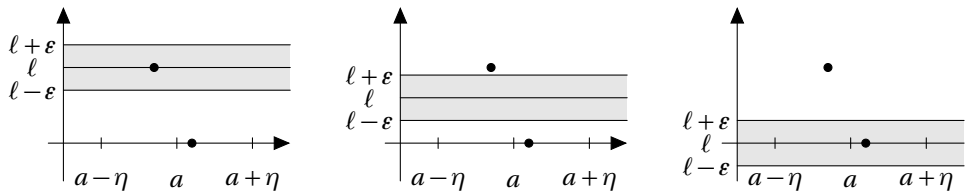
▹ En revanche, la réciproque est fausse : par exemple $x \mapsto x^2$ est continue sur $\mathbb{R}$ mais n'est pas lipschitzienne ; si elle l'était, on aurait entre 0 et $x > 0$, $x^2 \leq kx$ ce qui devient contradictoire quand x devient grand.

Exemple

La fonction indicatrice des rationnels définie par

$$x \mapsto \begin{cases} 1 & \text{si } x \in \mathbb{Q}, \\ 0 & \text{sinon} \end{cases}$$

n'est continue nulle part.



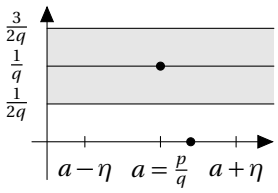
▹ En effet, tout intervalle $]a-\eta, a+\eta[$ contient aussi bien des rationnels que des irrationnels : la fonction y prend donc les valeurs 0 et 1 ce qui contredit la définition de limite finie ℓ en a (comme on peut le voir sur les schémas ci-dessus sur lesquels on a distingué les cas $\ell = 1$, $\ell \notin \{0,1\}$ et $\ell = 0$) avec $\varepsilon < \frac{1}{2}$.

Exemple

La fonction

$$f : x \mapsto \begin{cases} 0 & \text{si } x \notin \mathbb{Q}, \\ 1 & \text{si } x = 0, \\ \frac{1}{q} & \text{si } x = \frac{p}{q} \text{ avec } p \wedge q = 1,\ q > 0 \end{cases}$$

n'est continue que sur $\mathbb{R} \setminus \mathbb{Q}$.
▹ En effet, f est discontinue en tout rationnel car dans le voisinage de tout rationnel $a = \frac{p}{q}$, il y a des irrationnels donc f y prend la valeur 0, ce qui contredit la définition de limite finie en a avec $\varepsilon < \frac{1}{q}$.



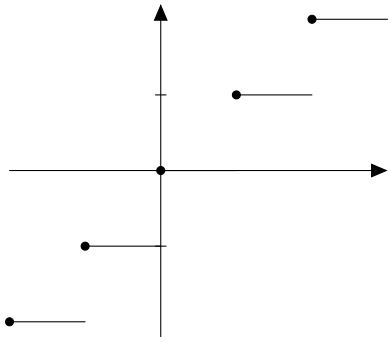
▹ Soit $a \in \mathbb{R} \setminus \mathbb{Q}$ et $\varepsilon > 0$. L'ensemble $A_\varepsilon = \{x \in]a-1, a+1[,\ f(x) \geq \varepsilon\}$ est fini (car les dénominateurs sont majorés par $\frac{1}{\varepsilon}$) ; notons $\eta = \frac{1}{2}\min\{|x-a|,\ x \in A_\varepsilon\}$. Alors, pour tout $x \in \mathbb{R}$ tel que $|x-a| \leq \eta$, $f(x) = 0$ (si $x \notin \mathbb{Q}$) ou $f(x) \leq \varepsilon$ (si $x \in \mathbb{Q}$ d'après la définition de η) donc

$$|f(x)-f(a)| \leq \varepsilon,$$

ce qui est la définition de la continuité de f en a.

Exemple

La fonction partie entière $x \mapsto \lfloor x \rfloor$ est continue sur $\mathbb{R} \setminus \mathbb{Z}$ et continue à droite mais pas à gauche en tout entier.



Remarque

Comme la continuité est une propriété locale définie par des limites de fonctions, on peut utiliser la caractérisation séquentielle des limites. En fait, il s'agit surtout d'un outil pour montrer qu'une fonction n'est pas continue.

Exemple

Soit $f : I \to \mathbb{R}$ continue et $(u_n)_n$ une suite vérifiant la relation de récurrence $u_{n+1} = f(u_n)$. Alors, si la suite $(u_n)_n$ converge vers $\ell \in I$, on a $\ell = f(\ell)$, c'est-à-dire que ℓ est un point fixe de f. Pour étudier les suites récurrentes d'ordre 1, on peut donc procéder en deux temps : montrer la convergence (par exemple avec de la monotonie) puis déterminer les limites possibles en cherchant les points fixes de la fonction.

Définition 7.13.

Une fonction continue $f : I \setminus \{a\} \to \mathbb{R}$ où $a \in I$ est prolongeable par continuité s'il existe $b \in \mathbb{R}$ tel que la fonction porlongée

$$\tilde{f} : x \mapsto \begin{cases} b & \text{si } x = a \\ f(x) & \text{si } x \in I \setminus \{a\} \end{cases}$$

est continue.

Exemple

La fonction sinus cardinal définie sur $\mathbb{R}^*$, par $x \mapsto \frac{\sin(x)}{x}$ admet la limite 1 en 0. Elle est donc prolongeable par continuité en 0.

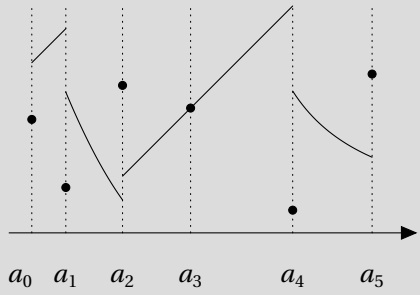
D'après la définition de limite, on obtient la caractérisation suivante pour le prolongement en un point intérieur.

Proposition 7.14.

Soit I un intervalle, a un point intérieur à I et $f : I \setminus \{a\} \to \mathbb{R}$ une fonction continue. La fonction f est prolongeable par continuité en a si, et seulement si, f admet des limites finies à gauche et à droite en a et que $\lim\limits_{a^-} f = \lim\limits_{a^+} f$.

Définition 7.15.

▹ Une fonction f est continue par morceaux sur le segment $[a,b]$ s'il existe une subdivision $a = a_0 < \ldots < a_n = b$ telle que, pour tout $k \in [\![1,n]\!]$, la restriction $f_{|]a_{k-1},a_k[}$ admette un prolongement continu sur $[a_{k-1},a_k]$.



▹ Une fonction f est continue par morceaux sur l'intervalle I si elle est continue par morceaux sur tout segment inclus dans I.

Remarque

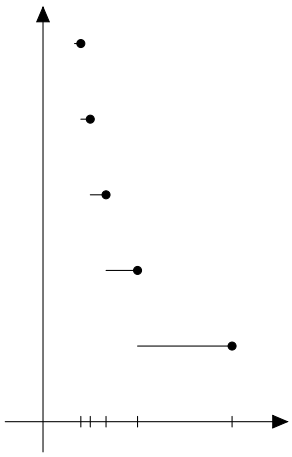
D'après la proposition précédente, une fonction est continue par morceaux sur un segment s'il existe une subdivision telle que

- la fonction est continue en dehors des points de la subdivision,
- la fonction admet des limites finies à gauche et à droite aux points de la subdivision.

En particulier, elle n'admet qu'un nombre fini de discontinuités (mais cette condition ne suffit pas).

Exemples

La fonction partie entière $x \mapsto \lfloor x \rfloor$ est continue par morceaux sur $\mathbb{R}$ avec une subdivision aux points d'abscisse entière (elle est même constante par morceaux).
La fonction $x \mapsto \left\lfloor \frac{1}{x} \right\rfloor$ est continue par morceaux sur $]0,1]$. dont voici le tracé de la courbe représentative entre $\frac{1}{5}$ et 1 :



En effet, pour tout $x \in]\frac{1}{k+1}, \frac{1}{k}]$, la partie entière $\lfloor \frac{1}{x} \rfloor$ vaut k.

1.3. Manipulations sur les limites

1.3.1. Manipulations algébriques

Proposition 7.16.

Soit $f, g : I \to \mathbb{R}$ et a une borne ou un point de l'intervalle I. On suppose que f et g admettent une limite finie en a. Alors,

▹ $|f|$ admet une limite finie en a et $\lim\limits_a |f| = |\lim\limits_a f|$.

▹ Pour tous $\lambda, \mu \in \mathbb{R}$, $\lambda f + \mu g$ admet une limite finie en a et

$$\lim_a (\lambda f + \mu g) = \lambda \lim_a f + \mu \lim_a g.$$

▹ $f.g$ admet une limite finie en a et $\lim\limits_a f.g = \lim\limits_a f . \lim\limits_a g$.

▹ Si $\lim\limits_a g \neq 0$, alors $\frac{f}{g}$ admet une limite finie en a et $\lim\limits_a \frac{f}{g} = \frac{\lim\limits_a f}{\lim\limits_a g}$.

Démonstration

Pour la démonstration, notons $\ell = \lim\limits_a f$ et $\ell' = \lim\limits_a g$. Écrivons pour chacune des assertions ci-dessus la majoration qui permet de conclure avec les définitions de limites en a de f et g.

▹ D'après l'inégalité triangulaire,

$$\forall x \in I, \qquad ||f(x)| - |\ell|| \leq |f(x) - \ell|.$$

On conclut avec la définition de limite de f en a. Plus précisément, pour tout $\varepsilon > 0$, il existe $\eta > 0$ tel que

$$\forall x \in I \cap]a - \eta, a + \eta[, \qquad |f(x) - \ell| \leq \varepsilon$$

donc tel que

$$\forall x \in I \cap]a - \eta, a + \eta[, \qquad ||f(x)| - |\ell|| \leq \varepsilon.$$

▹ D'après l'inégalité triangulaire,

$$\forall x \in I, \qquad |\lambda f(x) + \mu g(x) - \lambda \ell - \mu \ell'| \leq |\lambda||f(x) - \ell| + |\mu||g(x) - \ell'|,$$

et on conclut comme précédemment avec les définitions de limites de f et g.

▹ Remarquons que $|f|$ est bornée sur un intervalle ouvert contenant a par une constante M. Alors, par inégalité triangulaire, on a sur cet intervalle

$$\begin{aligned} |f(x)g(x) - \ell\ell'| &\leq |f(x)g(x) - f(x)\ell'| + |f(x)\ell' - \ell\ell'| \\ &\leq M|g(x) - \ell'| + |\ell'||f(x) - \ell|, \end{aligned}$$

et on conclut encore avec les définitions de limites de f et g.

▹ Remarquons que $|g|$ est minorée sur un intervalle ouvert contenant a par une constante $m > 0$. Alors, par inégalité triangulaire, on a sur cet intervalle

$$\left| \frac{f(x)}{g(x)} - \frac{\ell}{\ell'} \right| \leq \frac{|\ell' f(x) - \ell g(x)|}{m|\ell'|},$$

et on conclut toujours avec les définitions de limites de f et g.

■

Remarque

La démonstration ci-dessus ressemble beaucoup aux démonstrations analogues sur les limites de suites. C'est tout à fait cohérent puisque ces résultats peuvent également être établi avec la caractérisation séquentielle de limite en partant d'une suite quelconque à valeurs dans I et de limite a.

Avec les définitions de continuité en termes de limites, on obtient que la valeur absolue, les combinaisons linéaires et le produit de fonctions continues en a (respectivement sur I) sont encore des fonctions continues en a (respectivement sur I). Pour le quotient, il faut l'hypothèse supplémentaire que le dénominateur ne s'annule pas.

1.3.2. Composition

Proposition 7.17.

Soit $f : I \to J$ et $g : J \to K$ deux fonctions et a une borne ou un point de l'intervalle I tel que f admet la limite b en a et g admet la limite ℓ en b. Alors, $g \circ f$ admet la limite ℓ en a.

Démonstration

Soit $\varepsilon > 0$. D'après la définition de $\lim\limits_b g = \ell$, il existe $\eta > 0$ tel que

$$\forall y \in J \cap]b-\eta, b+\eta[, \qquad |g(y)-\ell| \leq \varepsilon.$$

Appliquons maintenant la définition de $\lim\limits_a f = b$: il existe $\nu > 0$ tel que

$$\forall x \in I \cap]a-\nu, a+\nu[, \qquad |f(x)-b| \leq \eta.$$

En combinant les deux propriétés,

$$\forall x \in I \cap]a-\nu, a+\nu[, \qquad |g(f(x))-\ell| \leq \varepsilon.$$

■

Remarque

Là encore, on pourrait effectuer la démonstration à partir de la caractérisation séquentielle de limite.

Avec les définitions de continuité en termes de limites, on obtient que la composée de deux fonctions continues est continue.

1.3.3. Inégalités

Toujours avec la caractérisation séquentielle de limite et les propositions analogues pour les suites, on obtient les résultats suivants d'encadrement.

Proposition 7.18.

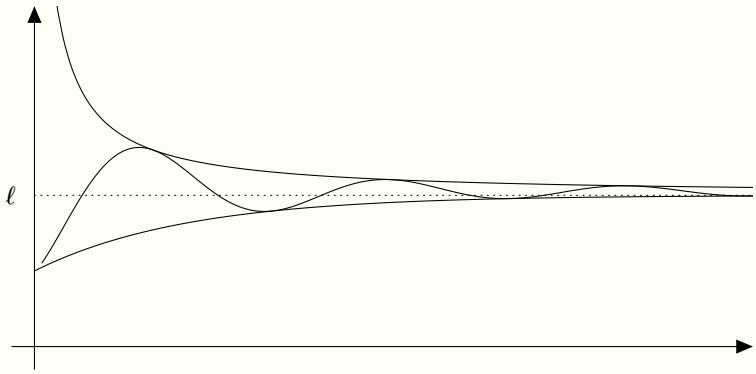
Soit $f, g : I \to \mathbb{R}$, $M \in \mathbb{R}$ et a une borne ou un point de l'intervalle I. Supposons que f et g admettent des limites finies en a.

- ▹ Si, sur un intervalle ouvert contenant a, $f(x) \leq M$, alors $\lim\limits_a f \leq M$.
- ▹ Si, sur un intervalle ouvert contenant a, $f(x) < M$, alors $\lim\limits_a f \leq M$.
- ▹ Si, sur un intervalle ouvert contenant a, $f(x) \leq g(x)$, alors $\lim\limits_a f \leq \lim\limits_a g$.

Proposition 7.19. Lemme d'encadrement

Soit $f, g, h : I \to \mathbb{R}$ et a une borne ou un point de l'intervalle I. Supposons que sur un voisinage de a dans I, on a $f(x) \leq g(x) \leq h(x)$ et que $\lim\limits_a f = \lim\limits_a h = \ell$. Alors, g admet une limite en a et $\lim\limits_a g = \ell$.

Dans le schéma ci-dessous, la fonction dont la courbe représentative oscille est « coincée » entre deux fonctions de même limite ℓ en $+\infty$. Elle admettra elle-aussi cette limite par le lemme d'encadrement.



Des résultats d'encadrement, on déduit le corollaire suivant.

Corollaire 7.20.

La limite du produit d'une fonction bornée par une fonction tendant vers 0 est nulle.

Exemple

En application directe du corollaire précédent avec $x \mapsto x$ de limite nulle en 0 et $x \mapsto \sin \frac{1}{x}$ bornée, on obtient $\lim\limits_{x \to 0} x \sin \frac{1}{x} = 0$.

1.4. Équations fonctionnelles

Une équation fonctionnelle est simplement une équation dont l'inconnue est une fonction. Elle est en général la conjonction d'une relation entre plusieurs valeurs de la fonction inconnue et d'une condition de « régularité » (limite, continuité, monotonie...).
Cette section n'ambitionne pas de traiter une théorie générale mais bien de fournir quelques exemples élémentaires afin de développer une intuition.

Proposition 7.21.

Les fonctions $f : \mathbb{R}_+^* \to \mathbb{R}$ admettant une limite finie ℓ en $+\infty$ et vérifiant

$$\forall x \in \mathbb{R}, \qquad f(x) = f(2x)$$

sont les fonctions constantes.

Démonstration

Remarquons que, pour tout $x \in \mathbb{R}$, une récurrence immédiate donne $f(x) = f(2^n x)$ pour tout $n \in \mathbb{N}$. En passant à la limite en n, on obtient que $f(x) = \ell$. Ainsi, f est constante égale à ℓ. ■

Proposition 7.22.

Les fonctions $f : \mathbb{R} \to \mathbb{R}$ continues en 0 et vérifiant

$$\forall x \in \mathbb{R}, \qquad f(x) = f(2x)$$

sont les fonctions constantes.

Démonstration
Remarquons que, pour tout x, une récurrence immédiate donne $f(x) = f(2^{-n}x)$ pour tout $n \in \mathbb{N}$. En passant à la limite en n et en utilisant la continuité de f en 0, on obtient que $f(x) = f(0)$. Ainsi f est constante égale à $f(0)$. ■

Dans les deux exemples précédents, on a exhibé une suite de réels ayant même image par f : d'un côté la suite $(2^n x)_n$ et de l'autre $(2^{-n}x)_n$. Cette façon de procéder est relativement générale même si l'on n'explicite pas toujours la suite considérée.

Proposition 7.23.

Les fonctions $f : \mathbb{R} \to \mathbb{R}$ continues et vérifiant

$$\forall x \in \mathbb{R}, \qquad f(x) = f\left(\frac{1+x}{2}\right)$$

sont les fonctions constantes.

Démonstration
Soit $x \in \mathbb{R}$. Considérons la suite $(x_n)_n$ définie par $x_0 = x$ et, pour tout $n \in \mathbb{N}$, $x_{n+1} = \frac{1+x_n}{2}$. Alors, une récurrence immédiate assure que $f(x) = f(x_n)$ pour tout $n \in \mathbb{N}$. Comme la suite arithmético-géométrique $(x_n)_n$ converge vers 1 et que la fonction f est continue en 1, on en déduit que $f(x) = f(1)$. En conclusion la fonction f inconnue est constante égale à $f(1)$. ■

On peut déduire de ces premiers exemples les solutions d'un certain nombre d'équations proches.

Proposition 7.24.

Les fonctions $f : \mathbb{R} \to \mathbb{R}$ continues en 0 et vérifiant

$$\forall x \in \mathbb{R}, \qquad f(x) + x = f(2x)$$

sont de la forme $x \mapsto c + x$ avec $c \in \mathbb{R}$.

Démonstration
Remarquons que l'équation se réécrit

$$\forall x \in \mathbb{R}, \qquad f(x) - x = f(2x) - 2x.$$

Considérons $g : x \mapsto f(x) - x$. Elle est continue et vérifie

$$\forall x \in \mathbb{R}, \qquad g(x) = g(2x)$$

donc est constante. En conclusion, f est de la forme $x \mapsto c + x$ avec $c \in \mathbb{R}$. ■

Passons à une autre équation pour illustrer l'importance de choisir des valeurs particulières des arguments.

Proposition 7.25.

La seule fonction $f : \mathbb{R} \to \mathbb{R}$ vérifiant

$$\forall x, y \in \mathbb{R}, \qquad f(xy) = xf(x) + yf(y)$$

est la fonction nulle.

Démonstration
▹ Commençons par remarquer que pour $x = y = 0$, la relation donne $f(0) = 0$.

▹ Ensuite pour $y=0$, on obtient que, pour tout x, $xf(x)=0$. Ainsi, f est nulle sur $\mathbb{R}^*$. En conclusion, la fonction f est nulle sur $\mathbb{R}$. ■

Après cet exemple élémentaire, passons à la plus « importante » des équations fonctionnelles classiques.

Proposition 7.26. Équation de Cauchy

Les fonctions $f:\mathbb{R}\to\mathbb{R}$ continues et vérifiant

$$\forall x,y\in\mathbb{R},\qquad f(x+y)=f(x)+f(y)$$

sont les fonctions linéaires.

Démonstration

▹ Appliquons la relation avec $x=y=0$: $f(0)=2f(0)$ donc $f(0)=0$.

▹ Soit $x\in\mathbb{R}$. On montre, par récurrence sur $n\in\mathbb{N}$, que $f(nx)=nf(x)$. Le résultat est vrai pour $n=0$. Soit $n\in\mathbb{N}$ tel que $f(nx)=nf(x)$. Alors,

$$f((n+1)x)=f(nx+x)=f(nx)+f(x)=nf(x)+f(x)=(n+1)f(x)$$

ce qui achève la récurrence.

▹ Soit $n\in\mathbb{N}$ et $x\in\mathbb{R}$; alors, $0=f(0)=f(nx-nx)=f(nx)+f(-nx)$ donc $f(-nx)=-nf(x)$.

Par conséquent, $f(nx)=nf(x)$ pour tout $n\in\mathbb{Z}$ et tout $x\in\mathbb{R}$.

▹ Soit $(p,q)\in\mathbb{Z}\times\mathbb{N}\setminus\{0\}$. Alors, en appliquant le point précédent avec $x=\frac{1}{q}$ et $n\in\{p,q\}$,

$$f(1)=f\Big(q\frac{1}{q}\Big)=qf\Big(\frac{1}{q}\Big)$$
$$f\Big(\frac{p}{q}\Big)=f\Big(p\frac{1}{q}\Big)=pf\Big(\frac{1}{q}\Big).$$

En combinant ces résultats, $f(\frac{p}{q})=\frac{p}{q}f(1)$, c'est-à-dire $f(r)=rf(1)$ pour tout $r\in\mathbb{Q}$.

▹ Soit $x\in\mathbb{R}$ et $(r_n)_n$ une suite de rationnels de limite x. Comme f est continue,

$$f(x)=\lim f(r_n)=\lim r_n f(1)=xf(1).$$

■

Conseils méthodologiques

La dernière étape peut être modifiée pour chercher les fonctions vérifiant la même équation mais avec la condition de régularité « f croissante ». En effet, dans cet exemple alternatif, les premiers points (construction jusqu'aux rationnels) restent identiques. Le passage à la limite s'obtient en encadrant une réel x quelconque par deux suites de rationnels de limite x et en utilisant le lemme d'encadrement.

Attention à ne pas conclure trop vite dans cet exemple que les fonctions solutions sont les fonctions $x\mapsto ax$ avec $a\in\mathbb{R}$; la vérification (qui était immédiate dans les exemples précédents) impose que $a\geq 0$ pour avoir la croissance.

Corollaire 7.27.

Les fonctions $f:\mathbb{R}_+^*\to\mathbb{R}$ continues et vérifiant

$$\forall x,y\in\mathbb{R}_+^*,\qquad f(xy)=f(x)+f(y)$$

sont les fonctions proportionnelles à ln.

Démonstration
Posons $g : t \mapsto f(e^t)$.
La relation devient
$$\forall s, t \in \mathbb{R}, \qquad g(t+s) = g(t) + g(s).$$
Comme g est continue, il existe, d'après la proposition précédente sur l'équation de Cauchy, $a \in \mathbb{R}$ tel que $g : t \mapsto at$ soit en revenant à f, tel que
$$f : x \mapsto a \ln x.$$
■

Remarque
Avec une méthode analogue, on peut montrer que les seules fonctions continues non nulles vérifiant
$$\forall x, y \in \mathbb{R}, \qquad f(x+y) = f(x)f(y)$$
sont les fonctions de la forme $x \mapsto e^{ax}$ avec $a \in \mathbb{R}$.

Reprenons la construction par étapes des propositions précédentes pour une autre équation.

Proposition 7.28.

Les fonctions $f : \mathbb{R} \to \mathbb{R}$ continues et vérifiant
$$\forall x, y \in \mathbb{R}, \qquad f(x+y) + f(x-y) = 2(f(x) + f(y))$$
sont les fonctions de la forme $x \mapsto ax^2$ avec $a \in \mathbb{R}$.

Démonstration
▹ Appliquons la relation avec $x = y = 0$: $2f(0) = 4f(0)$ donc $f(0) = 0$.
▹ Montrons, par récurrence sur $n \in \mathbb{N}$ et tout $x \in \mathbb{R}$, que $f(nx) = n^2 f(x)$. Le résultat est désormais établi pour $n = 0$ et $n = 1$. Soit $n \geq 1$ tel que la propriété soit vérifiée aux rangs n et $n-1$; montrons-la au rang $n+1$. Soit $x \in \mathbb{R}$. En appliquant la relation avec les arguments nx et x, on obtient
$$\begin{aligned} f((n+1)x) &= 2(f(nx) + f(x)) - f((n-1)x) \\ &= (2n^2 + 2 - (n-1)^2) f(x) = (n+1)^2 f(x). \end{aligned}$$
▹ Soit $x \in \mathbb{R}$ et $n \in \mathbb{N}$. En appliquant la relation avec les arguments 0 et nx,
$$f(nx) + f(-nx) = 2f(0) + 2f(nx)$$
d'où
$$f(-nx) = f(nx) = n^2 f(x) = (-n)^2 f(x).$$
▹ Soit $(p, q) \in \mathbb{Z} \times \mathbb{N} \setminus \{0\}$. Alors, en appliquant le point précédent avec $x = \frac{1}{q}$ et $n \in \{p, q\}$,
$$\begin{aligned} f(1) &= f\Big(q\frac{1}{q}\Big) = q^2 f\Big(\frac{1}{q}\Big) \\ f\Big(\frac{p}{q}\Big) &= f\Big(p\frac{1}{q}\Big) = p^2 f\Big(\frac{1}{q}\Big). \end{aligned}$$
En combinant ces résultats, $f(\frac{p}{q}) = \frac{p^2}{q^2} f(1)$, c'est-à-dire $f(r) = r^2 f(1)$ pour tout $r \in \mathbb{Q}$.
▹ Soit $x \in \mathbb{R}$ et $(r_n)_n$ une suite de rationnels de limite x. Comme f est continue,
$$f(x) = \lim f(r_n) = \lim r_n^2 f(1) = x^2 f(1).$$
■

Voici une équation dont les solutions se déduisent de l'équation précédente.

Corollaire 7.29.

Les fonctions $f : \mathbb{R} \to \mathbb{R}$ continues et vérifiant

$$\forall x, y \in \mathbb{R}, \qquad f(x+y)f(x-y) = f(x)^2 f(y)^2$$

sont les fonctions $x \mapsto 0$, $x \mapsto e^{ax^2}$ ou $x \mapsto -e^{ax^2}$ avec $a \in \mathbb{R}$.

Démonstration

La fonction nulle convient. Soit f une solution non nulle.

▹ Remarquons que si $f(0) = 0$, alors (en prenant $y = 0$ dans la relation),

$$f(x)^2 = f(x)^2 f(0)^2 = 0$$

pour tout $x \in \mathbb{R}$: f est nulle.

Par conséquent $f(0) \neq 0$.

▹ Considérons un réel x_0 où f s'annule. Alors, $f(x_0)f(0) = f(\frac{x_0}{2})^4$ donc $f(\frac{x_0}{2}) = 0$. Par une récurrence immédiate, $f(\frac{x_0}{2^n}) = 0$ pour tout $n \in \mathbb{N}$ puis en passant à la limite et en utilisant la continuité de f en 0, $f(0) = 0$: contradiction.

Ainsi, f ne s'annule pas ; d'après le théorème des valeurs intermédiaires, f est de signe constant. Supposons, quitte à remplacer f par $-f$, que f est strictement positive.

▹ Posons $g = \ln \circ f$. La fonction g vérifie l'équation précédente donc est de la forme $x \mapsto ax^2$ avec $a \in \mathbb{R}$. Par conséquent, f est de la forme $x \mapsto e^{ax^2}$. ■

2. Fonctions continues sur un intervalle

2.1. Théorème des valeurs intermédiaires

Théorème 7.30. Théorème des valeurs intermédiaires

Soit f une fonction continue sur l'intervalle I, a, $b \in I$ avec $a < b$. Pour tout y entre $f(a)$ et $f(b)$, il existe $x \in [a, b]$ tel que $y = f(x)$.

Remarque

Si on ajoute l'hypothèse de stricte monotonie pour f, l'antécédent x de y dans $[a, b]$ est alors unique.

Démonstration

Supposons, sans perte de généralité, $f(a) < f(b)$. Notons $A = \{x \in [a, b],\ f(x) \leq y\}$: cet ensemble est non vide ($a \in A$) et majoré par b. Soit $c = \sup A$. Comme f est continue, $f(c) \leq y$. De plus, c est la borne inférieure de l'intervalle $\{x \in [c, b],\ f(x) > y\}$ donc $f(c) \geq y$ (en passant à la limite encore avec la continuité de f). Par conséquent, $f(c) = y$. ■

Traduisons ce théorème en d'autres termes.

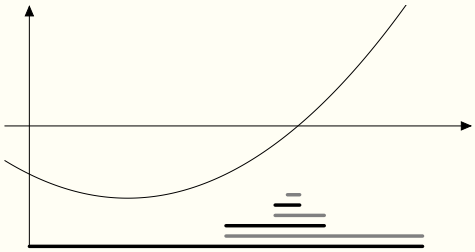
Corollaire 7.31.

L'image par une fonction continue d'un intervalle est un intervalle.

Remarque. Dichotomie

Comme une fonction continue ne peut changer de signe sans s'annuler, on peut en déduire un algorithme de recherche dichotomique d'un zéro d'une fonction continue changeant de signe. Partant d'un intervalle $[a,b]$ tel que $f(a)f(b)<0$, on introduit $c=\frac{a+b}{2}$ et on se ramène à l'un des intervalles $[a,c]$ ou $[c,b]$ où f change de signe. Renouvelant cette étape, on obtient une approximation de plus en plus précise d'un zéro de f.

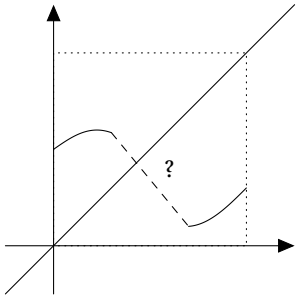
Voici pour illustrer la courbe représentative d'une fonction continue qui s'annule et les premiers intervalles de localisation du zéro obtenus par dichotomie (le plus bas est celui de départ et, à chaque montée, on ne garde qu'une moitié de l'intervalle précédent en fonction du signe de la fonction au milieu de celui-ci).



Exemple

Toute fonction $f:[0,1]\to[0,1]$ continue admet au moins un point fixe : si ni 0, ni 1 ne sont des points fixes, la fonction continue $x\mapsto f(x)-x$ prend des valeurs de signes distincts en 0 et 1 (car $f(0)>0$ et $f(1)<1$) donc s'annule d'après la propriété des valeurs intermédiaires.

Ce résultat est tout à fait naturel : dans le carré $[0,1]^2$, la courbe d'une fonction continue ne peut joindre le bord gauche au bord droit sans couper la diagonale.



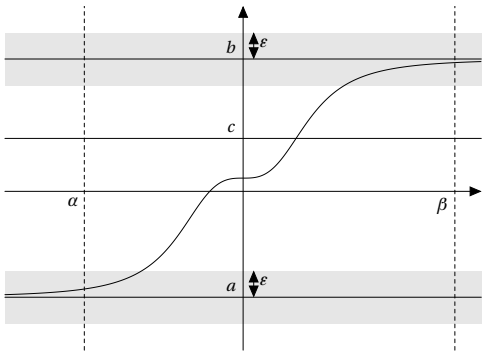
On remarque que ce raisonnement serait identique en remplaçant $[0,1]$ par n'importe quel autre segment.

Exemple Théorème des valeurs intermédiaires généralisé

Soit f une fonction continue sur $\mathbb{R}$ qui admet des limites finies a et $b>a$ en $-\infty$ et $+\infty$ respectivement. Montrons que toute valeur strictement entre a et b est atteinte.

Soit $c\in]a,b[$ et $\varepsilon<\min(c-a,b-c)$: ainsi, le réel c n'est ni dans la « bande » de demi-largeur ε autour de a ni dans la « bande » de demi-largeur ε autour de b.

On applique les définitions de limite en $\pm\infty$ avec ε. On a ainsi au moins deux réels α, β tels que $f(\beta)\geq b-\varepsilon$ et $f(\alpha)\leq a+\varepsilon$.



Par conséquent,

$$f(\beta) \geq b - \varepsilon \geq b - (b-c) = c = a + (c-a) \geq a + \varepsilon \geq f(\alpha).$$

On applique alors le théorème des valeurs intermédiaires sur l'intervalle entre α et β pour conclure.

Cet exemple se généralise évidemment à des limites infinies et permet d'obtenir, entre autres, l'exemple suivant.

Exemple

Un polynôme réel P de degré impair admet au moins une racine réelle. En effet, les limites en $\pm\infty$ sont celles de son monôme de plus haut degré :

- si le coefficient dominant est (strictement) positif, alors $\lim\limits_{+\infty} P = +\infty$ et $\lim\limits_{-\infty} P = -\infty$,
- sinon, $\lim\limits_{+\infty} P = -\infty$ et $\lim\limits_{-\infty} P = +\infty$.

En particulier, ces limites sont de signes distincts. Ainsi, ce polynôme change de signe et par conséquent s'annule d'après le théorème des valeurs intermédiaires (car la fonction $x \mapsto P(x)$ est continue).

2.2. Injectivité des fonctions continues

Commençons par énoncer un lemme.

Lemme 7.32.

Soit f une fonction continue sur un intervalle I. Si f est injective, alors f est strictement monotone.

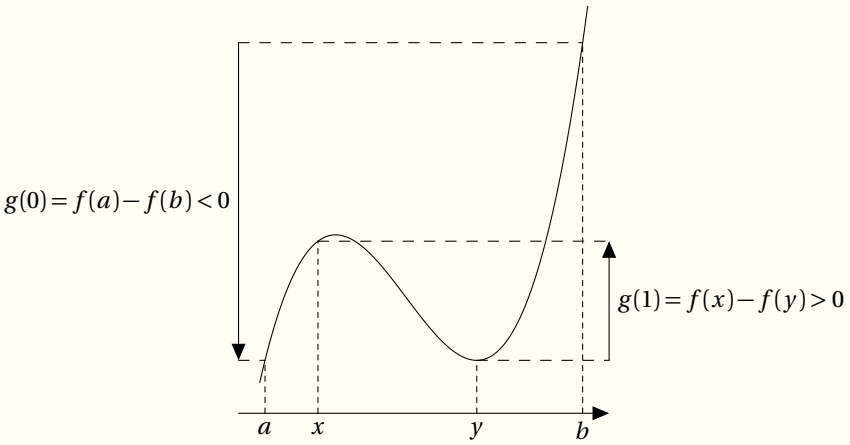
Remarquons que la réciproque est évidente : une fonction strictement monotone (continue ou pas) est injective.

Démonstration

Soit $a < b$ deux éléments de I. Comme f est injective, $f(a) \neq f(b)$. Supposons, sans perte de généralité, $f(a) < f(b)$. On cherche à montrer que f est croissante. Par l'absurde, s'il existe $a \leq x < y \leq b$ tel que $f(x) > f(y)$, alors on pose : $g : t \mapsto f(tx + (1-t)a) - f(ty + (1-t)b)$.
Remarquons que $g(0) = f(a) - f(b) < 0$ et $g(1) = f(x) - f(y) > 0$. Cette fonction est continue, change de signe sur $[0,1]$ donc s'annule en un certain $t \in [0,1]$: par conséquent,

$$f(tx + (1-t)a) = f(ty + (1-t)b) \quad \text{et} \quad tx + (1-t)a \leq x < y \leq ty + (1-t)b,$$

ceci contredit l'injectivité de f et clôt le raisonnement par l'absurde.



■

Remarque

Remarquons que le lemme précédent est faux si on enlève l'hypothèse de continuité : il existe des bijections de $\mathbb{R}$ dans $\mathbb{R}$ qui ne sont pas monotones.
Par exemple, la fonction

$$f : x \mapsto \begin{cases} x & \text{si } x \in \mathbb{Q}, \\ x+1 & \text{sinon} \end{cases}$$

est bijective (chaque rationnel est son unique antécédent et chaque irrationnel x admet pour unique antécédent $x-1$) mais pas monotone (par exemple, $f(3) < f(4) < f(\pi)$).

Proposition 7.33.

Soit f une fonction continue et strictement monotone sur un intervalle I. Alors, f réalise une bijection entre I et $f(I)$ et sa bijection réciproque est continue.

Démonstration

Il suffit d'étudier la continuité de f^{-1}. Comme f est strictement monotone, f^{-1} est aussi strictement monotone donc f^{-1} admet des limites à droite et à gauche en tout point de I. Soit $a \in I$. Par continuité de f,

$$f(\lim_{y \to a^-} f^{-1}(y)) = \lim_{y \to a^-} f(f^{-1}(y)) = \lim_{y \to a^-} y = a = f(f^{-1}(a)).$$

De même, $f(\lim_{y \to a^+} f^{-1}(y)) = f(f^{-1}(a))$, donc, avec l'injectivité de f, $\lim_{y \to a^-} f^{-1}(y) = \lim_{y \to a^+} f^{-1}(y) = f^{-1}(a)$ et f^{-1} est continue en a. ■

Remarque

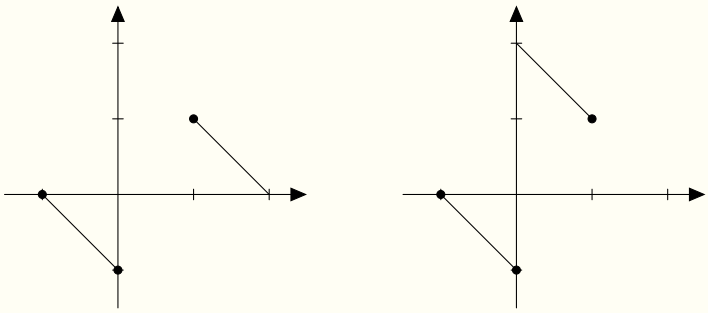
L'hypothèse « I intervalle » est essentielle comme on peut le constater avec la fonction

$$f : x \mapsto \begin{cases} -x-1 & \text{si } x \in [-1,0], \\ -x+2 & \text{si } x \in [1,2[\end{cases}$$

Cette fonction réalise une bijection continue de $[-1,0] \cup [1,2[$ dans $[-1,1]$ mais sa réciproque n'est pas continue en 0. En effet, après un calcul élémentaire,

$$f^{-1} : y \mapsto \begin{cases} -y-1 & \text{si } y \in [-1,0], \\ -y+2 & \text{si } y \in]0,1] \end{cases}$$

et $\lim\limits_{0^-} f^{-1} = f^{-1}(0) = -1$ et $\lim\limits_{0^+} f^{-1} = 2$. Voici leurs courbes représentatives :



3. Fonctions continues sur un segment

3.1. Image continue d'un segment

Voici un nouveau résultat qui ressemble « vaguement » à la propriété des valeurs intermédiaires mais qui est de nature différente : l'hypothèse forte ici est celle de segment, c'est-à-dire un intervalle de la forme $[a, b]$ avec a, $b \in \mathbb{R}$.

Proposition 7.34.

Toute fonction continue sur un segment y atteint ses bornes.

Avant de démontrer ce résultat, donnons une formulation alternative.

Corollaire 7.35.

L'image par une fonction continue d'un segment est un segment.

Démonstration

Soit f une fonction continue sur $[a, b]$.

▹ Si f n'est pas bornée, alors on peut construire une suite $(x_n)_n$ d'éléments de $[a, b]$ telle que $|f(x_n)| \geq n$ pour tout n. D'après le théorème de Bolzano-Weierstrass, on peut extraire une suite convergente de $(x_n)_n$ donc de $(|f(x_n)|)_n$ (car $|f|$ continue) : contradiction.

▹ Notons S la borne supérieure des valeurs de f sur $[a, b]$ (qui est finie d'après le point précédent). On construit une suite $(u_n)_n$ d'éléments de $[a, b]$ telle que

$$S - 2^{-n} \leq f(u_n) \leq S$$

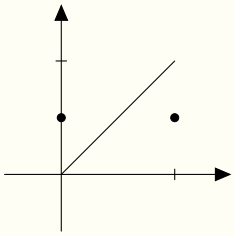
en utilisant que $S - 2^{-n}$ n'est pas un majorant de f.

On obtient, par le théorème de Bolzano-Weierstrass, une suite extraite de $(u_n)_n$ qui converge vers un réel $\ell \in [a, b]$. Par continuité de f en ℓ, $f(\ell) = S$.

On procède de même pour montrer que la borne inférieure I est atteinte. Ainsi, $f([a, b]) \subset [I, S]$ par définition des bornes supérieure et inférieure et $[I, S] \subset f([a, b])$ d'après la propriété des valeurs intermédiaires : on obtient l'égalité par double inclusion. ■

Remarque

▹ Le résultat devient faux pour une fonction seulement continue par morceaux sur un segment comme on peut le voir avec la fonction $f : [0,1] \to \mathbb{R}$ définie par $f(0) = f(1) = \frac{1}{2}$ et $f(x) = x$ sinon qui n'atteint pas ses bornes sur le segment $[0,1]$.

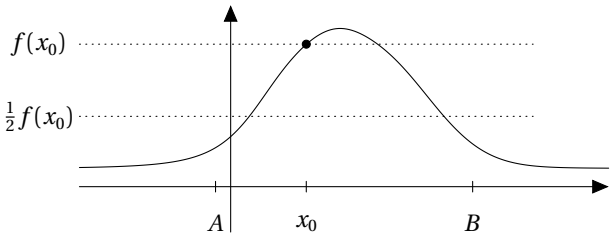


▹ En revanche, une fonction continue par morceaux sur un segment est toujours bornée. En effet, supposons que f soit continue par morceaux sur un segment $[a,b]$ et considérons une subdivision associée $(a_k)_{k\in[0,N]}$. Pour tout $k \in [\![1,N]\!]$, la restriction $f_{|]a_{k-1},a_k[}$ se prolonge en une fonction continue sur le segment $[a_{k-1},a_k]$ donc est bornée. Comme il y a un nombre fini de telles restrictions et de valeurs aux points de la subdivision, la fonction f est bornée.

Exemple

Si $f : \mathbb{R} \to \mathbb{R}$ est une fonction continue de limites nulles en $\pm\infty$, alors f atteint au moins l'une de ses bornes.
▹ Si f est la fonction nulle, alors le résultat est immédiat.
▹ Sinon, il existe $x_0 \in \mathbb{R}$ tel que $f(x_0) \neq 0$. Supposons sans perte de généralité $f(x_0) > 0$.
Comme les limites de f en $\pm\infty$ sont nulles, il existe $A < x_0 < B$ tels que

$$\forall x \notin [A,B], \qquad |f(x)| \leq \frac{1}{2} f(x_0).$$



Sur le segment $[A,B]$, f atteint sa borne supérieure. Or, cette quantité est supérieure à $f(x_0)$, car $x_0 \in [A,B]$, donc à $\frac{1}{2}f(x_0)$: c'est donc la borne supérieure de f sur $\mathbb{R}$.

Voici un exemple d'application où l'on utilise que le minimum est atteint sur un segment.

Exemple

Soit $f,g : [a,b] \to \mathbb{R}$ deux fonctions continues telles que $f(x) > g(x)$ pour tout $x \in [a,b]$. Il existe $m > 0$ tel que

$$\forall x \in [a,b], \qquad f(x) \geq g(x) + m.$$

En effet, la fonction continue $f - g$ est strictement positive et atteint sur le segment $[a,b]$ son minimum $m > 0$; pour tout $x \in [a,b]$, $f(x) - g(x) \geq m$.

Remarque
Une fois encore, l'hypothèse « I segment » est essentielle : le résultat peut être facilement mis en défaut que $[1,+\infty[$ avec la fonction $g : x \mapsto \frac{1}{x}$ et la fonction f nulle.

3.2. Continuité uniforme

Définition 7.36.

Une fonction $f : I \to \mathbb{R}$ est uniformément continue si

$$\forall \varepsilon > 0,\ \exists \eta > 0, \forall (x,x') \in I^2, \qquad |x-x'| \leq \eta \quad \Rightarrow \quad |f(x)-f(x')| \leq \varepsilon.$$

Proposition 7.37.

Soit $f : \mathbb{R}_+ \to \mathbb{R}$ uniformément continue. Alors, il existe $a,\ b \in \mathbb{R}_+$ tels que

$$\forall x \in \mathbb{R}_+, \qquad |f(x)| \leq ax+b.$$

Démonstration
Posons $\varepsilon = 1$ et considérons le réel $\eta > 0$ associé par la définition de continuité uniforme. Définissons, pour tout $x \in \mathbb{R}_+$, $n = \left\lfloor \frac{x}{\eta} \right\rfloor$. Alors,

$$\begin{aligned} |f(x)-f(0)| &\leq \sum_{k=0}^{n-1} |f(k\eta)-f((k+1)\eta)| + |f(x)-f(n\eta)| \\ &\leq (n+1)\varepsilon \leq \varepsilon + \frac{\varepsilon}{\eta}x. \end{aligned}$$

Par conséquent,

$$|f(x)| \leq |f(0)| + \varepsilon + \frac{\varepsilon}{\eta}x.$$

D'où le résultat avec $a = \frac{\varepsilon}{\eta}$ et $b = |f(0)| + \varepsilon$. ■

Proposition 7.38.

Une fonction lipschitzienne est uniformément continue.

Démonstration
Soit f une fonction k-lipschitzienne. Pour tout $\varepsilon > 0$, posons $\eta = \frac{\varepsilon}{k}$.
Alors, pour tous x, x' tels que $|x-x'| \leq \eta$,

$$|f(x)-f(x')| \leq k|x-x'| \leq k\eta = \varepsilon.$$

Donc f est uniformément continue. ■

On peut aussi obtenir le résultat suivant en intervertissant les quantificateurs (et donc en appauvrissant la propriété requise) :

Proposition 7.39.

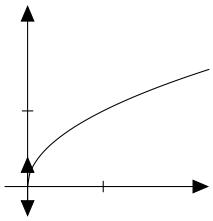
Une fonction uniformément continue est continue.

Exemple

▹ La fonction définie sur $\mathbb{R}$ par $x \mapsto x^2$ est continue mais pas uniformément continue (c'est, par exemple, une conséquence de la propriété 7.37).
▹ La fonction définie sur $[0,1]$ par $x \mapsto \sqrt{x}$ est uniformément continue (ce sera une conséquence du théorème suivant) mais n'est pas lipschitzienne. En effet, si cette fonction était lipschitzienne, alors, pour tous $x \neq y$, on aurait $|\sqrt{x}-\sqrt{y}| \leq k|x-y|$ et donc :

$$\frac{1}{\sqrt{x}+\sqrt{y}} \leq k,$$

ce qui entraîne une contradiction pour x, y suffisamment petits.
Graphiquement, la fonction $x \mapsto \sqrt{x}$ est de pente verticale au voisinage de 0 alors que les pentes d'une fonction k-lipschitzienne sont majorées par k.



Théorème 7.40. Heine

Une fonction continue sur un segment est uniformément continue.

Démonstration
Raisonnons par l'absurde, si f n'est pas uniformément continue, alors il existe $\varepsilon > 0$ tel que l'on peut construire deux suites $(u_n)_n$ et $(v_n)_n$ d'éléments de I vérifiant, pour tout $n > 0$:

$$|u_n - v_n| \leq \frac{1}{n}, \qquad |f(u_n) - f(v_n)| \geq \varepsilon.$$

D'après le théorème de Bolzano-Weierstrass, on peut obtenir une suite extraite $(u_{\varphi(n)})_n$ qui converge. Par conséquent, la suite $(v_{\varphi(n)})_n$ converge vers la même limite. Comme f est continue, les deux suites $(f(u_{\varphi(n)}))_n$ et $(f(v_{\varphi(n)}))_n$ convergent vers la même limite, ce qui est contradictoire (puisqu'elles sont séparées d'au moins ε). ■

3.3. Approximation uniforme par les fonctions en escalier

Définition 7.41.

Une fonction f est en escalier sur un segment $[a,b]$ s'il existe une subdivision $(a_k)_{k\in[0,n]}$ de $[a,b]$ telle que f soit constante sur les intervalles $]a_k, a_{k+1}[$ pour $k \in [\![0, n-1]\!]$.

Proposition 7.42.

Soit f continue sur un segment $[a,b]$ et $\varepsilon > 0$. Il existe des fonctions en escaliers φ et ψ telles que, pour tout $x \in [a,b]$,

$$\varphi(x) \leq f(x) \leq \psi(x), \qquad \psi(x) - \varphi(x) \leq \varepsilon.$$

Démonstration

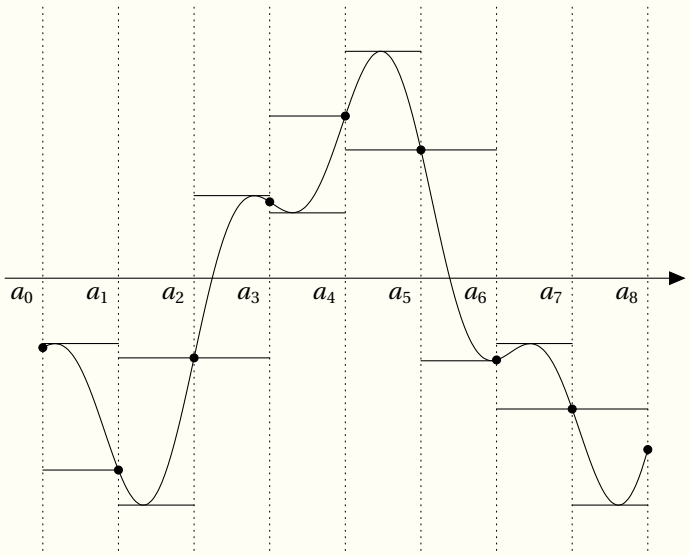
Soit $\varepsilon > 0$. D'après le théorème de Heine, la fonction continue f sur $[a,b]$ est uniformément continue : il existe $\eta > 0$ tel que

$$\forall (x,x') \in I^2, \qquad |x - x'| \leq \eta \quad \Rightarrow \quad |f(x) - f(x')| \leq \varepsilon.$$

Considérons $n \in \mathbb{N} \setminus \{0\}$ tel que $\frac{b-a}{n} \leq \eta$ et définissons la subdivision régulière $(a_k)_{k \in [0,n]}$ de $[a,b]$ par

$$\forall k \in [\![0,n]\!], \qquad a_k = a + k\,\frac{b-a}{n}.$$

Définissons alors les fonctions en escaliers φ et ψ sur chaque intervalle $]a_k, a_{k+1}[$ comme les fonctions constantes respectivement égales au minimum et au maximum de f sur l'intervalle fermé correspondant $[a_k, a_{k+1}]$.



Sur chaque intervalle $]a_k, a_{k+1}[$, les valeurs de φ et ψ sont des valeurs atteintes par f sur $[a_k, a_{k+1}]$; l'intervalle étant de taille inférieure à η, l'écart entre les deux valeurs, donc entre φ et ψ, est majoré par ε d'après la propriété d'uniforme continuité.

On complète la définition de ces fonctions en escalier sans perdre cette propriété sur l'écart en posant par exemple $\varphi(a_k) = \psi(a_k) = f(a_k)$ pour tout $k \in [\![0,n]\!]$. ■

Remarque

Ce résultat pourra se reformuler avec le cours de seconde année de la façon suivante : toute fonction continue f sur un segment $[a,b]$ est limite uniforme d'une suite de fonctions $(\varphi_n)_n$ en escalier.

En effet, en prenant pour chaque n, $\varepsilon = 2^{-n}$ dans la proposition précédente, on obtient une fonction en escalier φ_n telle que

$$\forall x \in [a,b], \qquad |f(x) - \varphi_n(x)| \leq 2^{-n}.$$

On utilisera ce résultat d'approximation ou plutôt le corollaire suivant afin de construire l'intégrale d'une fonction continue par morceaux sur un segment.

Corollaire 7.43.

Soit f continue par morceaux sur un segment $[a,b]$ et $\varepsilon > 0$. Il existe des fonctions en escaliers φ et ψ telles que, pour tout $x \in [a,b]$,

$$\varphi(x) \leq f(x) \leq \psi(x), \qquad \psi(x) - \varphi(x) \leq \varepsilon.$$

Démonstration

Il suffit d'appliquer le résultat précédent à chaque restriction de f aux intervalles ouvert de la subdivision convenablement prolongée aux extrémités de celui-ci. ■

FICHE

SYNTHÈSE

La première partie de ce chapitre est consacrée à la définition et aux propriétés de limite et de continuité.

Limites

Soit $f : I \to \mathbb{R}$ et a une borne ou un point de l'intervalle I.

- f admet la limite $\ell \in \mathbb{R}$ en a si
$$\forall \varepsilon > 0,\ \exists \eta > 0,\ \forall x \in]a-\eta, a+\eta[\cap I, \qquad |f(x)-\ell| \leq \varepsilon.$$
On note alors $f(x) \underset{x\to a}{\to} \ell$.
- f admet la limite $+\infty$ en a si
$$\forall M > 0,\ \exists \eta > 0,\ \forall x \in]a-\eta, a+\eta[\cap I, \qquad f(x) \geq M.$$
On note alors $f(x) \underset{x\to a}{\to} +\infty$.
- f admet la limite $-\infty$ en a si
$$\forall m < 0,\ \exists \eta > 0,\ \forall x \in]a-\eta, a+\eta[\cap I, \qquad f(x) \leq m.$$
On note alors $f(x) \underset{x\to a}{\to} -\infty$.

Les définitions de limites en $\pm\infty$ se déduisent de celles-ci en remplaçant les voisinages de a par des voisinages de $\pm\infty$.
Celles des limites à gauche et à droite sont les analogues en se limitant à $I\cap]-\infty, a[$ et $I\cap]a, +\infty[$.

Limite monotone

Soit $f :]a, b[\to \mathbb{R}$ croissante. Alors, f admet des limites finies à gauche et à droite en tout point $c \in]a, b[$. De plus,
$$\lim_{c^-} f \leq f(c) \leq \lim_{c^+} f.$$

La proximité de ces définitions avec les définitions pour les suites permet de démontrer de la même manière les règles de manipulation, l'utilisation des inégalités ou de la monotonie.

Continuité

Une fonction $f : I \to \mathbb{R}$ est continue en $a \in I$ si elle admet une limite finie en a. Elle est continue sur I si elle est continue en tout point de I (propriété locale).

Comme la continuité se définit en termes de limites, les propriétés opératoires sont aussi valables pour la continuité.

Il est bon de savoir utiliser la caractérisation séquentielle des limites pour montrer qu'une fonction n'admet pas de limite en un point (qu'elle n'est pas continue).

Voici le bilan des propriétés essentielles des fonctions continues sur un intervalle.

Fonctions continues sur un intervalle

▹ Soit f une fonction continue sur l'intervalle I, a, $b \in I$ avec $a < b$. Pour tout y entre $f(a)$ et $f(b)$, il existe $x \in [a, b]$ tel que $y = f(x)$. En d'autres termes, l'image par une fonction continue d'un intervalle est un intervalle.
▹ Soit f une fonction continue sur un intervalle I et injective. Alors, f est strictement monotone.
▹ Soit f une fonction continue et strictement monotone sur un intervalle I. Alors, f réalise une bijection entre I et $f(I)$ et sa bijection réciproque est continue.

Voici le bilan des propriétés essentielles des fonctions continues sur un segment.

Fonctions continues sur un segment

▹ Une fonction continue sur un segment est bornée et atteint ses bornes. En d'autres termes, l'image par une fonction continue d'un segment est un segment.
▹ Une fonction f continue sur un segment I est uniformément continue, c'est-à-dire

$$\forall \varepsilon > 0,\ \exists \eta > 0, \forall (x, x') \in I^2, \qquad |x - x'| \leq \eta \quad \Rightarrow \quad |f(x) - f(x')| \leq \varepsilon.$$

L'approximation uniforme (une fonction continue sur un segment peut être approchée arbitrairement près par des fonctions en escalier) est une conséquence de ces propriétés qui sera utilisée plus tard dans le cours d'intégration.

EXERCICES

Vrai ou faux ?

		Vrai	Faux
a)	f admet la limite $+\infty$ en $-\infty$ si $$\forall M>0,\ \exists A<0,\ \forall x\geq A,\qquad f(x)\geq M.$$	☐	☐
b)	Une fonction monotone admet une limite en tout point intérieur à son domaine de définition.	☐	☐
c)	Si $\lim\limits_{x\to a}(f(x)-g(x))=0$, alors $\lim\limits_{x\to a}f(x)=\lim\limits_{x\to a}g(x)$.	☐	☐
d)	La composée de deux fonctions continues est continue.	☐	☐
e)	Une fonction continue est bornée.	☐	☐
f)	Une fonction périodique est bornée.	☐	☐
g)	Une fonction bornée atteint ses bornes.	☐	☐
h)	Une fonction continue bornée atteint ses bornes.	☐	☐
i)	Soit $f,g:[a,b]\to\mathbb{R}$ continues telles que $f(x)>g(x)$ pour tout $x\in[a,b]$. Alors, $$\sup_{x\in[a,b]}f(x)>\sup_{x\in[a,b]}g(x).$$	☐	☐
j)	Le polynôme $X^4-X^3-19X+1$ admet au moins une racine dans $[3,+\infty[$.	☐	☐
k)	Une fonction continue à valeurs dans $\mathbb{Q}$ est constante.	☐	☐
l)	Une fonction f continue vérifiant $f(x)^2=1$ pour tout $x\in\mathbb{R}$ est constante égale à 1 ou constante égale à -1.	☐	☐
m)	L'image réciproque d'un intervalle par une fonction continue est un intervalle.	☐	☐
n)	La fonction $x\mapsto\dfrac{x}{\lvert x\rvert}$ est prolongeable par continuité en 0.	☐	☐
o)	La fonction partie fractionnaire $x\mapsto x-\lfloor x\rfloor$ est continue à gauche en tout point.	☐	☐
p)	Une fonction strictement monotone réalise une bijection entre un intervalle I et $f(I)$.	☐	☐
q)	Une fonction injective est strictement monotone.	☐	☐
r)	Une fonction strictement monotone est injective.	☐	☐

s)	Une fonction lipschitzienne sur $\mathbb{R}$ est uniformément continue.	☐	☐
t)	Une fonction bornée sur $\mathbb{R}$ est uniformément continue.	☐	☐
u)	La somme de deux fonctions uniformément continues est uniformément continue.	☐	☐
v)	La fonction $x \mapsto \sqrt{x}$ est lipschitzienne sur $[0,1]$.	☐	☐
w)	La fonction $x \mapsto \sqrt{x}$ est lipschitzienne sur $[1,+\infty[$.	☐	☐

Exercices d'application du chapitre 7

●●○ **Exercice 1**

La fonction $x \mapsto \cos(x+\frac{1}{x}) - \cos\frac{1}{x}$ est-elle prolongeable par continuité en 0 ?

●●○ **Exercice 2**

Soit $f : x \mapsto 2\lfloor x \rfloor - \cos(3\pi(x - \lfloor x \rfloor))$.

1. Déterminer les points où f est continue.
2. Calculer $\operatorname{card} f^{-1}(\{y\})$ pour $y \in \mathbb{R}$.

●●● **Exercice 3**

Montrer qu'il n'existe pas de fonction $f : \mathbb{R} \to [0,1]$ continue, bijective de réciproque continue.

●●○ **Exercice 4**

Montrer que deux fonctions croissantes sur un intervalle ouvert dont la somme est continue sont continues.

●○○ **Exercice 5**

Étudier la continuité sur $\mathbb{R}_+$ de $f : x \mapsto \sup\limits_{n\in\mathbb{N}} \frac{x^n}{n!}$.

●●○ **Exercice 6**

Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction continue. Posons $g : x \mapsto \lfloor x \rfloor + f(x - \lfloor x \rfloor)$. Déterminer à quelle condition sur f la fonction g est continue.

●●● **Exercice 7**

Soit $f : \mathbb{R}_+^* \to \mathbb{R}$ continue, telle que, pour tout $x > 0$, la suite $(f(nx))_n$ est croissante. Montrer que f est croissante.

●●● **Exercice 8**

Soit $f : \mathbb{R}_+^* \to \mathbb{R}$ nulle en tous les irrationnels telle que l'image du rationnel d'écriture irréductible $\frac{p}{q}$ est $\frac{1}{p+q}$.

1. Soit $\varepsilon > 0$. Montrer que l'ensemble $\{x \in \mathbb{R}_+^*,\ f(x) > \varepsilon\}$ est fini.
2. Déterminer l'ensemble des points où f est continue.

●●● **Exercice 9**

Déterminer les fonctions $f:\mathbb{R}^*\to\mathbb{R}$ continues telles que $f(x^2)=\frac{1}{x^2}f(x)$ pour $x\neq 0$.

●○○ **Exercice 10**

Soit $f:[0,1]\to\mathbb{R}$ une fonction continue et $x_1,\dots,x_n\in]0,1[$. Montrer qu'il existe $c\in]0,1[$ tel que

$$f(c)=\frac{f(x_1)+\ldots+f(x_n)}{n}.$$

●○○ **Exercice 11**

Soit n réels $a_1<a_2<\cdots<a_n$ et la fonction

$$f:x\mapsto\frac{1}{x-a_1}+\frac{1}{x-a_2}+\ldots+\frac{1}{x-a_n}.$$

Montrer que f s'annule exactement $n-1$ fois sur son ensemble de définition.

●●○ **Exercice 12 Théorème des cordes universelles**

Soit $f:[0,1]\to\mathbb{R}$ une application continue telle que $f(0)=f(1)$ et $n\in\mathbb{N}\setminus\{0\}$. Définissons

$$g:\left\{\begin{array}{ccc}[0,1-\frac{1}{n}] & \to & \mathbb{R}\\ x & \mapsto & f(x+\frac{1}{n})-f(x)\end{array}\right.$$

1. Calculer $\sum_{k=0}^{n-1}g\left(\frac{k}{n}\right)$.
2. En déduire qu'il existe $\alpha_n\in[0,1-\frac{1}{n}]$ tel que $f(\alpha_n+\frac{1}{n})=f(\alpha_n)$.

●●○ **Exercice 13**

Soit $(a_1,\dots,a_n)\in[0,1]^n$. Montrer qu'il existe $a\in[0,1]$ tel que $\frac{1}{n}\sum_{k=1}^{n}|a-a_k|=\frac{1}{2}$.

●○○ **Exercice 14**

Soit $f:\mathbb{R}_+\to\mathbb{R}_+$ continue telle que $\lim\limits_{x\to\infty}\frac{f(x)}{x}<1$. Montrer que f admet un point fixe.

●●● **Exercice 15**

Soit $f:\mathbb{R}\to\mathbb{R}$ continue injective telle qu'il existe $n\in\mathbb{N}\setminus\{0\}$ vérifiant

$$\forall x\in\mathbb{R},\qquad f^n(x)=x.$$

Montrer que f ou f^2 est l'identité.

●○○ **Exercice 16**

Soit f et g deux fonctions continues sur $[0,1]$ telles que

$$\forall x\in[0,1],\qquad 0<f(x)<g(x).$$

Montrer qu'il existe un réel $C>1$ tel que

$$\forall x\in[0,1],\ Cf(x)\leq g(x).$$

●●○ **Exercice 17**

Soit $f:\mathbb{R}_+\to\mathbb{R}$ continue et surjective. Montrer que $f^{-1}(\{0\})$ est infini.

EXERCICES

●●● **Exercice 18**

Soit $f:[a,b]\to\mathbb{R}$ continue et $x_0\in[a,b]$ tels que

$$f(x_0)<\sup_{[a,x_0]} f.$$

Montrer qu'il existe $u<x_0$ tel que $\sup_{[a,u]} f=\sup_{[a,x_0]} f$.

●●○ **Exercice 19**

Soit $f:\mathbb{R}\to\mathbb{R}$ T-périodique et continue. Montrer qu'il existe $a\in\mathbb{R}$ tel que $f(\mathbb{R})=f([a,a+\frac{T}{2}])$.

●●○ **Exercice 20**

Soit $f:[a,b[\to\mathbb{R}$ uniformément continue. Montrer que f est bornée.

●●● **Exercice 21**

Soit $f:\mathbb{R}_+\to\mathbb{R}$ une fonction continue de limite nulle en $+\infty$. Montrer que f est uniformément continue.

Exercices d'approfondissement du chapitre 7

●●○ **Exercice A**

Soit $f:\mathbb{R}\to\mathbb{R}$ une fonction continue telle que $\cos\circ f(x)\underset{x\to+\infty}{\longrightarrow}\ell$.

Montrer que f converge en $+\infty$.

●○○ **Exercice B**

Soit $f:]a,b[\to\mathbb{R}$ une fonction croissante. Montrer que la fonction $x\mapsto\lim\limits_{x^+} f$ est croissante.

●●○ **Exercice C**

1. Déterminer les fonctions $f:\mathbb{R}\to\mathbb{R}$ continues telles que

$$\forall x\in\mathbb{R},\qquad \sup_{t\leq x} f(t)=x.$$

2. Exhiber une fonction non continue vérifiant cette propriété.

●●● **Exercice D**

Soit $f:[a,b]\to\mathbb{R}$ telle que

- $f(a)>f(b)$;
- il existe $g:[a,b]\to\mathbb{R}$ continue telle que $f+g$ est croissante.

Montrer que, pour tout $z\in[f(b),f(a)]$, il existe $c\in[a,b]$ tel que $z=f(c)$.

●●● **Exercice E**

Montrer qu'il n'existe pas de fonction $f:\mathbb{R}\to\mathbb{R}$ continue telle que pour tout $c\in\mathbb{R}$,

$$\operatorname{card} f^{-1}(\{c\})\in\{0,2\}.$$

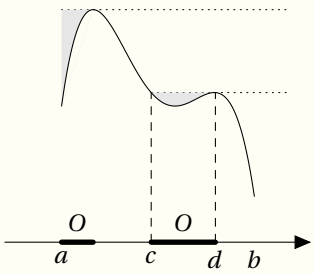
●●○ **Exercice F**

Déterminer les fonctions $f:\mathbb{R}\to\mathbb{R}$ continues telles que $f(\mathbb{Q})\subset\mathbb{R}\setminus\mathbb{Q}$ et $f(\mathbb{R}\setminus\mathbb{Q})\subset\mathbb{Q}$.

●●○ **Exercice G Lemme du soleil levant**

Soit $f : [a,b] \to \mathbb{R}$ continue et $O = \{x \in [a,b],\ \exists y > x,\ f(y) > f(x)\}$.

Intuitivement, O est l'ensemble des abscisses des points de la courbe représentative de f à l'ombre si on éclaire depuis la droite.



Supposons qu'il existe $c < d$ tels que $c \notin O$, $d \notin O$, mais $]c,d[\subset O$.

1. Montrer que, pour tout $t \in]c,d[$, $f(t) \leq f(d)$.
2. En déduire que $f(c) = f(d)$.

●○○ **Exercice H**

1. Montrer qu'il n'existe pas de fonction $f : [0,1] \to]0,1[$ continue et surjective.
2. Trouver une fonction $f :]0,1[\to [0,1]$ continue et surjective.

●●○ **Exercice I**

Soit $f : [a,b] \to \mathbb{R}$ continue. Montrer que $\sup\limits_{[a,b]} f = \sup\limits_{]a,b[} f$.

Problèmes du chapitre 7

●○○ **Problème 1 Fonctions semi-continues inférieurement**

Une fonction f définie sur un intervalle $I \subset \mathbb{R}$ est SCI (pour semi-continue inférieurement) si

$$\forall x_0 \in I,\ \forall \varepsilon > 0, \exists \eta > 0, \forall x \in I, \qquad |x - x_0| \leq \eta \quad \Rightarrow \quad f(x) \geq f(x_0) - \varepsilon.$$

1. a. Montrer qu'une fonction continue sur un intervalle est SCI.
   b. Déterminer une fonction SCI qui n'est pas continue.
2. a. Montrer que la somme de deux fonctions SCI sur un intervalle I est encore une fonction SCI.
   b. Montrer que le produit de deux fonctions SCI n'est pas forcément une fonction SCI.
3. Soit f une fonction SCI sur un segment I.
   a. Montrer que f est minorée.
   b. Montrer qu'il existe $x_0 \in I$ tel que $f(x_0) = \inf\limits_{x \in I} f(x)$.

●○○ **Problème 2 Module de continuité uniforme**

Le module de continuité d'une fonction continue $f : [0,1] \to \mathbb{R}$ est la fonction $\omega_f : [0,1] \to \mathbb{R}$ définie par :

$$\omega_f : t \mapsto \sup\{|f(x) - f(y)|,\ (x,y) \in [0,1]^2,\ |x - y| \leq t\}.$$

1. Calculer ω_g pour la fonction $g : x \mapsto \sqrt{x}$.
   Indication : on pourra vérifier que $\sqrt{x} - \sqrt{y} \leq \sqrt{x - y}$ pour tous $x > y \geq 0$.

PROBLÈMES

2. Soit f une fonction continue.
   a. Montrer que ω_f est une fonction croissante et calculer $\omega_f(0)$.
   b. Montrer que la fonction ω_f est sous-additive, c'est-à-dire que, pour tout $(u,v)\in[0,1]^2$ tel que $u+v\leq 1$, $\omega_f(u+v)\leq\omega_f(u)+\omega_f(v)$.
   c. Montrer que la fonction ω_f est continue.
   Indication : on pourra justifier puis utiliser l'uniforme continuité de f.
   d. Montrer que si ω_f s'annule hors de 0, alors ω_f est identiquement nulle.
   Préciser les fonctions f pour lesquelles $\omega_f=0$.
3. Soit $\omega:[0,1]\to\mathbb{R}$ une fonction continue, croissante, nulle à l'origine et sous-additive. Montrer qu'il existe une fonction dont ω est le module de continuité.

●○○ Problème 3 Transformée de Legendre

Pour toute fonction $\varphi : I\to\mathbb{R}$, on définit D_φ^* l'ensemble des $y\in\mathbb{R}$ telle que la fonction $x\mapsto xy-\varphi(x)$ est majorée puis

$$\varphi^*:\left\{\begin{array}{ccc} D_\varphi^* & \to & \mathbb{R} \\ y & \mapsto & \sup\{xy-\varphi(x),\ x\in I\}\end{array}\right.$$

1. Montrer que D_φ^* est un intervalle.
   Indication : on montrera que si $a,b\in D_\varphi^*$, alors $[a,b]\subset D_\varphi^*$.
2. Montrer que si φ est continue sur un segment I, alors $D_\varphi^*=\mathbb{R}$.
3. Déterminer D_φ^* et φ^* dans les cas suivants :
   a. $I=\mathbb{R}$ et $\varphi(x)=k$, $k\in\mathbb{R}$
   b. $I=\mathbb{R}_+$ et $\varphi(x)=\frac{x^\alpha}{\alpha}$, $\alpha>1$
   c. $I=\mathbb{R}$ et $\varphi(x)=e^x$
4. Montrer que, pour tout $k\in\mathbb{R}$ et toute fonction φ, $D_{\varphi+k}^*=D_\varphi^*$ et $(\varphi+k)^*=\varphi^*-k$.
5. Montrer que, si $I\subset\mathbb{R}_+$, alors φ^* est croissante. Que dire si $I\subset\mathbb{R}_-$?
6. Montrer que φ^* est une fonction convexe c'est-à-dire

$$\forall(x,y)\in(D_\varphi^*)^2,\quad \forall\lambda\in[0,1],\qquad \varphi^*(\lambda x+(1-\lambda)y)\leq\lambda\varphi^*(x)+(1-\lambda)\varphi^*(y)$$

7. Soit φ une fonction convexe, continue, positive.
   a. Montrer que, pour tout $x\in I$, $y\mapsto xy-\varphi^*(y)$ est majorée par $\varphi(x)$.
   b. Montrer que $(\varphi^*)^*=\varphi$.

●●● Problème 4 Théorème de Sarkovskii

Soit $f:\mathbb{R}\to\mathbb{R}$ une fonction continue. Un réel a est un point de période $n\in\mathbb{N}^*$ pour f si $f^n(a)=a$ et $f^k(a)\neq a$ pour tout $k\in[\![1,n-1]\!]$, c'est-à-dire si l'ensemble $\{f^k(a),\ k\in\mathbb{N}\}$ est de cardinal n.

Partie A – Préliminaires

1. Soit g la fonction définie par $g(x)=-\frac{3}{2}x^2+\frac{5}{2}x+1$ pour tout $x\in\mathbb{R}$.
   a. Montrer que 0 est un point de période 3 pour g.
   b. Déterminer les points fixes (éventuels) de g.
2. Soit ϕ une fonction continue sur le segment $[\alpha,\beta]$ telle que $[\alpha,\beta]\subset\phi([\alpha,\beta])$. Montrer que ϕ admet (au moins) un point fixe.
3. Supposons que x_0 est un point de période 3 pour f. Notons

$$a=\min\{x_0,f(x_0),f^2(x_0)\}.$$

   a. Justifier que

$$\{a,f(a),f^2(a)\}=\{x_0,f(x_0),f^2(x_0)\}.$$

b. Supposons que $a < f(a) < f^2(a)$. Montrer que

$$[f(a), f^2(a)] \subset [a, f^2(a)] \subset f([f(a), f^2(a)]).$$

En déduire que f admet un point fixe.

c. Supposons que $a < f^2(a) < f(a)$. Montrer que f admet un point fixe.

On suppose dorénavant que f admet un point a de période 3 tel que $a < f(a) < f^2(a)$, puis on note $J_0 = [a, f(a)]$ et $J_1 = [f(a), f^2(a)]$. Il a déjà été montré que f admet un point fixe ; l'objectif est de montrer que f admet un point de période n pour tout $n > 1$.

Partie B – Construction

On cherche dans cette partie à établir l'existence d'une suite décroissante (pour l'inclusion) de segments $(I_k)_{k \leq n}$ vérifiant les quatre conditions suivantes :

(C1) $I_0 = J_1$,

(C2) $\forall k \in [\![1, n-2]\!], \quad f(I_k) = I_{k-1}$,

(C3) $f^{n-1}(I_{n-1}) = J_0$,

(C4) $f^n(I_n) = J_1$.

1. Soit I un segment, $\phi : I \to I$ continue et $K = [\alpha, \beta]$ un segment non vide et non réduit à un point inclus dans $\phi(I)$.
   a. Montrer l'existence de $a, b \in I$ tels que $\phi(a) = \alpha$ et $\phi(b) = \beta$. On suppose, par symétrie, que $a < b$.
   b. Soit $A = \{x \in [a, b],\ \phi(x) = \beta\}$. Justifier l'existence de $v = \min A$.
   On considère de même $u = \max B$ où $B = \{x \in [a, v],\ \phi(x) = \alpha\}$.
   c. Montrer l'existence d'un segment $L \subset I$ tel que $K = \phi(L)$.
2. a. Construire une suite $(I_k)_{k \leq n-2}$ satisfaisant **(C1)** et **(C2)**.
   b. Montrer que cette suite vérifie que $\forall k \in [\![1, n-2]\!], \quad f^k(I_k) = J_1$.
3. Montrer que $J_0 \subset f(J_1) = f^{n-1}(I_{n-2})$; en déduire l'existence d'un segment I_{n-1} satisfaisant la condition **(C3)**.
4. Montrer que $J_1 \subset f^n(I_{n-1})$; en déduire l'existence de I_n satisfaisant la condition **(C4)**.

Partie C – Théorème de Sarkovskii

On considère dorénavant une suite décroissante de segments $(I_k)_{k \leq n}$ vérifiant les quatre conditions **(C1)**, **(C2)**, **(C3)** et **(C4)**.

1. Montrer que f^n admet au moins un point fixe dans I_n. On notera x_0 l'un de ces points fixes.
2. Montrer que $x_0 \notin \{f(a), f^2(a)\}$.
3. Conclure que x_0 est de période n.

Problème 5 Théorème de Coppel

Définition 7.44.

▹ Un réel $c \in [a, b]$ est un point fixe de l'application $f : [a, b] \to [a, b]$ si $f(c) = c$.
▹ Un réel $c \in [a, b]$ est un 2-cycle de l'application $f : [a, b] \to [a, b]$ si $f(c) \neq c$ et $f \circ f(c) = c$.

Dans ce problème, on note f^n la composée $n^{\text{ième}}$ de l'application $f : [a, b] \to I$; par exemple, $f^0 = \mathrm{Id}$, $f^1 = f$, $f^2 = f \circ f$, $f^3 = f \circ f \circ f \ldots$
L'objectif est d'étudier la convergence de la suite $(f^n(c))_n$ pour une application $f : [a, b] \to [a, b]$ continue sans 2-cycle et $c \in [a, b]$.

PROBLÈMES

Partie A – Exemples

1. Montrer que tout point fixe d'une application $f:[a,b]\to[a,b]$ est un point fixe de f^n pour $n>0$.
2. Donner un exemple d'application continue $f:[a,b]\to[a,b]$ qui admet un 2-cycle.
3. Donner un exemple d'application continue $f:[a,b]\to[a,b]$ qui n'admet pas de 2-cycle.
4. Déterminer, selon la valeur de $\lambda\in]0,4]$, les points fixes et les 2-cycles de l'application

$$f:\left\{\begin{array}{ccc}[0,1] & \to & [0,1]\\ x & \mapsto & \lambda x(1-x)\end{array}\right.$$

Partie B – Résultats d'existence

1. Montrer que toute application continue $f:[a,b]\to[a,b]$ admet un point fixe.
2. Soit un segment $[c,d]\subset[a,b]$ et une application continue $f:[a,b]\to[a,b]$ telle que
$$[c,d]\subset f([c,d]).$$
Montrer que f admet un point fixe dans $[c,d]$.
Indication : on pourra introduire des antécédents de c et d par f dans $[c,d]$.
3. Soit un segment $[c,d]\subset[a,b]$ et une application continue $f:[a,b]\to[a,b]$ telle que
$$[c,d]\subset f([a,b]).$$
Montrer qu'il existe un intervalle $[c',d']\subset[a,b]$ tel que $[c,d]=f([c',d'])$.
4. Soit $c\in[a,b]$ et une application continue $f:[a,b]\to[a,b]$ telle que
$$[c,b]\subset f([a,c])\qquad\text{et}\qquad[a,c]\subset f([c,b]).$$
Montrer que c est un point fixe ou que f admet un 2-cycle.
Indication : on pourra montrer que f^2 admet un point fixe dans un segment $J\subset[a,c]$ vérifiant $f(J)=[c,b]$.

Partie C – Lemme de Coppel (I)

Dans cette partie, on suppose que $f:[a,b]\to[a,b]$ est continue, n'admet pas de 2-cycle et on considère $c\in[a,b]$ tel que $f(c)>c$.

1. Justifier que $f^2(c)\neq c$.
2. On suppose que $c>f^2(c)$.
   a. Montrer que f^2 admet un point fixe sur l'intervalle $[a,c]$.
   Posons d la borne supérieure de l'ensemble des points fixes de f^2 sur $[a,c]$.
   b. Justifier que $d<c$.
   c. Montrer que f n'admet pas de point fixe sur l'intervalle $]d,c[$.
   d. Soit $x\in]d,c[$. Comparer les réels x, $f(x)$ et $f^2(x)$.
   e. Montrer qu'il existe $e\in]d,c[$ tel que $f(e)\in]d,c[$.
   Indication : on rappelle que $f(d)<c$.
   f. Comparer les résultats des deux dernières questions et conclure pour l'hypothèse $c>f^2(c)$.

Partie D – Lemme de Coppel (II)

Dans cette partie, on suppose encore que $f:[a,b]\to[a,b]$ est continue, n'admet pas de 2-cycle et on considère $c\in[a,b]$ tel que $f(c)>c$. On cherche à montrer que, pour tout $n\geq 3$, $f^n(c)>c$.
Considérons, pour tout $n\in\mathbb{N}$, $u_n=f^n(c)$ et considérons, en supposant qu'il existe,

$$N=\min\{j\geq 3,\ u_j\leq c\}.$$

Soit k le plus petit entier tel que la suite finie $(u_n)_{n\in[k+1,N]}$ est strictement décroissante et $j\in[\![1,N-k]\!]$ tel que $u_{k+j}\geq u_k>u_{k+j+1}$.

1. Justifier que $u_N < u_k \leq u_{k+1}$.
2. Montrer que $[u_k, u_{k+j}] \subset f([u_{k+j}, u_{k+1}])$ et $[u_{k+j}, u_{k+1}] \subset f([u_k, u_{k+j}])$.
3. En déduire que f admet un 2-cycle.

Partie E – Théorème de convergence

Dans cette partie, on suppose encore que $f : [a, b] \to [a, b]$ est continue, n'admet pas de 2-cycle, on considère $c \in [a, b]$ et, pour tout $n \in \mathbb{N}$, $u_n = f^n(c)$.

1. Montrer que s'il existe un rang N tel que $u_N = u_{N+1}$, alors la suite $(u_n)_n$ converge.
2. Montrer que s'il existe un rang N tel que, pour tout $n \geq N$, $u_n < u_{n+1}$, alors la suite $(u_n)_n$ converge.
   On suppose dorénavant que la suite $(u_n)_n$ n'est pas monotone à partir d'un certain rang.
3. Considérons $(y_n)_n$ (respectivement $(z_n)_n$) la suite extraite de $(u_n)_n$ composée des valeurs u_k strictement inférieures (respectivement supérieures) au terme suivant u_{k+1}.
   a. Étudier la monotonie des suites $(y_n)_n$ et $(z_n)_n$ puis justifier leur convergence.
   b. Montrer que $\lim y_n = \lim z_n$.
   c. En déduire que la suite $(u_n)_n$ converge.
4. Conclure.

PROBLÈMES

CORRIGÉS

Corrigés des Vrai/Faux

a) Faux, il faudrait remplacer $\forall x \geq A$ par $\forall x \leq A$.

b) Faux, à droite et à gauche seulement.

c) Faux, les fonctions f et g n'admettent pas forcément de limites en a.

d) Vrai.

e) Faux, $x \mapsto x^2$ sur $\mathbb{R}$.

f) Faux avec, par exemple, la fonction f définie par $f(x) = 0$ si $x = \frac{\pi}{2}[\pi]$, $f(x) = \tan x$ sinon.

g) Faux, arctan.

h) Faux, arctan.

i) Vrai car sur un segment la borne supérieure d'une fonction continue est un plus grand élément.

j) Vrai car en 3 ce polynôme vaut -2 et il tend vers $+\infty$ en $+\infty$: c'est ensuite le théorème des valeurs intermédiaires.

k) Vrai d'après le théorème des valeurs intermédiaires.

l) Vrai d'après le théorème des valeurs intermédiaires.

m) Faux, avec la fonction continue $f : x \mapsto x^2$ et $I = [1,4]$: $f^{-1}(I) = [-2,-1] \cup [1,2]$.

n) Faux.

o) Faux pas aux entiers.

p) Vrai.

q) Faux, avec, par exemple, la fonction f définie par $f(x) = x$ si $x \in \mathbb{Q}$, et $f(x) = x + 1$ sinon.

r) Vrai.

s) Vrai.

t) Faux.

u) Vrai.

v) Faux, en revanche, elle est uniformément continue d'après le théorème de Heine.

w) Vrai.

Corrigés des exercices d'application du chapitre 7

Exercice 1

Grâce à une formule de trigonométrie, on peut tout d'abord réécrire, pour tout $x \neq 0$,

$$f(x) = -2\sin\left(\frac{x}{2}\right)\sin\left(\frac{x}{2}+\frac{1}{x}\right).$$

Ainsi, f est le produit d'une fonction de limite nulle en 0 (à savoir, $x \mapsto -2\sin\left(\frac{x}{2}\right)$) et d'une fonction bornée (à savoir, $x \mapsto \sin\left(\frac{x}{2}+\frac{1}{x}\right)$) donc est de limite nulle en 0. Ainsi, f est prolongeable par continuité en 0.

Exercice 2

1. D'après les règles usuelles de composition et d'addition, la fonction f est continue sur $\mathbb{R}\setminus\mathbb{Z}$. En $k \in \mathbb{Z}$, on vérifie

$$\lim_{x\to k^-} f(x) = \lim_{x\to k^-} 2(k-1)-\cos(3\pi(x-k+1)) = 2(k-1)+1 = 2k-1$$
$$\lim_{x\to k^+} f(x) = \lim_{x\to k^+} 2k-\cos(3\pi(x-k)) = 2k-1.$$

Par conséquent, f est également continue en tout entier donc sur tout $\mathbb{R}$.

2. Commençons par remarquer que si $y = f(x)$ alors $2\lfloor x\rfloor - 1 \leq y \leq 2\lfloor x\rfloor + 1$ donc

$$\frac{y-1}{2} \leq \lfloor x\rfloor \leq \frac{y+1}{2}.$$

▹ Si y n'est pas un entier impair, l'intervalle $\left[\frac{y-1}{2}, \frac{y+1}{2}\right]$ ne contient qu'un entier donc les antécédents x de y par f vérifient $\lfloor x\rfloor = \left\lfloor \frac{y+1}{2}\right\rfloor$ et

$$\cos 3\pi(x-\lfloor x\rfloor) = 2\left\lfloor \frac{y+1}{2}\right\rfloor - y \in]-1,1[.$$

Or, $t \mapsto \cos 3\pi t$ réalise des bijections de chacun des intervalles $]0,\frac{1}{3}[$, $]\frac{1}{3},\frac{2}{3}[$ et $]\frac{2}{3},1[$ dans $]-1,1[$ donc trois valeurs de $x-\lfloor x\rfloor$ qui satisfont la dernière condition. En conclusion, card $f^{-1}(\{y\}) = 3$.

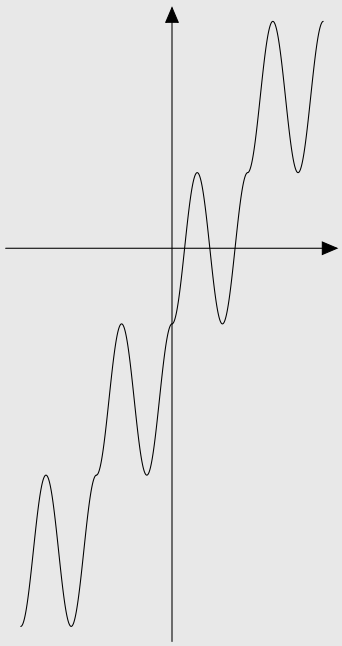
▹ Si $y = 2n+1$, alors les antécédents x de y par f vérifient $\lfloor x\rfloor \in \{n, n+1\}$.

- Si $\lfloor x\rfloor = n$, alors $\cos 3\pi(x-\lfloor x\rfloor) = -1$ donc $x = n+\frac{1}{3} = \frac{3y-1}{6}$.
- Si $\lfloor x\rfloor = n+1$, alors $\cos 3\pi(x-\lfloor x\rfloor) = 1$ donc $x = n = \frac{y-1}{2}$ ou $x = n+\frac{2}{3} = \frac{3y+1}{6}$.

En conclusion, card $f^{-1}(\{y\}) = 3$.

Voici le tracé de la courbe de f entre -2 et 2. On observe que chaque droite horizontale coupe exactement 3 fois la courbe.

CORRIGÉS



Exercice 3

Raisonnons par l'absurde. Avec une telle fonction, la suite $(f(n))_n$ est bornée donc on peut, avec le théorème de Bolzano-Weierstrass, en extraire une suite convergente $(f(\varphi(n)))_n$. En appliquant la fonction continue f^{-1}, on en déduit que $(\varphi(n))_n$ est convergente ce qui est absurde.

Exercice 4

Intuitivement, les fonctions croissantes ne peuvent avoir que des discontinuités vers le haut et deux discontinuités ne peuvent pas se compenser pour donner une continuité.

Soit f et g deux fonctions croissantes telles que $f+g$ soit continue. Soit $a \in \mathbb{R}$. D'après la monotonie, f et g admettent des limites à gauche et à droite en a. Par continuité de $f+g$,

$$\lim_{a^-} f + \lim_{a^-} g = \lim_{a^+} f + \lim_{a^+} g.$$

Comme $\lim\limits_{a^-} f \leq \lim\limits_{a^+} f$ et $\lim\limits_{a^-} g \leq \lim\limits_{a^+} g$, on a forcément $\lim\limits_{a^-} f = \lim\limits_{a^+} f$ et $\lim\limits_{a^-} g = \lim\limits_{a^+} g$ donc f et g sont continues en a.

Exercice 5

Commençons par chercher une expression explicite de $f(x)$. Soit $x \geq 0$. Notons $u_n = \frac{x^n}{n!}$. Alors, pour tout $n \in \mathbb{N}$,

$$\frac{u_{n+1}}{u_n} = \frac{x}{n+1}.$$

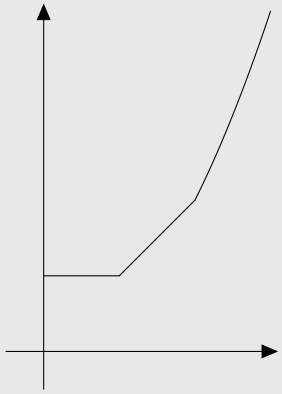
Ainsi, $\frac{u_{n+1}}{u_n} \geq 1$ si, et seulement si, $n+1 \leq x$. Par conséquent, $f(x) = u_{\lfloor x \rfloor} = \frac{x^{\lfloor x \rfloor}}{\lfloor x \rfloor !}$.

- La fonction f est continue sur chaque intervalle de $\mathbb{R}_+ \setminus \mathbb{N}$ comme composée de fonctions continues.

- Pour tout $k \in \mathbb{N}$,

$$\lim_{x \to k-} f(x) = \lim_{x \to k-} \frac{x^{k-1}}{(k-1)!} = \frac{k^{k-1}}{(k-1)!}$$
$$\lim_{x \to k+} f(x) = \lim_{x \to k+} \frac{x^{k}}{k!} = \frac{k^{k}}{k!} = \frac{k^{k-1}}{(k-1)!}$$

Par conséquent, f est continue en k.
En conclusion, f est continue sur $\mathbb{R}_+$.
Voici le tracé de la courbe de f entre 0 *et* 3.



Exercice 6

Par composition, la fonction g est continue sur $\mathbb{R} \setminus \mathbb{Z}$. De plus, pour tout $x \in \mathbb{R}$ et tout $k \in \mathbb{Z}$, $g(x+k) = g(x) + k$: il suffit donc d'étudier la continuité en 1. Or,

$$\lim_{1^-} g = f(1), \qquad \lim_{1^+} g = 1 + f(0).$$

Par conséquent, g est continue sur $\mathbb{R}$ si, et seulement si, $f(1) = f(0) + 1$.

Exercice 7

Soit $x < y$.

▹ Supposons que $x, y \in \mathbb{Q}$; il existe des entiers q et $a < b$ tels que $x = \frac{a}{q}$ et $y = \frac{b}{q}$. Comme la suite $(f(n\frac{1}{q}))_n$ est croissante, $f(x) \leq f(y)$.

▹ Sinon, il existe une suite de rationnels $(r_n)_n$ décroissante de limite x et une suite de rationnels $(r'_n)_n$ croissante de limite y. À partir d'un certain rang, $r_n \leq r'_n$ et donc, d'après le premier point, $f(r_n) \leq f(r'_n)$. En passant à la limite (grâce à la continuité de f), on obtient que $f(x) \leq f(y)$.

En conclusion, f est croissante.

Exercice 8

1. Pour tout $\varepsilon > 0$, notons $A_\varepsilon = \{x \in \mathbb{R}_+^*,\ f(x) > \varepsilon\}$. Remarquons tout d'abord que $A_\varepsilon \subset \mathbb{Q} \cap \mathbb{R}_+^*$ car l'image par f de tout irrationnel est nulle donc inférieure à ε. De plus, le rationnel d'écriture irréductible $\frac{p}{q}$ appartient à A_ε si, et seulement si, $\frac{1}{p+q} > \varepsilon$, c'est-à-dire $p + q < \frac{1}{\varepsilon}$; on a alors nécessairement $p, q \in \{0, \ldots, \lfloor \frac{1}{\varepsilon} \rfloor\}$ donc un nombre fini de rationnels dans A_ε.
En conclusion, pour tout $\varepsilon > 0$, A_ε est fini.

2. ▹ Si f est continue en un rationnel $x > 0$, alors il existe un voisinage de x inclus dans $A_{\frac{1}{2}f(x)}$. Or, d'après la question précédente, $A_{\frac{1}{2}f(x)}$ est fini donc ne contient pas d'intervalle ouvert non vide. Par conséquent, f n'est continue en aucun rationnel.
▹ Soit $x > 0$ irrationnel et $\varepsilon > 0$. L'ensemble A_ε est composé d'un nombre fini de rationnels ; notons $\eta = \frac{1}{2}\min_{y\in A_\varepsilon}|x-y|$. Par construction, pour tout $z \in]x-\eta, x+\eta[\cap\mathbb{R}_+^*$, $z \notin A_\varepsilon$ et donc

$$[f(z)-f(x)| = |f(z)| \leq \varepsilon.$$

Par conséquent, f est continue en chaque irrationnel de $\mathbb{R}_+^*$.

Exercice 9

Remarquons que la fonction $g : x \mapsto \frac{1}{x^2}$ ne s'annule pas. Alors, la fonction $h = \frac{f}{g}$ vérifie $h(x^2) = h(x)$ pour tout $x \in \mathbb{R}$. En utilisant la continuité en 1,

- Pour $x \in]0,1]$, $h(x) = h(x^{2^n}) \to h(1)$.
- Pour $x \geq 1$, $h(x) = h(x^{2^{-n}}) \to h(1)$.

Par conséquent, h est constante sur $\mathbb{R}_+^*$. Par ailleurs, h est paire car, pour tout $x \in \mathbb{R}$,

$$h(-x) = h((-x)^2) = h(x^2) = h(x).$$

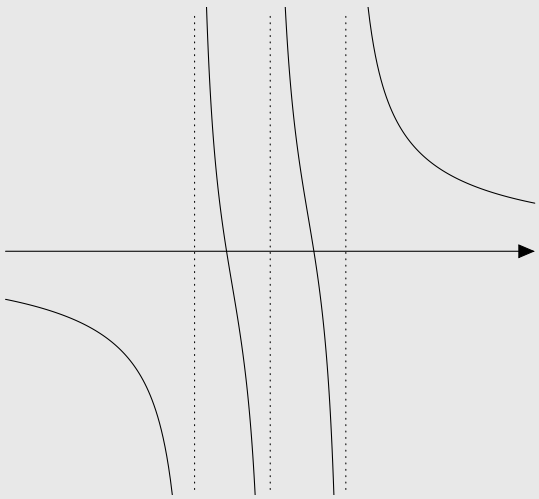
Ainsi, h est constante sur $\mathbb{R}^*$ et donc, il existe $a \in \mathbb{R}$ tel que $f : x \mapsto \frac{a}{x^2}$.

Exercice 10

La valeur moyenne $\frac{f(x_1)+\ldots+f(x_n)}{n}$ est comprise entre le plus grand des $f(x_k)$ et le plus petit des $f(x_k)$ donc entre deux valeurs prises par la fonction continue f : elle est donc atteinte d'après le théorème des valeurs intermédiaires.

Exercice 11

L'ensemble de définition de f est $\mathbb{R}\setminus\{a_1,\ldots,a_n\}$. Étudions les variations de f sur chacun des $n+1$ intervalles qui compose cet ensemble.
À titre d'illustration, voici la courbe pour $n = 3$, $a_1 = 1$, $a_2 = 2$, $a_3 = 3$.



▹ Sur chacun des intervalles $]a_k, a_{k+1}[$, la fonction est continue, strictement décroissante, $\lim\limits_{a_k^+} f = +\infty$, $\lim\limits_{a_{k+1}^-} f = -\infty$: f réalise donc une bijection de $]a_k, a_{k+1}[$ dans $\mathbb{R}$ donc s'annule exactement une fois sur $]a_k, a_{k+1}[$.
▹ Sur l'intervalle $]-\infty, a_1[$, la fonction est continue, strictement décroissante, $\lim\limits_{-\infty} f = 0$, $\lim\limits_{a_1^-} f = -\infty$: f réalise donc une bijection de $]-\infty, a_1[$ dans $\mathbb{R}_-^*$ donc ne s'annule pas sur $]-\infty, a_1[$.
▹ Sur l'intervalle $]a_n, +\infty[$, la fonction est continue, strictement décroissante, $\lim\limits_{+\infty} f = 0$, $\lim\limits_{a_n^+} f = +\infty$: f réalise donc une bijection de $]a_n, +\infty[$ dans $\mathbb{R}_+^*$ donc ne s'annule pas sur $]a_n, +\infty[$.
En conclusion, f s'annule exactement $n-1$ fois.

Exercice 12

1. Il s'agit d'une somme télescopique

$$\begin{aligned}\sum_{k=0}^{n-1} g\left(\frac{k}{n}\right) &= \sum_{k=0}^{n-1}\left(f\left(\frac{k+1}{n}\right) - f\left(\frac{k}{n}\right)\right)\\ &= f(1) - f(0) = 0.\end{aligned}$$

2. Il s'agit ici de montrer que la fonction g s'annule.
Si la fonction g s'annule en tous les réels de la forme $\frac{k}{n}$, le résultat est établi. Sinon, d'après la question précédente, il existe au moins deux valeurs de g de signes différents et le théorème des valeurs intermédiaires appliqué à la fonction continue g donne le résultat.

Exercice 13

Considérons la fonction $f : x \mapsto \frac{1}{n}\sum_{k=1}^{n} |x - a_k|$ définie et continue sur $[0,1]$. Alors,

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

Si $f(0) = \frac{1}{2}$, le résultat est établi avec $a = 0$; sinon, $\frac{1}{2}$ est entre $f(0)$ et $1 - f(0) = f(1)$ et on conclut avec le théorème des valeurs intermédiaires.

Exercice 14

Si $f(0) = 0$, alors 0 est fixe. Sinon, $f(0) > 0$ et $\lim\limits_{x\to 0} \frac{f(x)}{x} = +\infty$. Le théorème des valeurs intermédiaires pour la fonction $x \mapsto \frac{f(x)}{x}$ indique qu'il existe $t \in \mathbb{R}_+$ tel que $f(t) = t$.

Exercice 15

Remarquons tout d'abord que comme f est continue injective sur un intervalle, alors f est strictement monotone.
▹ Supposons f strictement croissante et qu'elle n'est pas l'identité : ainsi, il existe $x \in \mathbb{R}$ tel que $f(x) \neq x$.
- Si $f(x) > x$, alors, en composant l'inégalité par f, on obtient, pour tout n,

$$f^n(x) > f^{n-1}(x) > \ldots > f(x) > x$$

ce qui contredit l'hypothèse sur f.
- Si $f(x) < x$, alors on obtient la même contradiction en se retournant les inégalités.

Ainsi, la seule fonction strictement croissante vérifiant l'hypothèse est l'identité.

▹ Supposons f strictement décroissante. Alors, f^2 est strictement croissante. Par hypothèse sur f, pour tout $x \in \mathbb{R}$, il existe n tel que $f^n(x) = x$ donc $(f^2)^n(x) = f^n(f^n(x)) = f^n(x) = x$. Ainsi, f^2 vérifie encore la condition de l'énoncé et, d'après le premier point, est égal à l'identité.

Exercice 16

La fonction $\frac{g}{f}$ est continue (car f ne s'annule pas) donc atteint son minimum sur $[0,1]$ en un réel a. D'où

$$\forall x \in [0,1], \qquad g(x) \geq \frac{g(a)}{f(a)} f(x).$$

Par hypothèse, $\frac{g(a)}{f(a)} > 1$.

Exercice 17

Supposons $f^{-1}(\{0\})$ fini. Cet ensemble est en particulier majoré par un certain $A > 0$.

Sur l'intervalle $]A, +\infty[$, la fonction continue f ne change plus de signe (car sinon le théorème des valeurs intermédiaires impliquerait l'existence d'un zéro, c'est-à-dire un élément de $f^{-1}(\{0\})$). Quitte à changer f en $-f$, on peut supposer la fonction f positive sur cet intervalle.

Par ailleurs, sur le segment $[0, A]$, la fonction continue f atteint sa borne inférieure M.

Toute valeur strictement inférieure à $\min(0, M)$ ne peut être atteinte ni sur le segment $[0, A]$ ni sur $]A, +\infty[$ ce qui contredit la surjectivité de f à partir de $\mathbb{R}_+$ et termine le raisonnement par l'absurde.

Exercice 18

Sur le segment $[a, x_0]$, la fonction continue f atteint son maximum en un réel u. D'après l'hypothèse sur x_0, on a $u < x_0$. Par ailleurs

$$\sup_{[a,u]} f = f(u) = \sup_{[a,x_0]} f.$$

Exercice 19

Comme f est T-périodique, $f(\mathbb{R}) = f([x, x+T])$ pour tout $x \in \mathbb{R}$.

Soit $\alpha \in [0, T]$ où la fonction continue f admet son maximum (sur le segment $[0, T]$) et β où la fonction continue f admet son minimum (sur le segment $[\alpha, \alpha + T]$)

▹ Si $\beta \leq \alpha + \frac{T}{2}$, alors

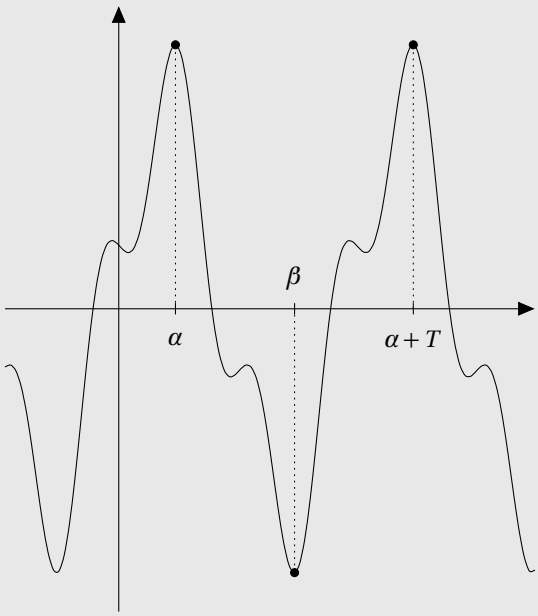
$$f(\mathbb{R}) = f([\alpha, \beta]) \subset f([\alpha, \alpha + \frac{T}{2}]) \subset f(\mathbb{R}).$$

D'où $f(\mathbb{R}) = f([\alpha, \alpha + \frac{T}{2}])$.

▹ Sinon $\alpha + \frac{T}{2} \leq \beta \leq \alpha + T$ donc $\beta - T \leq \alpha \leq \beta - \frac{T}{2}$; alors

$$f(\mathbb{R}) = f([\beta - T, \alpha]) \subset f([\beta - T, \beta - \frac{T}{2}]) \subset f(\mathbb{R}).$$

D'où $f(\mathbb{R}) = f([\beta - T, \beta - T + \frac{T}{2}])$.



Exercice 20

Comme f est uniformément continue, il existe $\eta > 0$ tel que

$$\forall x, y \in [a, b[, \qquad |x-y| \leq \eta \quad \Rightarrow \quad |f(x)-f(y)| \leq 1.$$

- f est bornée sur $[a, b-\eta]$ car continue sur ce segment.
- Pour tout $x \in [b-\eta, b[$, $|f(x)| \leq |f(b-\eta)|+1$. Ainsi, f est bornée sur $[b-\eta, b[$.

En conclusion, f est bornée sur $[a, b[$.

Exercice 21

Soit $\varepsilon > 0$. Il existe $A > 0$ tel que

$$\forall x \geq A, \qquad |f(x)| \leq \frac{1}{2}\varepsilon.$$

Par ailleurs, sur $[0, A+1]$, f est uniformément continue d'après le théorème de Heine : il existe $\eta > 0$ tel que

$$\forall x, y \in [0, A+1], \qquad |x-y| \leq \eta \quad \Rightarrow \quad |f(x)-f(y)| \leq \varepsilon.$$

Soit $x,\ y \in \mathbb{R}_+$ tels que $|x-y| \leq \min(\eta, \frac{1}{2})$.

- Si x et $y \in [0, A+1]$, alors $|f(x)-f(y)| \leq \varepsilon$.
- Sinon, l'un des deux est dans $[A+1, +\infty[$ donc $x,\ y \in [A, +\infty[$. Alors

$$|f(x)-f(y)| \leq |f(x)|+|f(y)| \leq \varepsilon.$$

Par conséquent, f est uniformément continue.

Corrigés des exercices d'approfondissement du chapitre 7

Exercice A

▹ Remarquons tout d'abord que, quitte à remplacer f par $f+\pi$, on peut se limiter au cas où $\ell \in [0,1]$.

▹ Supposons $\ell = 1$, c'est-à-dire $\cos(f(x)) - 1 \underset{x\to+\infty}{\longrightarrow} 0$. D'après la formule de duplication de cos, ceci entraîne immédiatement $\sin(\frac{f(x)}{2}) \underset{x\to+\infty}{\longrightarrow} 0$.

Par conséquent, pour $\varepsilon > 0$, il existe A tel que, pour tout $x \geq A$,

$$\left|\sin\frac{f(x)}{2}\right| \leq \varepsilon.$$

En d'autres termes, pour tout $x \geq A$,

$$\frac{f(x)}{2} \in \bigcup_{k\in\mathbb{Z}} [-\arcsin\varepsilon + k\pi, \arcsin\varepsilon + k\pi].$$

Les intervalles sont deux à deux disjoints (car $2\arcsin\varepsilon < \pi$) donc, d'après le théorème de valeurs intermédiaires, il existe un entier k tel que

$$f([A,+\infty[) \subset [-2\arcsin\varepsilon + 2k\pi, 2\arcsin\varepsilon + 2k\pi]$$

ou encore, pour $a \geq A$,

$$f(x) - 2k\pi \in [-2\arcsin\varepsilon, 2\arcsin\varepsilon] \subset [-\pi,\pi].$$

Par conséquent, pour tout $x \geq A$,

$$\begin{aligned}\frac{f(x)}{2} - k\pi &= \arcsin(\sin(\frac{f(x)}{2} - k\pi)) \\ &= \arcsin((-1)^k \sin\frac{f(x)}{2})\end{aligned}$$

Ainsi, $f(x) \to 2k\pi$.

▹ Supposons dorénavant $\ell \in [0,1[$ et considérons θ tel que $\ell = \cos\theta$. Par hypothèse, $\cos f(x) \to \cos\theta$ donc $\sin^2 f(x) \to \sin^2\theta$. Comme $\sin\theta \neq 0$ (car $\ell \in [0,1[$), on en déduit $\sin f(x) \to \sin\theta$ ou $\sin f(x) \to -\sin\theta$ (propriété des valeurs intermédiaires). Supposons, quitte à changer f en $-f$ que le premier cas est réalisé. Alors,

$$\begin{aligned}\cos(f(x)+\pi-\theta) &= \sin f(x)\sin(\pi-\theta) - \cos f(x)\cos(\pi-\theta) \\ &= \sin f(x)\sin\theta + \cos f(x)\cos\theta\end{aligned}$$

tend vers 1. D'après le cas précédent, $f+\pi-\theta$ converge en $+\infty$ donc f converge en $+\infty$.

Exercice B

Soit $x < y$.

▹ Pour tout $t \geq y$, $f(t) \geq f(y)$ d'où $\lim\limits_{y^+} f \geq f(y)$.

▹ Pour tout $t \in [x,y]$, $f(t) \leq f(y)$ d'où $\lim\limits_{x^+} f \leq f(y)$.

En combinant ces deux inégalités, on obtient $\lim\limits_{x^+} f \leq \lim\limits_{y^+} f$ d'où la propriété de croissance recherchée.

Exercice C

1. Soit f une telle fonction et $x \in \mathbb{R}$. Pour tout $\varepsilon > 0$, il existe $\eta > 0$ tel que
$$\forall x' \in [x-\eta, x], \qquad |f(x') - f(x)| \leq \varepsilon.$$
Par conséquent, en passant à la borne supérieure,
$$\left| \sup_{x' \in [x-\eta, x]} f(x') - f(x) \right| \leq \varepsilon.$$
Comme $\sup\limits_{x' \in [x-\eta, x]} f(x') = \sup\limits_{x' \in]-\infty, x]} f(x') = x$, on a $|x - f(x)| \leq \varepsilon$. Comme la propriété est vraie pour tout $\varepsilon > 0$, $f(x) = x$: la seule fonction continue qui vérifie cette propriété est l'identité.
2. Soit A une partie dense de $\mathbb{R}$ distincte de $\mathbb{R}$ (par exemple $\mathbb{Q}$). Il suffit de considérer
$$x \mapsto \begin{cases} x & \text{si } x \in A, \\ x-1 & \text{sinon.} \end{cases}$$

Exercice D

L'idée simple est de montrer que f ne peut pas avoir de « sauts vers le bas ».

▹ Supposons qu'il existe $x \in [a, b[$ tel que
$$\exists \varepsilon > 0,\ \forall \eta > 0,\ \exists y \in [x, x+\eta[, \qquad f(x) \geq f(y) + \varepsilon.$$
Écrivons la continuité de g en x pour ce ε :
$$\exists \eta > 0, \forall y \in]x-\eta, x+\eta[, \qquad |g(y) - g(x)| \leq \frac{1}{2}\varepsilon.$$
Considérons η d'après cette deuxième propriété et y choisi par la première. Alors
$$f(y) + g(y) \geq f(x) + g(x) \geq f(y) + \varepsilon + g(y) - \frac{1}{2}\varepsilon = f(y) + g(y) + \frac{1}{2}\varepsilon.$$
Contradiction.

▹ Supposons qu'il existe $x \in]a, b]$ tel que
$$\exists \varepsilon > 0,\ \forall \eta > 0,\ \exists y \in]x-\eta, x], \qquad f(x) + \varepsilon \geq f(y).$$
Écrivons la continuité de g en x pour ce ε :
$$\exists \eta > 0, \forall y \in]x-\eta, x+\eta[, \qquad |g(y) - g(x)| \leq \frac{1}{2}\varepsilon.$$
Considérons η d'après cette deuxième propriété et y choisi par la première. Alors
$$f(y) + g(y) \leq f(x) + g(x) \leq f(y) - \varepsilon + g(y) + \frac{1}{2}\varepsilon = f(y) + g(y) - \frac{1}{2}\varepsilon.$$
Contradiction.

▹ Passons à la conclusion. Soit $c = \sup\{x \in [a, b],\ f(x) \geq z\}$ (qui est bien défini car cet ensemble contient a et est majoré par b). Il existe par définition de la borne supérieure, une suite (c_n) de limite c telle que, pour tout $n \in \mathbb{N}$, $f(c_n) \geq z$. D'après ce que l'on a prouvé précédemment, pour tout $\varepsilon > 0$, il existe $\eta > 0$ tel que pour tout $y \in]c-\eta, c]$, $f(c) + \varepsilon > f(y)$. En particulier, pour n suffisamment grand, $f(c) + \varepsilon > f(c_n) \geq z$. Comme le résultat est vrai pour tout $\varepsilon > 0$, on a $f(c) \geq z$.

Par ailleurs, il existe une suite (d_n) à valeurs supérieures à c telle que, pour tout $n \in \mathbb{N}$, $f(d_n) \leq z$. D'après ce que l'on a prouvé précédemment, pour tout $\varepsilon > 0$, il existe $\eta > 0$ tel que pour tout $y \in [c, c+\eta[$, $f(c) < \varepsilon + f(y)$. En particulier, pour n suffisamment grand, $f(c) < \varepsilon + f(d_n) \leq \varepsilon + z$. Comme le résultat est vrai pour tout $\varepsilon > 0$, on a $f(c) \leq z$.

En conclusion, $f(c) = z$.

Exercice E

Soit $y \in f(\mathbb{R})$ et a, b les antécédents de cette valeur. La fonction continue f atteint ses extremums sur le segment $[a,b]$ et l'un au moins, disons un maximum, est atteint en $c \in]a,b[$. Notons d l'autre antécédent de $f(c)$.

- Si $d \in]a,b[$, alors f atteint son minimum sur l'intérieur du segment d'extrémités c et d (sinon f serait constante sur ce segment). Toute valeur entre ce minimum et $f(c)$ est atteinte plus de deux fois.
- Si $d > b$, alors toute valeur entre $f(a)$ et $f(c)$ est atteinte trois fois d'après le théorème des valeurs intermédiaires (sur $[a,c]$, $[c,b]$ et $[b,d]$).
- Si $d < a$, on procède de même.

On aboutit dans tous les cas à une contradiction.

Exercice F

Posons $g : x \mapsto f(x)+x$. La fonction g est continue et à valeurs dans $\mathbb{R}\setminus\mathbb{Q}$ donc d'après le théorème des valeurs intermédiaires est constante. Ainsi, il existe $a \in \mathbb{R}\setminus\mathbb{Q}$ tel que pour tout $x \in \mathbb{R}$, $f(x) = a - x$. Mais dans ce cas, $f(2a) = -a \notin \mathbb{Q}$: contradiction.

Exercice G

1. La fonction f est continue sur le segment $[c,d]$ donc y atteint son maximum $\max_{[c,d]} f$. Notons A l'ensemble des $x \in [c,d]$ tel que $f(x)$ est égal à $\max_{[c,d]} f$ et $t = \sup A$.

▹ Montrons que $t \in A$, c'est-à-dire que $f(t)$ est égal à $\max_{[c,d]} f$.

Pour cela, introduisons une suite $(x_n)_n$ à valeurs dans A et de limite t (ce qui est possible d'après la définition de borne supérieure. Pour tout n, $f(x_n) = \max_{[c,d]} f$. En passant à la limite et avec la continuité de f, on obtient $f(t) = \max_{[c,d]} f$.

En particulier, pour tout $y \in]t,d]$, $f(y) < f(t)$.

▹ Supposons $f(t) > f(d)$. Comme $d \in O$, pour tout $y > d$, $f(y) < f(d)$ donc $f(y) < f(t)$.

Ainsi, pour tout $y > t$, $f(y) < f(t)$ donc $t \notin O$: contradiction.

2. En faisant tendre t vers c dans l'inégalité de la question précédente et en utilisant la continuité de f, on obtient $f(c) \leq f(d)$.

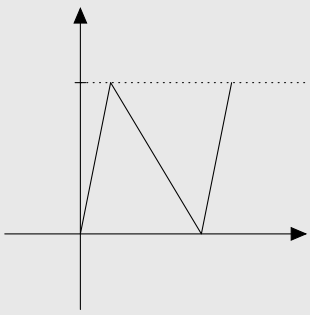
Or, comme $c \notin O$, pour tout $t > c$, $f(t) \leq f(c)$; en particulier, $f(d) \leq f(c)$ donc, avec l'inégalité précédente, $f(c) = f(d)$.

Exercice H

1. Soit $f : [0,1] \to]0,1[$ continue. L'ensemble $f([0,1])$ est un segment donc ne peut être égal à $]0,1[$: f n'est pas surjective.

2. Il suffit de choisir une fonction continue qui atteint son maximum 1 et son minimum 0 à l'intérieur de l'intervalle $]0,1[$.

Voici un exemple de telle fonction affine par morceaux



Exercice I

Sur le segment $[a,b]$, la fonction f atteint son maximum en $c \in [a,b]$.
Si $c \in]a,b[$, il est évident que $f(c) = \sup\limits_{[a,b]} f = \sup\limits_{]a,b[} f$.
Si $c = a$ (on procède de même pour $c = b$), on utilise la continuité de f en c : pour tout $\varepsilon > 0$, il existe $\eta > 0$ tel que

$$\forall x \in]c, c+\eta[, \qquad f(c) - \varepsilon \leq f(x) \leq f(c).$$

Ainsi, $f(c)$ est un majorant de $f(]a,b[)$ mais, pour tout $\varepsilon > 0$, $f(c) - \varepsilon$ n'est pas un majorant de $f(]a,b[)$: en conclusion,

$$\sup_{]a,b[} f = f(c) = \sup_{[a,b]} f.$$

Corrigés des problèmes du chapitre 7

Problème 1

1. a. Une fonction continue sur un intervalle est SCI car $|f(x) - f(x_0)| \leq \varepsilon$ implique $f(x) \geq f(x_0) - \varepsilon$.
b. La fonction f_0, définie par $f(x) = -x$ pour $x \leq 0$ et $f(x) = x+1$ si $x > 0$, est SCI mais n'est pas continue.
2. a. Soit f, g deux fonctions SCI $x_0 \in I$ et $\varepsilon > 0$. Il existe $\eta > 0$ et $\eta' > 0$ tels que, pour tout $x \in I$:

$$\begin{array}{lcl} |x - x_0| \leq \eta & \Rightarrow & f(x) \geq f(x_0) - \varepsilon \\ |x - x_0| \leq \eta' & \Rightarrow & g(x) \geq g(x_0) - \varepsilon \end{array}$$

D'où, pour tout $x \in I$ tel que $|x - x_0| \leq \min(\eta, \eta')$, $f(x) + g(x) \geq f(x_0) + g(x_0) - 2\varepsilon$ et $f + g$ est SCI.
b. Le produit de deux fonctions SCI n'est pas forcément SCI : en effet, la fonction constante égale à -1 et la fonction f_0 exhibée à la question 1. sont SCI et pourtant $-f_0$ n'est pas SCI.
3. a. Si f n'est pas minorée, il existe une suite $(x_n)_n$ telle que, pour tout $n \in \mathbb{N}$, $f(x_n) \leq -n$. D'après le théorème de Bolzano-Weierstrass (car I est un segment), on peut extraire une sous-suite $(x_{\varphi(n)})_n$ convergente vers l. Alors, l'hypothèse de semi-continuité inférieure implique qu'à partir d'un certain rang, $f(x_{\varphi(n)}) \geq f(l) - \varepsilon$ ce qui contredit la définition de la suite $(x_n)_n$.
b. Comme f est minorée, elle admet une borne inférieure m. Par définition de la borne inférieure, il existe une suite $(x'_n)_n$ de I telle que $m + \frac{1}{n} \geq f(x'_n) \geq m$. D'après le théorème de Bolzano-Weierstrass (car I est un segment), on peut extraire une sous-suite $(x'_{\varphi(n)})_n$ convergente ; notons l sa limite. L'hypothèse de semi-continuité inférieure de f implique que l'on a, à partir d'un certain rang, $f(x'_{\varphi(n)}) \geq f(l) - \varepsilon$ d'où $m + \frac{1}{n} + \varepsilon \geq f(l) \geq m$ et donc $f(l) = m$.

Problème 2

1. Avec l'indication de l'énoncé, on a, pour tout (x,y) tel que $|x - y| \leq t$, $g(x) - g(y) \leq \sqrt{t}$; or cette borne est atteinte pour $y = 0$ et $x = t$. Par conséquent, $\omega_g(t) = \sqrt{t}$ pour tout $t \leq 1$.
2. a. Soit $t \leq t'$. Alors

$$\{|f(x) - f(y)|,\ |x - y| \leq t\} \subset \{|f(x) - f(y)|,\ |x - y| \leq t'\},$$

donc $\omega_f(t) \leq \omega_f(t')$: ω_f est croissante ; par définition, $\omega_f(0) = 0$.

b. Soit $(u,v)\in[0,1]^2$ tel que $u+v\leq 1$ et x,y tels que $|x-y|\leq u+v$; il existe z tel que $|x-z|\leq u$ et $|z-y|\leq v$, d'où

$$|f(x)-f(y)|\leq|f(x)-f(z)|+|f(z)-f(y)|\leq\omega_f(u)+\omega_f(v),$$

et par définition de la borne supérieure comme plus petit des majorants : $\omega_f(u+v)\leq\omega_f(u)+\omega_f(v)$.
c. f est continue sur le segment $[0,1]$ donc y est uniformément continue (théorème de Heine) ce qui se traduit par

$$\forall\varepsilon>0,\ \exists\eta>0,\qquad \omega_f(\eta)\leq\varepsilon.$$

Soit $\varepsilon>0$ et $\eta>0$ associées par la propriété d'uniforme continuité. Pour tous $t>t'$ tels que $t-t'\leq\eta$, les propriétés de sous-additivité et de croissance (établies aux questions précédentes) donnent

$$\omega_f(t)-\omega_f(t')\leq\omega_f(t-t')\leq\omega_f(\eta)\leq\varepsilon.$$

D'où la continuité de ω_f.
d. Si ω_f s'annule en $x_0>0$, alors ω_f est nulle sur $[0,x_0]$ d'après la croissance de la fonction positive ω_f. Pour tout $x\in[0,1]$, il existe $n\in\mathbb{N}$ et $y<x_0$ tels que $x=nx_0+y$. Alors, par sous-additivité,

$$\omega_f(x)\leq n\omega_f(x_0)+\omega_f(y)=0.$$

D'où ω_f est identiquement nulle.
De plus, $\omega_f=0$ si, et seulement si, f est constante.
3. Soit ω une fonction vérifiant les propriétés de l'énoncé. Montrons que ω est son propre module de continuité. En effet, pour tous $y\leq x\leq y+t$, on a :

$$\omega(x)-\omega(y)\leq\omega(y+t)-\omega(y)\leq\omega(t)=\omega(t)-\omega(0),$$

en utilisant successivement la croissance, la sous-additivité et la valeur en 0.
Par conséquent, $\omega_\omega(t)=\omega(t)$ pour tout $t\leq 1$.

Problème 3

1. Soit $a<b\in D_\varphi^*$, alors, pour tout $y\in[a,b]$, $ax-\varphi(x)\leq yx-\varphi(x)\leq bx-\varphi(x)$ pour tout $x\geq 0$ (et l'inégalité contraire pour $x<0$). Par conséquent $x\mapsto yx-\varphi(x)$ est bornée et $y\in D_\varphi^*$; d'où $[a,b]\subset D_\varphi^*$ et D_φ^* est un intervalle.
2. Si φ est continue sur le segment I, alors pour tout $y\in\mathbb{R}$, la fonction $x\mapsto xy-\varphi(x)$ est continue sur ce même segment donc bornée d'après le théorème des bornes atteintes). Ainsi, $D_\varphi^*=\mathbb{R}$.
3. On trouve successivement en étudiant les variations de $x\mapsto xy-\varphi(x)$.
a. $D_\varphi^*=\{0\}$ et $\varphi^*(0)=-k$.
b. $D_\varphi^*=\mathbb{R}_+$ et $\varphi^*(y)=\frac{y^\beta}{\beta}$ avec $\frac{1}{\alpha}+\frac{1}{\beta}=1$.
c. $D_\varphi^*=\mathbb{R}_+$ et $\varphi^*(y)=y\ln y-y$ si $y>0$ et $\varphi^*(0)=0$.
4. Soit $k\in\mathbb{R}$ et φ une fonction. Alors, $x\mapsto xy-\varphi(x)$ est majorée si, et seulement si, $x\mapsto xy-\varphi(x)-k$ est majorée, d'où $D_{\varphi+k}^*=D_\varphi^*$. La propriété $(\varphi+k)^*=\varphi^*-k$ vient de la définition de la borne supérieure.
5. Supposons que $I\subset\mathbb{R}_+$, et considérons $y\leq y'\in D_\varphi^*$. Pour tout $x\in I$, $x\geq 0$ donc

$$xy-\varphi(x)\leq xy'-\varphi(x)$$

et par conséquent $\varphi^*(y)\leq\varphi^*(y')$ et φ^* est croissante. De même, si $I\subset\mathbb{R}_-$, φ^* est décroissante.
6. Soit $y,y'\in D_\varphi^*$, $\lambda\in[0,1]$. Alors, pour tout $x\in I$:

$$x(\lambda y+(1-\lambda)y')-\varphi(x)=\lambda(xy-\varphi(x))+(1-\lambda)(xy'-\varphi(x))\leq\lambda\varphi^*(y)+(1-\lambda)\varphi^*(y').$$

D'où $\varphi^*(\lambda y+(1-\lambda)y')\leq\lambda\varphi^*(y)+(1-\lambda)\varphi^*(y')$ et φ^* est une fonction convexe.
7. a. Pour tout $x\in I$, $xy-\varphi(x)\leq\varphi^*(y)$ (car φ^* est la borne supérieure) d'où $xy-\varphi^*(y)\leq\varphi(x)$. Par conséquent, $I=D_{\varphi^*}^*$.

b. D'après la question précédente, $(\varphi^*)^* \leq \varphi$. Supposons que $(\varphi^*)^* \neq \varphi$, c'est-à-dire qu'il existe x_0 tel que $(\varphi^*)^*(x_0) < \varphi(x_0)$. Comme φ est convexe, il existe a et b tels que $(\varphi^*)^*(x_0) < a\,x_0 + b$ et

$$\forall x \in I, \qquad \varphi(x) > a\,x + b.$$

Alors, $\varphi^*(a) = \sup\limits_{x\in I}(x\,a - \varphi(x)) \leq -b$ et $a\,x_0 + b \leq a\,x_0 - \varphi^*(a) \leq (\varphi^*)^*(x_0)$ d'où la contradiction.
Par conséquent, $(\varphi^*)^* = \varphi$.

Problème 4

Partie A – Préliminaires

1. a. On trouve $g(0) = 1$, $g^2(0) = g(1) = 2$ et $g^3(0) = g(2) = 0$.
b. Soit $x \in \mathbb{R}$. Alors, $g(x) = x$ si, et seulement si, $-\frac{3}{2}x^2 + \frac{3}{2}x + 1 = 0$, soit $x = \frac{1}{2} \pm \frac{\sqrt{33}}{6}$.
2. Soit $c, d \in [\alpha, \beta]$ tels que $\phi(c) = \alpha$ et $\phi(d) = \beta$. Posons $\psi : x \mapsto \phi(x) - x$. Alors,

$$\psi(c) = \phi(c) - c = \alpha - c \leq 0 \qquad \psi(d) = \phi(d) - d = \beta - d \geq 0.$$

D'après le théorème des valeurs intermédiaires appliqué à la fonction continue ψ, il existe $x_0 \in [\alpha, \beta]$ tel que $\phi(x_0) = x_0$.
3. a. Il suffit de remarquer que $f^{3k+r}(x_0) = f^r(x_0)$ pour tout $k \in \mathbb{N}$.
b. L'inclusion $[f(a), f^2(a)] \subset [a, f^2(a)]$ provient de l'ordre supposé entre les trois éléments de l'orbite de a. L'inclusion $[a, f^2(a)] \subset f([f(a), f^2(a)])$ est une conséquence du théorème des valeurs intermédiaires appliqué à la fonction continue f.
L'existence du point fixe provient de l'application de la question précédente avec $\phi = f$ et le segment $[\alpha, \beta] = [f(a), f^2(a)]$.
c. On procède comme à la question précédente en montrant

$$[a, f^2(a)] \subset [a, f(a)] \subset f([a, f^2(a)])$$

Partie B – Construction

1. a. Les réels α et β appartiennent à $\phi(\mathbb{R})$ donc admettent des antécédents.
b. La partie $A = \{x \in [a, b],\ \phi(x) = \beta\}$ est non-vide (car $b \in A$) et minorée (par a) donc admet une borne inférieure.
Soit $(u_n)_n$ une suite d'éléments de A qui converge vers $\inf A$; alors, pour tout n, $\phi(u_n) = \beta$ or ϕ est continue d'où $\phi(\inf A) = \beta$ et la borne inférieure est un plus petit élément.
c. D'après le théorème des valeurs intermédiaires (appliqué à la fonction continue ϕ), $K \subset \phi([u, v])$. Soit $y \in \phi([u, v])$. Par définition de u et de v (et par continuité de ϕ), $y \in [\alpha, \beta] = K$.
2. a. On construit la suite $(I_k)_{k \leq n-2}$ par récurrence avec la question 1.
▹ Notons tout d'abord que $I_0 \subset f(\mathbb{R})$ donc il existe un segment $I_1 \subset I_0$ tel que $f(I_1) = I_0$.
▹ Supposons la suite satisfaisant **(C1)** et **(C2)** construite jusqu'au rang $k < n-2$; donc $I_k \subset I_{k-1} = f(I_k)$. D'après la question 1, il existe un segment $I_{k+1} \subset I_k$ tel que $f(I_{k+1}) = I_k$.
b. Pour tout $k \leq n-2$, $f^k(I_k) = f^{k-1}(I_{k-1}) = \ldots = f^0(I_0) = J_1$.
3. D'après le théorème des valeurs intermédiaires appliqué à la fonction continue f, $J_0 \subset f(J_1)$; par ailleurs, $f(J_1) = f(f^{n-2}(I_{n-2})) = f^{n-1}(I_{n-2})$.
En appliquant la question 1. à la fonction f^{n-1} il existe I_{n-1} tel que $f^{n-1}(I_{n-1}) = J_0$.
4. On a $J_1 \subset f(J_1) = f^n(I_{n-1})$; en appliquant la question 1. à la fonction f^n il existe I_n tel que $f^n(I_n) = J_1$.

Partie C – Théorème de Sarkovskii

1. $I_n \subset J_1 = f^n(I_n)$ et f^n continue donc f^n admet un point fixe d'après la question 2 de la Partie A.
2. Commençons par remarquer que, par construction, $\{f^k(x_0),\ k \in [\![0, n-2]\!]\} \subset J_1$ et que $f^{n-1}(x_0) \in J_0$.
- Si $x_0 = f^2(a)$, alors $f(x_0) = x_0 \notin J_1$: contradiction.
- Si $x_0 = f(a)$, alors $f^2(x_0) = x_0 \notin J_1$: contradiction.

3. Remarquons que $f^{n-1}(x_0) \neq x_0$. Par ailleurs, si x_0 est de période $k < n$, alors $k < n-1$ et

$$\{f^j(x_0),\ j \in \mathbb{N}\} = \{f^j(x_0),\ j \leq k-1\} \subset J_1.$$

Or, $f^{n-1}(x_0) \notin J_1$ d'où la contradiction.

Problème 5

Partie A – Exemples

1. Soit c un point fixe de f et montrons par récurrence sur $n \in \mathbb{N}^*$ que $f^n(c) = c$.
▹ Le résultat pour $n = 1$ est la définition de point fixe.
▹ Soit $n \in \mathbb{N}^*$ tel que $f^n(c) = c$. Alors,

$$f^{n+1}(c) = f^n(f(c)) = f^n(c) = c.$$

2. La fonction continue $x \mapsto a + b - x$ admet un 2-cycle en a.
3. La fonction continue $x \mapsto x$ n'admet pas de 2-cycle car tous les points sont fixes.
4. Les points fixes sont les solutions de l'équation $f(x) = x$, c'est-à-dire 0 et $\frac{\lambda-1}{\lambda}$ (cette dernière n'existe dans $[0,1]$ que pour $\lambda \geq 1$). Les 2-cycles sont les solutions de $f^2(x) = x$ distinctes des points fixes (qui sont aussi solutions de cette équation). Comme

$$f^2(x) - x = -x(\lambda x + 1 - \lambda)(\lambda^2 x^2 - \lambda(\lambda+1)x + \lambda + 1),$$

on obtient les 2-cycles suivants pour $\lambda > 3$

$$\frac{1}{2} + \frac{1}{2\lambda} \pm \frac{\sqrt{(\lambda+1)(\lambda-3)}}{2\lambda}.$$

Partie B – Résultats d'existence

1. Considérons $g : x \mapsto f(x) - x$. g est continue comme différence de fonctions continues, $g(a) = f(a) - a \geq 0$ (car f est à valeurs dans $[a,b]$) et $g(b) = f(b) - b \leq 0$. D'après le théorème des valeurs intermédiaires, g s'annule sur $[a,b]$ donc f admet un point fixe sur ce segment.
2. Considérons c' et $d' \in [c,d]$ tels que $f(c') = c$ et $f(d') = d$ puis l'application $g : x \mapsto f(x) - x$. g est continue comme différence de fonctions continues, $g(c') = f(c') - c' = c - c' \leq 0$, $g(d') = f(d') - d' = d - d' \geq 0$. D'après le théorème des valeurs intermédiaires, g s'annule sur $[c', d'] \subset [c,d]$ (ou $[d', c']$ si l'ordre est différent) donc f admet un point fixe sur ce segment.
3. Considérons c' et $d' \in [c,d]$ tels que $f(c') = c$ et $f(d') = d$. On suppose sans perte de généralité $c < d$ et $c' < d'$. Considérons

$$\begin{aligned} c'' &= \sup\{x \geq c',\ f(x) = c\}, \\ d'' &= \inf\{x \geq c'',\ f(x) = d\}. \end{aligned}$$

▹ Remarquons que par continuité de f, $f(c'') = c$ (car c'' est la limite d'une suite d'antécédents de c par f). De même, $f(d'') = d$.
▹ D'après le théorème des valeurs intermédiaires, $[c,d] \subset f([c'', d''])$.
▹ Supposons qu'il existe $x_0 \in [c'', d'']$ tel que $f(x_0) < c$. Alors, le théorème des valeurs intermédiaires pour la fonction continue f assure l'existence d'un réel $c''' \in [x_0, d'']$ tel que $f(c''') = c$ ce qui contredit la maximalité de c''.
De même, il n'existe pas $x_0 \in [c'', d'']$ tel que $f(x_0) > d$. En conclusion, $[c,d] = f([c'', d''])$.

4. Considérons un segment $J \subset [a,c]$ vérifiant $f(J) = [c,b]$ (qui existe d'après la question précédente). Les hypothèses de l'énoncé entraînent

$$J \subset [a,c] \subset f([c,b]) \subset f^2(J).$$

D'après la question 2, f^2 admet un point fixe en $x_0 \in J$.
▹ Si $f(x_0) \neq x_0$, alors x_0 est un 2-cycle de f.
▹ Si $f(x_0) = x_0$, alors $x_0 \in J \cap f(J) \subset [a,c] \cap [c,b]$ d'où $x_0 = c$.

Partie C – Lemme de Coppel (I)

1. Si $f^2(c) = c$, alors $f(c) = c$ car f n'admet pas de 2-cycle. Or $f(c) > c$ d'où la contradiction.
2. a. L'application $g : x \mapsto f^2(x) - x$ est continue et vérifie $g(a) = f^2(a) - a \geq 0$, $g(c) = f^2(c) - c < 0$. D'après le théorème des valeurs intermédiaires, g admet un zéro sur $[a,c[$ donc f y admet un point fixe.
b. On remarque que par continuité de f, d est un point fixe. Or, c n'est pas fixe. Par conséquent, $d < c$.
c. Si f admet un point fixe sur l'intervalle $]d,c[$, f^2 y admet aussi un point fixe ce qui contredit la maximalité de d.
d. Comme f n'admet pas de point fixe sur $]d,c[$, la fonction $x \mapsto f(x) - x$ est de signe constant sur $]d,c[$ et positive en c. Ainsi,

$$\forall x \in]d,c[, \qquad f(x) > x.$$

Comme f^2 n'admet pas de point fixe sur $]d,c[$, la fonction $x \mapsto f^2(x) - x$ est de signe constant sur $]d,c[$ et négative en c. Ainsi,

$$\forall x \in]d,c[, \qquad f^2(x) < x.$$

Ainsi, pour tout $x \in]d,c[$, $f(x) > x > f^2(x)$.
e. Comme f est continue et $f(d) < c$, il existe un voisinage (à droite) de d sur lequel f reste strictement inférieure à d. Pour tout e dans ce voisinage, $f(e) < c$. Par ailleurs, $f(e) > e > d$ (car $x \mapsto f(x) - x$ est strictement positive).
f. D'une part, $e \in]d,c[$ donc

$$f(e) > e > f^2(e).$$

D'autre part, $f(e) \in]d,c[$ donc

$$f^2(e) > f(e) > f^3(e).$$

On obtient une contradiction avec les positions relatives de $f(e)$ et $f^2(e)$. En conclusion, l'hypothèse $c > f^2(c)$ est irréaliste.
On a ainsi établi que $f^2(c) > c$.

Partie D – Lemme de Coppel (II)

1. Par construction, la suite finie $(u_n)_{n\in[k+1,N]}$ est décroissante mais pas la suite finie $(u_n)_{n\in[k,N]}$ donc $u_k \leq u_{k+1}$.
Par ailleurs, $u_1 > c$, et $u_2 > c$ d'après la partie C. Donc, pour tout $n < N$, $u_n > c \geq u_N$; en particulier, $u_k > u_N$.
2. D'une part, $f(u_k) = u_{k+1}$ et $f(u_{k+j}) = u_{k+j+1} \leq u_{k+j}$ donc (avec le théorème des valeurs intermédiaires) $[u_{k+j}, u_{k+1}] \subset f([u_k, u_{k+j}])$.
D'autre part, $f(u_{k+j}) = u_{k+j+1} < u_{k+j}$ et $f(u_{k+1}) = u_{k+2} \geq u_{k+j}$ (car $j \geq 2$ d'après la partie C). Donc (avec le théorème des valeurs intermédiaires), $[u_k, u_{k+j}] \subset f([u_{k+j}, u_{k+1}])$.
3. D'après la dernière question de la partie B, u_{k+j} est un point fixe (impossible par définition de k) ou f admet un 2-cycle.
Ainsi, si f n'admet pas de 2-cycle et $f(c) > c$ entraîne, pour tout $n \in \mathbb{N}^*$, $f^n(c) > c$.

Partie E – Théorème de convergence

1. Par une récurrence immédiate, pour tout $k \geq N$, $u_k = u_N$; la suite $(u_n)_n$ est donc stationnaire donc converge.
2. Par hypothèse, la suite est croissante à partir d'un certain rang et bornée donc converge.
3. a. Fixons n et considérons u_k égal à z_n. Par définition de z_n, $f(u_k) > u_k$. D'après la partie D, pour tout j, $f^j(u_k) > u_k$ donc tous les termes de la suite au-delà de u_k sont strictement supérieurs à u_k. En particulier $z_{n+1} > z_n$. De manière analogue, on a $y_{n+1} < y_n$.
Les suites $(y_n)_n$ et $(z_n)_n$ sont monotones, bornées donc convergent.
b. La suite $(u_n)_n$ n'est pas stationnaire ni monotone. Il existe dont une infinité de termes de la suite $(y_n)_n$ suivi (dans la suite $(u_n)_n$) par un terme de la suite $(z_n)_n$. Considérons la sous-suite de ces termes : elle converge vers $\lim y_n$ et son image par f est une sous-suite de $(z_n)_n$ d'où

$$f(\lim y_n) = \lim z_n.$$

De même, en échangeant les rôles de $(y_n)_n$ et $(z_n)_n$,

$$f(\lim z_n) = \lim y_n.$$

Comme f n'admet pas de 2-cycle, $\lim y_n = \lim z_n$.
c. Tous les termes de la suite $(u_n)_n$ appartiennent à l'une des suites $(y_n)_n$ et $(z_n)_n$ qui ont la même limite. Par conséquent, $(u_n)_n$ converge vers cette limite.
4. Dans tous les cas, la suite $(u_n)_n$ converge.

Chapitre 8

Fonctions dérivables

1. Fonctions une fois dérivables — **2.** Fonctions de régularité supérieure — **3.** Propriété des accroissements finis

Objectifs et compétences du programme

- Comprendre le caractère local de la définition de dérivation.
- Calculer des dérivées (et dérivées $n^{\text{ièmes}}$) de fonctions déduites des fonctions usuelles.
- Utiliser la dérivation pour obtenir des variations ou rechercher des extremums.
- Maîtriser les différentes propriétés des accroissements finis.

Michel Rolle, mathématicien français (1651-1719)
Disons le tout de suite, il n'a pas vraiment démontré le théorème désormais éponyme... en tout cas, pas avec toute la généralité où il est énoncé dans ce chapitre. Son travail sur ce sujet, publié en 1691, concernait, et c'est déjà beaucoup pour l'époque, les fonctions polynomiales réelles. Parmi ses sujets de prédilection, on peut citer la théorie des nombres et notamment l'étude des solutions entières à des équations algébriques. Enfin, on lui doit également la notation $\sqrt[n]{x}$ pour la racine $n^{\text{ième}}$ positive d'un réel x (c'est-à-dire pour $x^{\frac{1}{n}}$) et l'extension de la relation d'ordre naturelle aux nombres négatifs.

COURS

Après la continuité, on passe à l'étude de la dérivabilité des fonctions d'une variable réelle. Après les propriétés calculatoires (pour la dérivation ou les dérivations successives), on s'intéresse à une famille de résultats regroupés dans une partie intitulée « propriétés des accroissements finis » : les recherches d'extremas avec la dérivation, le théorème de Rolle, l'inégalité et le théorème des accroissements finis, le lien entre signe de la dérivée et sens de variation...

Dans tout ce chapitre, I désigne un intervalle.

1. Fonctions une fois dérivables

1.1. Définitions et premières propriétés

Définition 8.1.

Soit $f : I \to \mathbb{R}$ et $a \in I$.

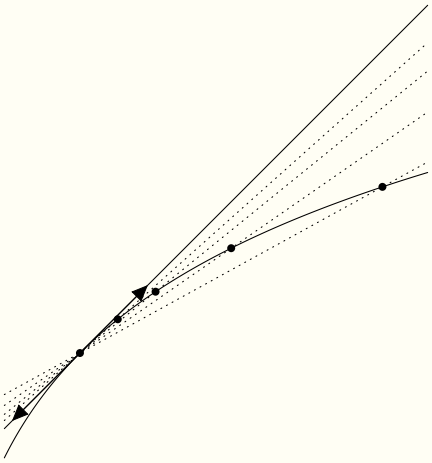
▹ La fonction f est dérivable en a si la fonction $x \longmapsto \frac{f(x)-f(a)}{x-a}$ admet une limite finie en a.

Cette limite est le nombre dérivé en a et notée $f'(a) = \frac{\mathrm{d}f}{\mathrm{d}x}(a)$.

▹ La fonction f est dérivable sur I si f est dérivable en tout point de I.

Remarque

La dérivabilité, comme la continuité, est une propriété locale : être dérivable sur un intervalle c'est être dérivable en chacun des points de celui-ci (ce qui se définit à partir de la notion de limite).



La dérivée en un point a s'interprète géométriquement comme la pente de la tangente puisqu'il s'agit de la limite de la pente de la corde joignant les points d'abscisses a et x lorsque $x \to a$.

Exemple

Soit $n \in \mathbb{N} \setminus \{0\}$. La fonction $f : x \mapsto x^n$ est dérivable en tout $a \in \mathbb{R}$. En effet, quand $x \to a$,

$$\frac{f(x)-f(a)}{x-a} = \sum_{k=0}^{n-1} x^k a^{n-1-k} \to \sum_{k=0}^{n-1} a^k a^{n-1-k} = na^{n-1}.$$

On retrouve la dérivée bien connue $f' : x \mapsto nx^{n-1}$.

Exemple

Soit $\alpha > 1$. La fonction $f : x \mapsto |x|^\alpha \cos \frac{1}{x}$ complétée par continuité en 0 est dérivable en 0. En effet, pour tout $x \neq 0$,

$$\frac{f(x)-f(0)}{x-0} = \frac{|x|^\alpha \cos \frac{1}{x}}{x} = |x|^{\alpha-1} \operatorname{sgn}(x) \cos \frac{1}{x}.$$

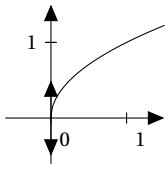
où $\operatorname{sgn}(x)$ est le signe du réel $x \neq 0$. Ce taux d'accroissement est donc de limite nulle comme produit d'une fonction de limite nulle et d'une fonction bornée.

Exemple

La fonction $f : x \mapsto \sqrt{x}$ n'est pas dérivable en 0 puisque, lorsque $x \to 0$,

$$\frac{\sqrt{x}-\sqrt{0}}{x-0} = \frac{\sqrt{x}}{x} \to +\infty.$$

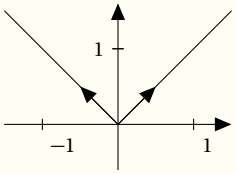
Cela se comprend graphiquement par la tangente verticale en 0 à la courbe représentative de f :



Remarque

Si on remplace la notion de limite par les limites à gauche ou à droite, on obtient les dérivées à gauche ou à droite traditionnellement notées $f'_g(a)$ et $f'_d(a)$.

Par exemple, la fonction valeur absolue $f : x \mapsto |x|$ n'est pas dérivable en 0 mais admet des dérivées à gauche et à droite $f'_g(0) = -1$ et $f'_d(0) = 1$.



Définition 8.2.

Soit $f : I \to \mathbb{R}$ et $a \in I$.

▹ La fonction f est dérivable par morceaux sur le segment $[a, b] \subset I$ s'il existe une subdivision

$a = a_0 < \ldots < a_n = b$ telle que, pour tout $k \in [\![1, n]\!]$, la restriction $f_{|]a_{k-1}, a_k[}$ admette un prolongement dérivable sur $[a_{k-1}, a_k]$.
▷ La fonction f est dérivable par morceaux sur I si elle est dérivable par morceaux sur tout segment inclus dans I.

Proposition 8.3.

Si une fonction $f : I \to \mathbb{R}$ est dérivable en a, alors f est continue en a.

Démonstration
Si $x \mapsto \frac{f(x)-f(a)}{x-a}$ admet une limite finie en a, alors le numérateur tend nécessairement vers 0 en a (puisque le dénominateur tend vers 0) ; par conséquent, f admet la limite $f(a)$ en a donc est continue en a. ■

Remarque
La réciproque est fausse comme on peut le voir avec la fonction $x \mapsto |x|$ continue mais non dérivable en 0.

Traduisons la définition de dérivabilité.

Proposition 8.4.

La fonction $f : I \to \mathbb{R}$ est dérivable en $a \in I$ si, et seulement s'il existe $(\alpha, \beta) \in \mathbb{R}^2$ et ε une fonction définie sur un voisinage V de a telle que $\lim\limits_a \varepsilon = 0$ tels que, pour tout $x \in V$,

$$f(x) = \alpha + \beta(x-a) + (x-a)\varepsilon(x).$$

Dans ce cas, $\alpha = f(a)$ et $\beta = f'(a)$.

Démonstration
($\Rightarrow$) La condition $\lim\limits_{x \to a} \frac{f(x)-f(a)}{x-a} - f'(a) = 0$ se traduit par $\frac{f(x)-f(a)}{x-a} - f'(a) = \varepsilon(x)$, avec ε une fonction telle que $\lim\limits_a \varepsilon = 0$. Ainsi, $f(x) = f(a) + f'(a)(x-a) + (x-a)\varepsilon(x)$.
($\Leftarrow$) En faisant tendre x vers a, on obtient $\lim\limits_a f = \alpha$: f est continue en a et $f(a) = \alpha$. De plus,

$$\frac{f(x)-f(a)}{x-a} = \beta + \varepsilon(x).$$

Ainsi, $\lim\limits_{x \to a} \frac{f(x)-f(a)}{x-a} = \beta$: f est dérivable en a et $f'(a) = \beta$. ■

Définition 8.5.

La fonction $f : I \to \mathbb{C}$ est dérivable en $a \in I$ si la fonction $x \mapsto \frac{f(x)-f(a)}{x-a}$ admet une limite finie en a.

Cette définition équivaut à la dérivabilité des fonctions réelles $\mathfrak{Re}(f)$ et $\mathfrak{Im}(f)$ en a. La dérivée de f en a est alors

$$f'(a) = \mathfrak{Re}(f)'(a) + i\mathfrak{Im}(f)'(a).$$

Ainsi, c'est cohérent avec la définition réelle et les propriétés des limites pour les combinaisons linéaires.

Définition 8.6.

La fonction $f : I \to \mathbb{R}$ est de classe $\mathscr{C}^1$ sur I si f est dérivable sur I et si f' est continue sur I.

Exemple

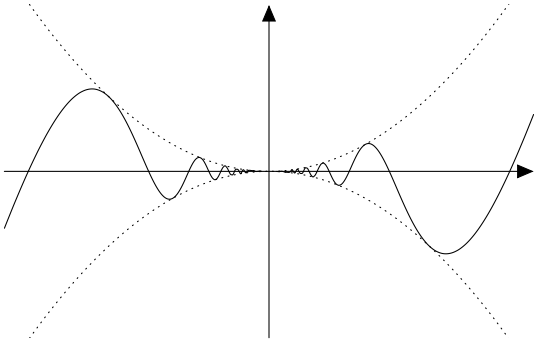
La fonction

$$f : x \mapsto \begin{cases} x^2 \sin \frac{1}{x} & \text{si } x \neq 0 \\ 0 & \text{sinon} \end{cases}$$

est dérivable en 0 mais n'est pas de classe $\mathscr{C}^1$. En effet, sa dérivée est

$$f' : x \mapsto \begin{cases} 2x \sin \frac{1}{x} - \cos \frac{1}{x} & \text{si } x \neq 0, \\ 0 & \text{sinon.} \end{cases}$$

Soit la suite $(u_n)_n$ définie par $u_n = \frac{1}{2n\pi + \pi/2}$. Alors, $u_n \to 0$ et $f'(u_n) = (4n+1)\pi \to +\infty$: ainsi, f' n'est pas continue en 0 en vertu de la caractérisation séquentielle des limites.
Traçons les courbes représentatives de f (et des fonctions $x \mapsto \pm x^2$ qui l'encadrent). On visualise bien qu'elle tend vers 0 en continuant à osciller « de plus en plus fort » ce qui indique que la dérivée n'admet pas de limite en 0.



1.2. Opérations

D'après les résultats sur les limites, on déduit la proposition suivante.

Proposition 8.7.

Soit $f, g : I \to \mathbb{R}$ deux fonctions dérivables en a, $\lambda, \mu \in \mathbb{R}$.
▹ La fonction $\lambda f + \mu g$ est dérivable en a et $(\lambda f + \mu g)'(a) = \lambda f'(a) + \mu g'(a)$.
▹ La fonction $f.g$ est dérivable en a et $(f.g)'(a) = f'(a)g(a) + f(a)g'(a)$.
▹ Si $g(a) \neq 0$, alors la fonction $\frac{f}{g}$ est dérivable en a et $\left(\frac{f}{g}\right)'(a) = \frac{f'(a)g(a) - f(a)g'(a)}{g(a)^2}$.

Proposition 8.8.

Soit $f : I \to J$ dérivable en a et $g : J \to \mathbb{R}$ dérivable en $f(a)$. Alors, la fonction composée $g \circ f$ est dérivable en a et $(g \circ f)'(a) = g' \circ f(a) f'(a)$.

Démonstration
Il suffit d'écrire la dérivabilité de g en $f(a)$ comme l'existence d'une fonction ε de limite nulle en $f(a)$ telle que

$$g(y)=g(f(a))+g'(f(a))(y-f(a))+\varepsilon(y)(y-f(a)),$$

puis la dérivabilité de f en a comme l'existence d'une fonction θ de limite nulle en a telle que

$$f(x)=f(a)+f'(a)(x-a)+\theta(x)(x-a),$$

Alors, pour tout $x\neq a$,

$$\begin{aligned}\frac{g\circ f(x)-g\circ f(a)}{x-a}&=\frac{g\circ f(a)+g'(f(a))(f(x)-f(a))+\big(f(x)-f(a)\big)\varepsilon(f(x))-g\circ f(a)}{x-a}\\&=g'(f(a))\frac{f(x)-f(a)+(f(x)-f(a))\varepsilon(f(x))}{x-a}\\&=g'(f(a))\left(f'(a)+\theta(x)+\frac{(f(x)-f(a))\varepsilon(f(x))}{x-a}\right)\\&=g'(f(a))f'(a)+\varphi(x),\end{aligned}$$

avec $\varphi : x\mapsto g'(f(a))\big(\theta(x)+\frac{(f(x)-f(a))\varepsilon(f(x))}{x-a}\big)$, fonction de limite nulle en a. Ainsi, quand $x\to a$,

$$\frac{g\circ f(x)-g\circ f(a)}{x-a}\to g'(f(a))f'(a).$$

■

D'après le calcul sur les fonctions puissances, on obtient immédiatement la dérivabilité des fonctions polynomiales.

Exemple

Les fonctions polynomiales sont dérivables et la dérivée de $x\mapsto\sum_{k=0}^{n}\alpha_k x^k$ est la fonction

$$x\mapsto\sum_{k=1}^{n}k\alpha_k x^{k-1}.$$

Par composition, les fonctions polynomiales trigonométriques sont dérivables.

Exemple

Une fonction homographique $x\mapsto\frac{ax+b}{cx+d}$ est dérivable en tout point de son ensemble de définition et sa dérivée est $x\mapsto\frac{ad-bc}{(cx+d)^2}$.

Corollaire 8.9.

Soit $f : I\to J$ dérivable en a, $g : J\to K$ dérivable en $f(a)$ et $h : K\to\mathbb{R}$ dérivable en $g(f(a))$. La fonction composée $h\circ g\circ f$ est dérivable en a et

$$(h\circ g\circ f)'(a)=h'\circ g\circ f(a)g'\circ f(a)f'(a).$$

Exemple

La fonction $f : x \mapsto \ln(\ln(\ln(x)))$ est dérivable sur son domaine de définition $]e,+\infty[$ comme composée de fonctions dérivables. Avec le corollaire précédent, sa dérivée est

$$f' : x \mapsto \ln'(\ln(\ln(x)))\ln'(\ln(x))\ln'(x) = \frac{1}{x\ln(x)\ln(\ln(x))}.$$

1.3. Dérivée d'une bijection réciproque

Terminons les propriétés des bijections réciproques après le résultat de continuité établi au chapitre précédent.

Proposition 8.10.

Soit $f : I \to \mathbb{R}$ continue, strictement monotone et dérivable en $a \in I$. Alors, la bijection réciproque $f^{-1} : f(I) \to I$ est dérivable en $f(a)$ si, et seulement si, $f'(a) \neq 0$. De plus,

$$(f^{-1})'(f(a)) = \frac{1}{f'(a)}.$$

Rappelons que nous avons déjà utilisé ces résultats pour déterminer les dérivées des fonctions trigonométriques réciproques :

$$\arccos'(x) = \frac{-1}{\sqrt{1-x^2}}, \qquad \arcsin'(x) = \frac{1}{\sqrt{1-x^2}}, \qquad \arctan'(x) = \frac{1}{1+x^2}.$$

Démonstration

($\Rightarrow$) Il suffit d'utiliser la règle de dérivation des fonctions composées pour obtenir $f'(a)(f^{-1})'(f(a)) = 1$ donc $f'(a) \neq 0$.

($\Leftarrow$) On remarque que

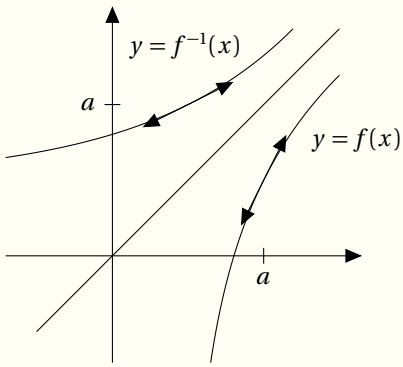
$$\lim_{y \to f(a)} \frac{f^{-1}(y) - f^{-1}(f(a))}{y - f(a)} = \lim_{x \to a} \frac{f^{-1}(f(x)) - f^{-1}(f(a))}{f(x) - f(a)} = \frac{1}{f'(a)}.$$

■

Remarque

Avec les notations de la proposition, on remarque que $(f^{-1})'(f(a))f'(a) = 1$ ou dit plus visuellement que le produit des pentes des tangentes aux points correspondants des courbes représentatives de f et f^{-1} est égal à 1.

Ce résultat ne devrait pas surprendre si l'on se souvient que ces courbes se déduisent l'une de l'autre par la symétrie par rapport à la première bissectrice et que l'on regarde la transformation d'une droite par cette transformation



2. Fonctions de régularité supérieure

2.1. Définition

Comme nous avons déjà défini les fonctions de classe $\mathscr{C}^0$ (les fonctions continues) et de classe $\mathscr{C}^1$, définissons récursivement les fonctions de classe $\mathscr{C}^n$ pour tout n.

Définition 8.11.

▹ Pour tout $n \in \mathbb{N}$, une fonction $f : I \to \mathbb{R}$ est de classe $\mathscr{C}^{n+1}$ sur I si f est dérivable sur I et f' est de classe $\mathscr{C}^n$.
▹ Une fonction $f : I \to \mathbb{R}$ est de classe $\mathscr{C}^\infty$ si f est de classe $\mathscr{C}^n$ pour tout $n \in \mathbb{N}$.

Remarque

Cette définition admet la définition équivalente suivante : f est de classe $\mathscr{C}^n$ si f est successivement n fois dérivable et si la dérivée $n^{\text{ième}}$ de f est continue.

La $n^{\text{ième}}$ dérivée est notée $f^{(n)}$ ou $\frac{\mathrm{d}^n f}{\mathrm{d}x^n}$. En particulier, $f^{(0)} = f$, $f^{(1)} = f'$ et $f^{(2)} = f''$.

La définition se réécrit $f^{(n+1)} = (f^{(n)})' = (f')^{(n)}$ et $f^{(0)} = f$.

Remarque

Plus généralement, les opérations qui conservent les fonctions dérivables préservent aussi les fonctions de classe $\mathscr{C}^n$.

Exemple

La dérivée $n^{\text{ième}}$ de $f : x \mapsto x^\alpha$ avec $\alpha \in \mathbb{R}$ est $f^{(n)} : x \mapsto \alpha(\alpha-1)\dots(\alpha-(n-1))x^{\alpha-n}$. Dans le cas où α est un entier supérieur à n, on peut réécrire ce résultat de manière plus « compacte » avec des factorielles

$$f^{(n)} : x \mapsto \frac{\alpha!}{(\alpha-n)!}\, x^{\alpha-n}.$$

Exemple

Déterminons la dérivée $n^{\text{ième}}$ de $f : x \mapsto e^{\sqrt{3}x}\sin(x)$. Pour cela remarquons que f est la partie imaginaire de la fonction $g : x \mapsto e^{(\sqrt{3}+i)x}$. Ainsi, pour tout $x \in \mathbb{R}$, on calcule la dérivée $n^{\text{ième}}$ de l'exponentielle g

$$g^{(n)}(x) = (\sqrt{3}+i)^n e^{(\sqrt{3}+i)x}$$

puis

$$\begin{aligned} f^{(n)}(x) = \mathfrak{Im}(g^{(n)}(x)) &= \mathfrak{Im}\left((\sqrt{3}+i)^n e^{(\sqrt{3}+i)x}\right) \\ &= 2^n \mathfrak{Im}\left(\left(\frac{\sqrt{3}}{2} + i\frac{1}{2}\right)^n e^{(\sqrt{3}+i)x}\right) \\ &= 2^n e^{\sqrt{3}x}\sin\left(x + n\frac{\pi}{6}\right). \end{aligned}$$

2.2. Formule de Leibniz

Proposition 8.12. Leibniz

Soit $f, g : I \to \mathbb{R}$ des fonctions n fois dérivables. Alors, la fonction produit $f.g$ est n fois dérivable et

$$(f.g)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} f^{(k)} g^{(n-k)}.$$

Cette formule ressemble beaucoup à la formule du binôme et pour cause, la preuve est formellement la même avec la récurrence et la formule du triangle de Pascal.

Démonstration

Procédons par récurrence. Le résultat est évident pour $n = 0$.

Soit $n \in \mathbb{N}$ et f, g des fonctions $n+1$ fois dérivables telles que $(f.g)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} f^{(k)} g^{(n-k)}$. Alors, $(f.g)^{(n)}$ est dérivable comme somme et produit de fonctions dérivables ; dérivons l'expression obtenue

$$\begin{aligned} (f.g)^{(n+1)} = ((f.g)^{(n)})' &= \sum_{k=0}^{n} \binom{n}{k} (f^{(k)} g^{(n-k)})' \\ &= \sum_{k=0}^{n} \binom{n}{k} (f^{(k+1)} g^{(n-k)} + f^{(k)} g^{(n+1-k)}) \\ &= \sum_{k=0}^{n} \binom{n}{k} f^{(k)} g^{(n+1-k)} + \sum_{k=1}^{n+1} \binom{n}{k-1} f^{(k)} g^{(n+1-k)} \\ &= f g^{(n+1)} + \sum_{k=1}^{n} \left[\binom{n}{k} + \binom{n}{k-1}\right] f^{(k)} g^{(n+1-k)} + f^{(n+1)} g \\ &= \sum_{k=0}^{n+1} \binom{n+1}{k} f^{(k)} g^{(n+1-k)}. \end{aligned}$$

■

Exemple Polynômes de Hermite

La fonction $f : x \mapsto e^{-x^2}$ est de classe $\mathscr{C}^\infty$. Pour tout $n \in \mathbb{N}$, $f^{(n)}$ s'écrit comme le produit de f par un polynôme H_n de degré n et de coefficient dominant $(-2)^n$. En effet, le résultat est établi pour $n \in \{0,1\}$. Procédons ensuite par récurrence en dérivant n fois la fonction dérivée f' :

$$f^{(n+1)}(x) = -2xf^{(n)}(x) - 2nf^{(n-1)}(x),$$

et donc $H_{n+1}(x) = -2xH_n(x) - 2nH_{n-1}(x)$.

Exemple

Déterminons la dérivée $n^{\text{ième}}$ de la fonction $f : x \mapsto x^{n-1}\ln x$ pour $n \geq 1$. Posons $g : x \mapsto x^{n-1}$ et $h : x \mapsto \ln(x)$ de sorte à avoir $f = g.h$. Alors,

$$\forall k \in \mathbb{N}, \qquad g^{(k)}(x) = \frac{(n-1)!}{(n-1-k)!}x^{n-1-k}.$$

$$\forall k \in \mathbb{N} \setminus \{0\}, \qquad h^{(k)}(x) = (-1)^{k-1}(k-1)!x^{-k}.$$

Ainsi, la formule de Leibniz donne

$$\begin{aligned} f^{(n)}(x) &= \sum_{k=1}^{n} \binom{n}{k} \frac{(n-1)!}{(k-1)!} x^{k-1}(-1)^{k-1}(k-1)!x^{-k} \\ &= \frac{(n-1)!}{x} \sum_{k=1}^{n} \binom{n}{k} (-1)^{k-1} \\ &= \frac{(n-1)!}{x}(1-(1-1)^n) = \frac{(n-1)!}{x}. \end{aligned}$$

3. Propriété des accroissements finis

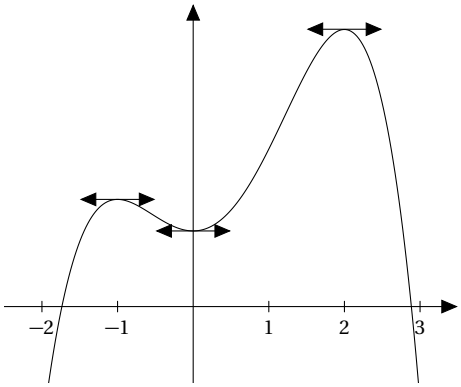
3.1. Extrema et théorème de Rolle

Définition 8.13.

Une fonction $f : I \to \mathbb{R}$ admet un extremum local en $a \in I$ s'il existe $\eta > 0$ tel que $f(a)$ est le plus petit ou le plus grand élément de $\{f(x),\ x \in]a-\eta, a+\eta[\cap I\}$.

Exemple

La fonction polynomiale $x \mapsto -\frac{1}{4}x^4 + \frac{1}{3}x^3 + x^2 + 1$ admet un maximum local en -1, un minimum local en 0 et un maximum global en 2.



En revanche, elle n'admet pas de minimum global sur $\mathbb{R}$.

Proposition 8.14.

Soit $f : I \to \mathbb{R}$ dérivable. Si f admet un extremum local en un point a intérieur à I, alors $f'(a) = 0$.

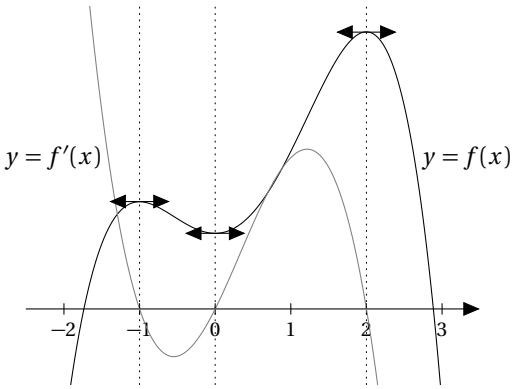
La réciproque est fausse : par exemple, la fonction $x \mapsto x^3$ est de dérivée nulle en 0 mais n'admet pas d'extremum local en 0.

Démonstration

De part et d'autre d'un extremum local, la fonction $x \mapsto \frac{f(x)-f(a)}{x-a}$ change de signe (car le numérateur est de signe constant au voisinage de a, alors que le dénominateur change de signe en a). Par conséquent, en passant à la limite (qui existe, d'après l'hypothèse de dérivabilité), on obtient $f'(a) = 0$. ■

Exemple

En ajoutant à l'exemple de la courbe de $f : x \mapsto -\frac{1}{4}x^4 + \frac{1}{3}x^3 + x^2 + 1$ celle de sa dérivée $f' : x \mapsto -x^3 + x^2 + 2x$, on vérifie l'annulation de f' aux extremums locaux.

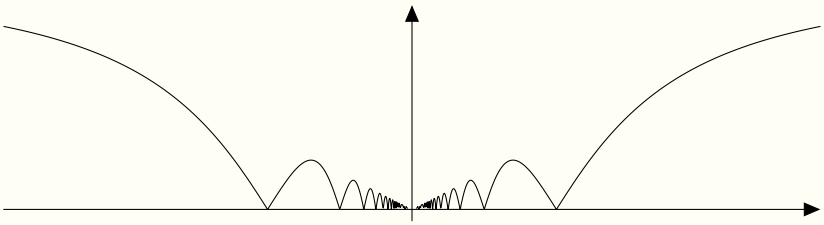


Conseils méthodologiques

Pour trouver les extremums locaux d'une fonction dérivable sur un intervalle, il suffit d'étudier les points où la dérivée s'annule et les éventuelles bornes de l'intervalle puis de déterminer s'ils sont ou non des extremums.

Remarque

Attention à ne pas faire une erreur en présumant l'existence de monotonie au voisinage d'un extremum. Par exemple, la fonction $x \mapsto \left|x \sin \frac{1}{x}\right|$ prolongée par continuité en 0 admet un minimum global en 0 mais n'est décroissante sur aucun intervalle de la forme $]-\eta, 0]$, ni croissante sur aucun intervalle de la forme $[0, \eta[$.



Théorème 8.15. Rolle

Soit $f : [a, b] \to \mathbb{R}$ telle que
- f est continue sur $[a, b]$,
- f est dérivable sur $]a, b[$,
- $f(a) = f(b)$.

Alors, il existe $c \in]a, b[$ tel que $f'(c) = 0$.

Démonstration

Si la fonction f est constante, le résultat est évident.

Sinon, comme f est continue sur le segment $[a, b]$, f atteint ses extremums dont l'un au moins n'est pas atteint au bord de l'intervalle (sinon la fonction est constante). D'après la proposition précédente, la dérivée f' s'annule en cet extremum. ■

Exemple

Si f est une fonction de classe $\mathscr{C}^n$ qui s'annule en $n+1$ points distincts, alors $f^{(n)}$ s'annule au moins une fois.
En effet, il suffit de procéder par récurrence.
▹ Le résultat est vrai pour $n = 0$ par hypothèse puisque $f^{(0)} = f$.
▹ Soit $n \in \mathbb{N}$; supposons le résultat vrai pour toute fonction de classe $\mathscr{C}^n$ qui s'annule en $n+1$ points distincts. Soit f une fonction de classe $\mathscr{C}^{n+1}$ qui s'annule en $n+2$ points distincts : sur chacun des $n+1$ intervalles déterminés par deux zéros consécutifs de f, f' s'annule d'après le théorème de Rolle donc $f^{(n+1)} = (f')^{(n)}$ s'annule au moins une fois.

Exemple Polynômes scindés à racines simples sur $\mathbb{R}$

▹ Soit $P \in \mathbb{R}[X]$ un polynôme de degré $n > 1$ scindé à racines simples : ainsi, P admet n racines distinctes.
En appliquant le théorème de Rolle sur chacun des $n-1$ intervalles déterminés par deux racines

consécutives, on obtient que P' admet $n-1$ racines sur $\mathbb{R}$ donc P' est lui aussi scindé à racines simples sur $\mathbb{R}$ (et par récurrence, il en est de même pour toutes les dérivées non constantes de P).
▹ On peut utiliser ce résultat pour obtenir des propriétés sur les coefficients des polynômes scindés à racines simples sur $\mathbb{R}$. Par exemple, un polynôme scindé à racines simples sur $\mathbb{R}$ ne peut avoir deux coefficients consécutifs nuls, car sinon il aurait une dérivée $k^{\text{ième}}$ qui se factorise par x^2 donc qui admet 0 comme racine double. Ainsi, il y aurait un polynôme scindé à racines simples dont une dérivée n'est pas scindée à racines simples : contradiction avec la propriété précédente.

Examinons un exemple avec des racines multiples.

Exemple Polynômes de Legendre

Soit $n \in \mathbb{N}$. Le polynôme de degré n défini par

$$L_n(x) = \frac{\mathrm{d}^n((x^2-1)^n)}{\mathrm{d}x^n}$$

admet n racines distinctes dans $]-1,1[$.
Posons $P(X) = (X^2-1)^n$ et montrons, par récurrence sur $k \leq n$, que $P^{(k)}$ admet au moins k racines dans $]-1,1[$ en utilisant que 1 et -1 sont racines de $P^{(k)}$ pour tout $k < n$ (simple vérification en écrivant $P = (x-1)^n Q$ puis en dérivant k fois).
▹ L'initialisation $k = 0$ est évidente.
▹ Soit $k \leq n-1$ tel que $P^{(k)}$ admet au moins k racines dans $]-1,1[$ donc $k+2$ dans $[-1,1]$ (en ajoutant -1 et 1). Appliquons le théorème de Rolle à $P^{(k)}$ sur chacun des $k+1$ intervalles délimités par deux de ces zéros consécutifs : $P^{(k+1)}$ s'annule au moins $k+1$ fois dans $]-1,1[$.

Remarque
En étant un peu plus soigneux et en combinant les idées des deux derniers exemples, on peut établir que la dérivée d'un polynôme réel scindé (pas forcément à racines simples) sur $\mathbb{R}$ est encore un polynôme scindé.

Voici un autre exemple d'application du théorème de Rolle.

Exemple

Soit f continue sur $[a,b]$, dérivable sur $]a,b[$ telle que $f(a) = f(b) = 0$. Alors, pour tout $\lambda \in \mathbb{R}$, il existe $c \in]a,b[$ tel que $f'(c) = \lambda f(c)$.
En effet, il suffit d'appliquer le théorème de Rolle à la fonction $g : x \mapsto e^{-\lambda x} f(x)$ continue sur $[a,b]$, dérivable sur $]a,b[$ et qui vérifie $g(a) = g(b) (= 0)$.

3.2. Propriété des accroissements finis

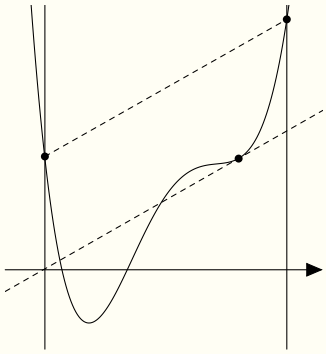
Proposition 8.16. Égalité des accroissements finis

Soit $f : [a,b] \to \mathbb{R}$ telle que
- f est continue sur $[a,b]$;
- f est dérivable sur $]a,b[$.

Alors, il existe $c \in]a,b[$ tel que $f(b) - f(a) = f'(c)(b-a)$.

Remarque

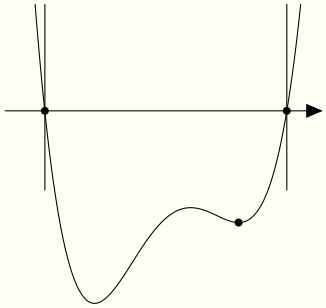
En termes géométriques, il existe un point de la courbe de f d'abscisse entre a et b où la tangente est parallèle à la corde entre les points de la courbe de f d'abscisses a et b.



Démonstration

Il suffit d'appliquer le théorème de Rolle à la fonction $x \mapsto f(x) - f(a) - \frac{f(b)-f(a)}{b-a}(x-a)$ qui a les bonnes propriétés de régularité et qui à la même valeur (nulle) en a et b. ■

L'idée de la preuve est relativement simple : on se ramène au cadre du théorème de Rolle en retranchant l'équation de la corde (ainsi, on « redresse » la courbe). Par exemple, la courbe représentative représentée ci-dessus devient après transformation



Commençons par le résultat classique sur les variations d'une fonction dérivable qui est une simple conséquence de l'égalité ci-dessus :

Proposition 8.17.

Soit f dérivable sur un intervalle I.
- f est constante si, et seulement si, $f' = 0$ sur I.
- f est croissante si, et seulement si, $f' \geq 0$ sur I.
- f est décroissante si, et seulement si, $f' \leq 0$ sur I.

Démonstration

Montrons la deuxième équivalence.

▹ Pour le sens direct, remarquons que si $y > x$, alors $\frac{f(y)-f(x)}{y-x} \geq 0$ donc en faisant tendre y vers x, $f'(x) \geq 0$.

▹ Pour le sens retour, considérons $y > x$ et écrivons la formule des accroissements finis entre x et y : il existe $c \in]x, y[$ tel que

$$f(y) - f(x) = f'(c)(y - x).$$

Comme $f'(c) \geq 0$ et $y \geq x$, on en déduit que $f(y) \geq f(x)$ ce qui est bien la croissance attendue. ■

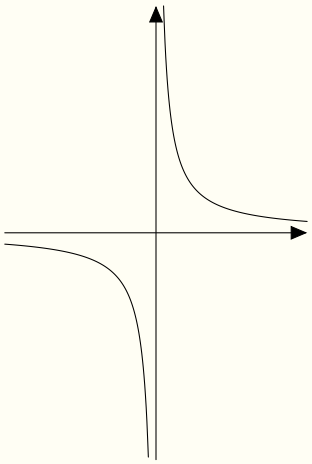
Remarque

Il n'y a plus de telle équivalence pour la stricte monotonie. La formule des accroissements finis (comme dans la preuve précédente) permet de montrer que si $f' > 0$ sur I, alors f est strictement croissante.

La réciproque est fausse comme on peut le constater avec la fonction strictement croissante $x \mapsto x^3$ dont la dérivée s'annule en 0.

Remarque

On ne peut pas en déduire de résultat « global » si le domaine de définition n'est pas un intervalle : par exemple, la fonction $x \mapsto \frac{1}{x}$ est à dérivée strictement négative sur $\mathbb{R}^*$ sans être décroissante sur tout $\mathbb{R}^*$ (mais elle l'est bien sur $\mathbb{R}_+^*$ et $\mathbb{R}_-^*$).



Cette remarque impose de résoudre les équations différentielles sur des intervalles pour pouvoir exploiter qu'une fonction à dérivée nulle y est constante. Par exemple, une fonction f à dérivée nulle sur l'ensemble $[0,1] \cup [2,3]$ n'est pas forcément constante mais de la forme

$$f : x \mapsto \begin{cases} A & \text{si } x \in [0,1] \\ B & \text{si } x \in [2,3] \end{cases}$$

avec A, $B \in \mathbb{R}$ des constantes (non forcément égales).

Remarque

On ne peut pas obtenir de résultat même « local » de la dérivée en un point : avoir une dérivée en un point strictement positive n'entraîne pas la croissance sur un intervalle autour de ce point.

Par exemple, la fonction

$$f : x \mapsto \begin{cases} x + x^2 \sin \frac{1}{x^2} & \text{si } x \neq 0, \\ 0 & \text{sinon,} \end{cases}$$

est continue sur $\mathbb{R}$, de classe $\mathscr{C}^1$ sur $\mathbb{R}^*$ et de dérivée en 0 donnée par

$$f'(0) = \lim_{x \to 0} \frac{f(x) - f(0)}{x - 0} = \lim_{x \to 0} 1 + x \sin \frac{1}{x^2} = 1 > 0.$$

Remarquons que, pour tout entier $n \geq 1$, $f'\left(\frac{1}{\sqrt{2n\pi}}\right) = 1 - 2\sqrt{2n\pi} < 0$.
Comme f est de classe $\mathscr{C}^1$ sur $\mathbb{R}^*$ (donc f' continue sur $\mathbb{R}^*$), il existe un voisinage de $\frac{1}{\sqrt{2n\pi}}$ sur lequel f' demeure strictement négative donc sur lequel f est strictement décroissante; tout voisinage de 0 contient un réel de la forme $\frac{1}{\sqrt{2n\pi}}$ donc un sous-intervalle sur lequel f est strictement décroissante.
En conclusion, cette fonction f n'est croissante sur aucun voisinage de 0 bien que $f'(0) = 1 > 0$.

Un autre conséquence de l'égalité des accroissements finis est la proposition suivante.

Proposition 8.18. Inégalité des accroissements finis

Soit $f : [a, b] \to \mathbb{R}$ telle que
- f est continue sur $[a, b]$,
- f est dérivable sur $]a, b[$,
- $|f'|$ est bornée par M sur $]a, b[$.

Alors, $|f(b) - f(a)| \leq M|b - a|$.

Ce résultat reste vrai pour une fonction à valeurs complexes (comme il sera établi plus loin dans le cours) contrairement à l'égalité des accroissements finis.

Démonstration
Il suffit d'utiliser l'hypothèse de majoration dans l'égalité des accroissements finis précédente. ■

Conseils méthodologiques

Les fonctions dérivables et à dérivée bornée vérifient les hypothèses de l'inégalité des accroissements finis donc sont lipschitziennes.
En particulier, les fonctions de classe $\mathscr{C}^1$ sont lipschitziennes sur tous les segments puisque leurs dérivées sont continues donc bornées sur les segments (par le théorème des bornes atteintes).

Proposition 8.19.

Soit $f : I \to \mathbb{R}$ continue sur I, dérivable sur $I \setminus \{a\}$ telle que f' admette une limite finie en a. Alors, f est dérivable sur I et $f'(a) = \lim\limits_{a} f'$.

Ce résultat est quelquefois appelé « prolongement » des fonctions de classe $\mathscr{C}^1$. Cette appellation est très maladroite puisque la dérivée existe bien en a et qu'on ne la prolonge pas.

Démonstration
Appliquons l'égalité des accroissements finis entre a et $x \in I$. Il existe c_x entre x et a tel que

$$\frac{f(x) - f(a)}{x - a} = f'(c_x).$$

Quand x tend vers a, c_x tend vers a (par le lemme d'encadrement). L'hypothèse sur la limite de f' donne

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a} = \lim_{x \to a} f'(x).$$

■

Exemple Fonction plate (lisse) en 0

Considérons la fonction $f : x \mapsto \begin{cases} e^{-\frac{1}{x}} & \text{si } x > 0, \\ 0 & \text{sinon.} \end{cases}$

Montrons que la fonction f est de classe $\mathscr{C}^\infty$ sur $\mathbb{R}$ en utilisant la propriété précédente pour toutes les dérivées. Remarquons tout d'abord que f est de classe $\mathscr{C}^\infty$ sur $\mathbb{R}^*$ donc il suffit d'étudier la régularité de f en 0.

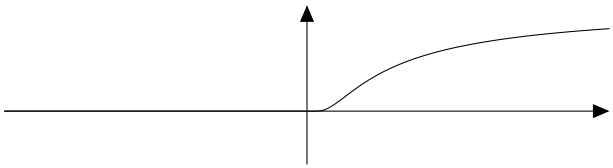
- La fonction f est continue en 0 car $e^{-\frac{1}{x}} \to 0$ quand $x \to 0$
- Pour tout $n \in \mathbb{N}$, il existe un polynôme P_n tel que

$$f^{(n)}(x) = \begin{cases} \frac{P_n(x)}{x^{2n}} e^{-\frac{1}{x}} & \text{si } x > 0, \\ 0 & \text{si } x < 0. \end{cases}$$

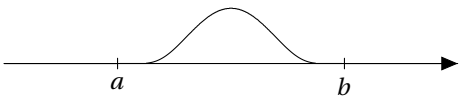
  En effet, le résultat est évident pour $n = 0$ et si la dérivée $n^{\text{ième}}$ est de cette forme, alors, pour tout $x > 0$,

$$f^{(n+1)}(x) = \frac{x^2 P_n'(x) + (1-2nx)P_n(x)}{x^{2(n+1)}} e^{-\frac{1}{x}},$$

  ce qui termine la récurrence avec $P_{n+1}(x) = x^2 P_n'(x) + (1-2nx)P_n(x)$.
- D'après le point précédent, pour tout $n \in \mathbb{N}$, $f^{(n)}(x) \to 0$ quand $x \to 0$. Ainsi, f est de classe $\mathscr{C}^\infty$ en 0 et toutes ses dérivées y sont nulles.



Cette fonction peut être utilisée pour construire des fonctions de classe $\mathscr{C}^\infty$ nulles en dehors d'un segment $[a, b]$. Par exemple, la fonction $x \mapsto f(a-x)f(b-x)$ ne s'annule pas en x si, et seulement si, $x - a > 0$ ou $b - x > 0$, c'est-à-dire si $x \in]a, b[$ (ce qui n'est pas apparent sur le tracé suivant à cause de l'écrasement de la courbe au voisinage de a et b).



Exemple

En guise d'application, résolvons sur $\mathbb{R}$ l'équation $x^2 y' + (x^2 - 1)y = 0$.

▹ Cette équation se réécrit sur $\mathbb{R}_+^*$ et $\mathbb{R}_-^*$ sous la forme $y' = \frac{1-x^2}{x^2} y$. On résout sur chaque intervalle et on obtient que les solutions sont de la forme

$$y : x \mapsto \begin{cases} Ae^{-\frac{1}{x}-x} & \text{si } x > 0 \\ Be^{-\frac{1}{x}-x} & \text{si } x < 0 \end{cases}$$

avec A et $B \in \mathbb{R}$.

▹ Ces solutions sont de classe $\mathscr{C}^\infty$ sur $\mathbb{R}^*$. Étudions le raccordement par continuité en 0. Remarquons tout d'abord que la limite à droite en 0 est nulle (car l'argument de l'exponentielle tend vers $-\infty$. En revanche, à gauche, si $B \neq 0$, la fonction n'admet pas de limite finie. Il n'y a

donc de raccordement continu que si, et seulement si, $B=0$.
▹ Étudions le caractère $\mathscr{C}^1$ de ces solutions. La limite à gauche de la dérivée en 0 est 0. La dérivée à droite est

$$\forall x>0, \qquad y'(x)=A\frac{x^2-1}{x^2}e^{-\frac{1}{x}-x}.$$

En particulier, $y'(x)\sim -1\frac{1}{x^2}e^{-\frac{1}{x}}$ en 0 donc $y'(x)\to 0$ quand $x\to 0$ d'après les croissances comparées. D'après le corollaire des accroissements finis, y est de classe $\mathscr{C}^1$.
▹ En conclusion, les solutions (de classe $\mathscr{C}^1$) sur $\mathbb{R}$ de l'équation $x^2y'+(x^2-1)y=0$ sont les fonctions

$$y : x \mapsto \begin{cases} Ae^{-\frac{1}{x}-x} & \text{si } x>0 \\ 0 & \text{si } x<0 \end{cases}$$

avec $A\in\mathbb{R}$.

3.3. Applications aux systèmes dynamiques discrets

Reprenons l'étude des suites récurrentes dans le cas d'une dynamique de classe $\mathscr{C}^1$.

Définition 8.20.

Soit $f : I\to I$ et $u_0\in I$. Le système dynamique discret de dynamique f et de condition initiale u_0 est la suite $(u_n)_n$ qui vérifie pour tout $n\in\mathbb{N}$, $u_{n+1}=f(u_n)$.

Remarque

De telles suites ont déjà été étudiées dans le chapitre sur les suites afin de déterminer leur monotonie (et donc leur convergence éventuelle). Rappelons que

- l'étude de la position de la courbe représentative de f par rapport à la première bissectrice détermine la monotonie du système dynamique correspondant $(u_n)_n$.
- Si ce critère ne suffit pas à conclure, la monotonie de la fonction f permet d'obtenir la monotonie des suites $(u_{2n})_n$ et $(u_{2n+1})_n$.

Proposition 8.21.

Soit $f : I\to I$ continue et $u_0\in I$. Si le système dynamique discret de dynamique f et de condition initiale u_0 converge vers ℓ, alors ℓ est un point fixe de f (c'est-à-dire $f(\ell)=\ell$).

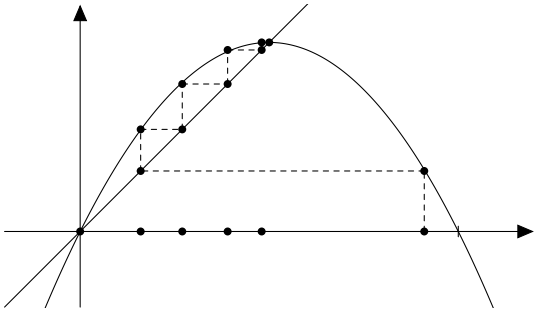
Démonstration

Pour tout $n\in\mathbb{N}$, $u_{n+1}=f(u_n)$. En passant à la limite et en utilisant la continuité de f en ℓ, on obtient immédiatement que la limite ℓ est un point fixe, c'est-à-dire $f(\ell)=\ell$. ■

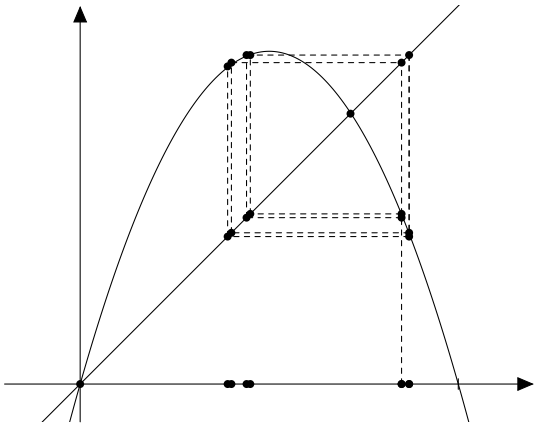
Exemple

Illustrons deux systèmes dynamiques discrets pour des fonctions appelées logistiques sur l'intervalle $I=[0,1]$. Dans tous les cas, on représente les valeurs de la suite en abscisse et on obtient graphiquement le point suivant en passant par la première bissectrice.

- Soit $f : x\mapsto 2x(1-x)$. Cette fonction admet deux points fixes sur l'intervalle I d'abscisses 0 et $\frac{1}{2}$. On observe que les systèmes issus de conditions initiales non nulles semblent converger vers $\frac{1}{2}$.



- Soit $f : x \mapsto 3{,}52x(1-x)$. Il y a encore deux points fixes mais le système dynamique tracé semble avoir un comportement asymptotique « périodique » ou, pour le dire mieux, il semble y avoir deux suites extraites $(u_{2n})_n$ et $(u_{2n+1})_n$ convergentes de limites différentes.



Pour préciser ces comportements, rappelons un théorème du point fixe.

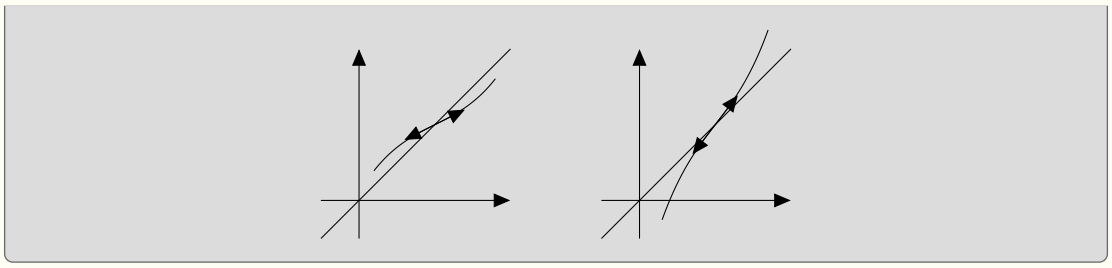
Proposition 8.22. Théorème du point fixe

Soit $f : I \to I$ k-lipschitzienne avec $k < 1$ sur un segment I. Alors, f admet un unique point fixe ℓ et, pour tout $u_0 \in I$, le système dynamique discret $(u_n)_n$ de dynamique f et de condition initiale u_0 converge vers ℓ.
Plus précisément, la convergence est « exponentielle » au sens suivant

$$\forall n \in \mathbb{N}, \qquad |u_n - \ell| \leq k^n |u_0 - \ell|.$$

Définition 8.23.

Soit $f : I \to I$ de classe $\mathscr{C}^1$ et ℓ un point fixe de f.
- Si $|f'(\ell)| < 1$, alors ℓ est attractif.
- Si $|f'(\ell)| > 1$, alors ℓ est répulsif.

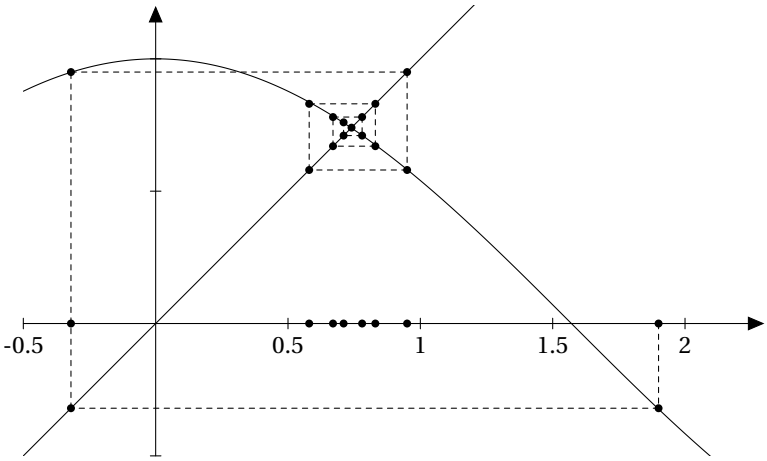


La définition d'un point fixe attractif correspond à l'intuition : en effet, si $|f'(\ell)| < 1$ et f' est continue, alors il existe un voisinage de ℓ sur lequel $|f'|$ est strictement plus petit que 1 donc sur lequel f est k-lipschitzienne avec $k < 1$ d'après la propriété des accroissements finis ; si u_0 appartient à cet intervalle, alors le système dynamique $(u_n)_n$ converge vers ℓ.

On procède de manière analogue pour le cas répulsif.

Exemple

La fonction $f : x \mapsto \cos x$ admet un unique point fixe sur $\mathbb{R}$ (facile à établir en étudiant les variations de $f - \mathrm{Id}$) qui est attractif car il n'appartient pas à l'ensemble $\frac{\pi}{2} + \pi\mathbb{Z}$ des réels où $f' : x \mapsto -\sin x$ vaut 1 ou -1.



FICHE

SYNTHÈSE

Ce chapitre est consacré à l'étude des fonctions dérivables.

Définitions

▹ Une fonction $f : I \to \mathbb{R}$ est dérivable en $a \in I$ si la fonction taux d'accroissement

$$x \mapsto \frac{f(x)-f(a)}{x-a}$$

admet une limite finie en a.
Cette limite est le nombre dérivé en a et notée $f'(a)$.
▹ Une fonction $f : I \to \mathbb{R}$ est dérivable sur I si f est dérivable en tout point de I.

Les propriétés opératoires (combinaison linéaire, produit, composition...) des limites se transmettent à la dérivation.
On se sert de la dérivation pour déterminer les variations ou rechercher des extremums.

Variations

Soit f dérivable sur un intervalle I.

- f est constante si, et seulement si, $f' = 0$ sur I.
- f est croissante si, et seulement si, $f' \geq 0$ sur I.
- f est décroissante si, et seulement si, $f' \leq 0$ sur I.

En itérant, on obtient la notion de dérivée $n^{\text{ième}}$ et de fonction de classe $\mathscr{C}^n$.

Dérivée $n^{\text{ième}}$

▹ Une fonction $f : I \to \mathbb{R}$ est de classe $\mathscr{C}^0$ si f est continue sur I.
▹ Pour tout $n \in \mathbb{N}$, une fonction $f : I \to \mathbb{R}$ est de classe $\mathscr{C}^{n+1}$ sur I si f est dérivable sur I et f' est de classe $\mathscr{C}^n$.
▹ Une fonction $f : I \to \mathbb{R}$ est de classe $\mathscr{C}^\infty$ sur I si f est de classe $\mathscr{C}^n$ sur I pour tout $n \in \mathbb{N}$.
Soit $f, g : I \to \mathbb{R}$ des fonctions n fois dérivables. Alors, la fonction produit $f.g$ est n fois dérivable et la dérivée $n^{\text{ième}}$ est donnée par la formule de Leibniz

$$(f.g)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} f^{(k)} g^{(n-k)}.$$

Le principal résultat théorique est le théorème de Rolle et ses corollaires : égalité et inégalité des accroissements finis.

On peut garder en mémoire que ce sont des conséquences directes du résultat indiquant que si une fonction dérivable admet un extremum en un point intérieur à son domaine de définition, sa dérivée s'y annule.

Théorèmes de Rolle et des accroissements finis

▹ Soit $f : [a, b] \to \mathbb{R}$ telle que
- f est continue sur $[a, b]$,
- f est dérivable sur $]a, b[$,
- $f(a) = f(b)$.

Alors, il existe $c \in]a, b[$ tel que $f'(c) = 0$.

▹ Soit $f : [a, b] \to \mathbb{R}$ telle que
- f est continue sur $[a, b]$;
- f est dérivable sur $]a, b[$.

Alors, il existe $c \in]a, b[$ tel que $f(b) - f(a) = f'(c)(b - a)$.

▹ Soit $f : [a, b] \to \mathbb{C}$ telle que
- f est continue sur $[a, b]$,
- f est dérivable sur $]a, b[$,
- $|f'|$ est bornée par M sur $]a, b[$.

Alors, $|f(b) - f(a)| \leq M|b - a|$.

Une conséquence utile est le théorème de dérivation suivant qui permet notamment d'étudier les raccordements de fonctions définies par morceaux.

Soit $f : I \to \mathbb{R}$ continue sur I, dérivable sur $I \setminus \{a\}$ telle que f' admette une limite finie en a. Alors, f est dérivable sur I et $f'(a) = \lim\limits_{a} f'$.

EXERCICES

Vrai ou faux ?

		Vrai	Faux
a)	Une fonction de classe $\mathscr{C}^1$ est dérivable.	☐	☐
b)	La fonction $x \mapsto x\lvert x\rvert$ est de classe $\mathscr{C}^1$.	☐	☐
c)	La fonction $x \mapsto \begin{cases} x^2 \sin \dfrac{1}{x} & \text{si } x \neq 0 \\ 0 & \text{sinon.} \end{cases}$ est de classe $\mathscr{C}^1$.	☐	☐
d)	La fonction $x \mapsto \begin{cases} x^3 \sin \dfrac{1}{x} & \text{si } x \neq 0 \\ 0 & \text{sinon.} \end{cases}$ est de classe $\mathscr{C}^1$.	☐	☐
e)	La fonction $x \mapsto \begin{cases} x^4 \sin \dfrac{1}{x} & \text{si } x \neq 0 \\ 0 & \text{sinon.} \end{cases}$ est de classe $\mathscr{C}^2$.	☐	☐
f)	Une fonction lipschitzienne est dérivable.	☐	☐
g)	f dérivable est paire si, et seulement si, f' est impaire.	☐	☐
h)	f dérivable est impaire si, et seulement si, f' est paire.	☐	☐
i)	Si f est à dérivée positive, alors f est croissante.	☐	☐
j)	Si f est dérivable et strictement croissante, alors f' est strictement positive.	☐	☐
k)	Soit $f : \mathbb{R} \to \mathbb{R}$ dérivable. $\lvert f\rvert$ est dérivable si, et seulement si, f ne s'annule pas.	☐	☐
l)	Si f est de classe $\mathscr{C}^n$ et admet $n+1$ zéros distincts, alors sa dérivée $n^{\text{ième}}$ s'annule au moins une fois.	☐	☐
m)	La dérivée d'un polynôme réel scindé est scindée.	☐	☐
n)	Si la dérivée de f s'annule en a, alors f réalise un extremum en a.	☐	☐
o)	Si la dérivée de f est strictement positive en a, alors f est strictement croissante sur un voisinage de a.	☐	☐
p)	Si f admet un maximum en a et est dérivable en a, alors $f'(a) = 0$.	☐	☐
q)	Si f admet un minimum en a, alors il existe $\eta > 0$ tel que f soit croissante sur $[a, a+\eta[$.	☐	☐

Exercices d'application du chapitre 8

●○○ **Exercice 1**

Montrer que $f : x \mapsto x^2 \tan \frac{1}{x} \sin \frac{2}{x}$ et $f(0) = 0$ est dérivable en 0.

●○○ **Exercice 2**

Soit $f : \mathbb{R} \to \mathbb{R}$ dérivable et $a \in \mathbb{R}$. Montrer que la fonction

$$x \mapsto \frac{xf(a) - af(x)}{x - a}$$

admet une limite en a que l'on précisera.

●○○ **Exercice 3**

Soit f une fonction de classe $\mathscr{C}^1$ et bornée sur $\mathbb{R}$. Supposons que f' possède une limite finie ℓ en $+\infty$. Déterminer ℓ.

●○○ **Exercice 4**

Soit $(a, b) \in \mathbb{R}^2$ et f la fonction définie par

$$f(x) = (x - a)^n (x - b)^n.$$

1. Calculer $f^{(n)}(x)$.
2. Trouver, à l'aide du cas $a = b$, l'expression de $\sum_{k=0}^{n} \binom{n}{k}^2$.

●●○ **Exercice 5**

Soit $f, g : [a, b] \to \mathbb{R}$ continues sur $[a, b]$, dérivables sur $]a, b[$. Montrer qu'il existe $c \in]a, b[$ tel que

$$g'(c)(f(b) - f(a)) = f'(c)(g(b) - g(a)).$$

●●● **Exercice 6**

1. Soit $f, g : [a, b] \to \mathbb{R}$ deux fonctions continues sur $[a, b]$, de classe $\mathscr{C}^1$ sur $]a, b[$. On suppose que $g' > 0$ et que $\frac{f'}{g'}$ est croissante. Montrer que $x \mapsto \frac{f(x)-f(a)}{g(x)-g(a)}$ et $x \mapsto \frac{f(x)-f(b)}{g(x)-g(b)}$ sont croissantes.
   Indication : on pourra utiliser l'exercice précédent.
2. Étudier la monotonie de $\varphi : x \in [1, +\infty[\mapsto \frac{x^p - 1}{x^q - 1}$ pour $p > q > 0$.

●○○ **Exercice 7 Théorème de Rolle généralisé**

Soit f dérivable sur $\mathbb{R}$ telle que

$$\lim_{+\infty} f = \lim_{-\infty} f \in \mathbb{R}.$$

Montrer qu'il existe $c \in \mathbb{R}$ tel que $f'(c) = 0$.

●○○ **Exercice 8**

Soit P un polynôme à coefficients réels. Montrer que l'équation $e^x = P(x)$ admet au plus $\deg P + 1$ solutions.

●●○ **Exercice 9**

Soit $P_1,\dots,P_n,Q_1,\dots,Q_n \in \mathbb{R}[X]$ avec $P_1,\dots,P_n$ non nuls et $Q_1,\dots,Q_n$ deux à deux distincts. Montrer que la fonction

$$f_n : x \mapsto \sum_{k=1}^{n} P_k(x)e^{Q_k(x)}$$

admet un nombre fini de zéros.

●●○ **Exercice 10**

Soit $n \geq 3$, $\alpha, \beta \in \mathbb{R}$. Montrer que la fonction $f : x \mapsto x^n + \alpha x + \beta$ s'annule au plus trois fois.

●●○ **Exercice 11**

Soit $f : [a,b] \to \mathbb{R}$ de classe $\mathscr{C}^n$ telle que $f(a) = 0$ et

$$f(b) = f'(b) = \dots = f^{(n-1)}(b) = 0.$$

Montrer qu'il existe $c \in]a,b[$ tel que $f^{(n)}(c) = 0$.

●●○ **Exercice 12**

Soit $f : [0,a] \to \mathbb{R}$ continue, dérivable sur $]0,a]$, vérifiant $f(0) = 0$ et $f(a)f'(a) < 0$. Montrer qu'il existe $c \in]0,a[$ tel que $f'(c) = 0$.

●●○ **Exercice 13**

Soit $f : [a,b] \to \mathbb{R}$ dérivable telle que $f(a) = f(b) = 0$. Montrer que, par tout point $(x,0)$ avec $x \notin [a,b]$, il passe au moins une tangente à la courbe représentative de f.

●●● **Exercice 14**

Soit $f : [a,b] \to \mathbb{R}$ dérivable telle que $f(a) = f(b) = 0$ et $f'(a)f'(b) > 0$. Montrer qu'il existe $a < c_1 < c_2 < c_3 < b$ tels que $f'(c_1) = f(c_2) = f'(c_3) = 0$.

●●○ **Exercice 15**

Soit f une fonction de classe $\mathscr{C}^n$ sur $[a,b]$, $a_1 < \dots < a_n$ des éléments de $[a,b]$. Le polynôme interpolateur P de f aux points $(a_i)_{i\leq n}$ est défini par

$$P(x) = \sum_{k=1}^{n} f(a_k) \prod_{j\neq k} \frac{x-a_j}{a_k-a_j}.$$

1. Calculer, pour tout i, la valeur $P(a_i)$.
   Fixons $x \in [a,b]$ différent des points $(a_i)_{i\leq n}$.
2. Montrer qu'il existe une constante $A \in \mathbb{R}$ telle que la fonction

$$g : t \mapsto f(t) - P(t) - A\frac{(t-a_1)\dots(t-a_n)}{n!}$$

   s'annule en x.
3. En déduire qu'il existe $\xi \in]a,b[$ tel que

$$f(x) - P(x) = \frac{(x-a_1)\dots(x-a_n)}{n!} f^{(n)}(\xi).$$

EXERCICES

Exercices d'approfondissement du chapitre 8

●○○ **Exercice A**

Soit $f, g : I \to \mathbb{R}$ dérivables en $a \in I$ et telles que $f(a) = g(a)$. Supposons qu'il n'existe pas de voisinage de a sur lequel $f \geq g$ ou $g \geq f$. Montrer que si $h : x \mapsto \max(f(x), g(x))$ est dérivable en a, alors $f'(a) = g'(a)$.

●●○ **Exercice B**

1. Montrer que, pour tout $x \in \mathbb{R}$, il existe un unique $\varphi(x)$ tel que
$$\int_x^{\varphi(x)} e^{t^2}\,\mathrm{d}t = 1.$$
2. Montrer que la fonction φ est de classe $\mathscr{C}^1$.

●●○ **Exercice C**

Déterminer les fonctions $f : \mathbb{R} \to \mathbb{R}$ tels que, pour tous x et $y \in \mathbb{R}$, $f(x) - f(y) \leq (x-y)^2$.

●●● **Exercice D**

Soit $f \in \mathscr{C}^0(\mathbb{R}, \mathbb{R})$ dérivable en $a \in \mathbb{R}$, $(x_n)_n$ et $(y_n)_n$ deux suites de limite a.
On suppose que, pour tout $n \in \mathbb{N}$, $y_n \neq x_n$ et on pose, pour tout $n \in \mathbb{N}$,
$$u_n = \frac{f(y_n) - f(x_n)}{y_n - x_n}.$$

1. On suppose dans cette question que $y_n > a > x_n$ pour tout $n \in \mathbb{N}$. Déterminer la limite de $(u_n)_n$.
2. On suppose dans cette question f de classe $\mathscr{C}^1$ sur un voisinage de a. Que dire de la suite $(u_n)_n$?
3. Examiner le cas général.

●●● **Exercice E**

Soit $f : \mathbb{R} \to \mathbb{R}$ deux fois dérivable telle qu'il existe $c \in \mathbb{R}$ vérifiant
$$\forall a \neq b, \qquad f'(c) \neq \frac{f(a) - f(b)}{a - b}.$$
Montrer que $f''(c) = 0$.

●●● **Exercice F**

Soit $f :]0, 1] \to \mathbb{R}$ dérivable et prolongeable par continuité en 0. Supposons que $x \mapsto x f'(x)$ admette une limite en 0. Calculer cette limite.

Problèmes du chapitre 8

●●○ **Problème 1 Fonctions à variations bornées**

Dans ce problème, on s'intéresse à une classe de fonctions réelles. Dans les deux premières parties, on considère deux propriétés dont l'équivalence est établie en partie C. Les parties D et E explorent le lien avec la propriété de dérivabilité.

Dans tout ce sujet, I désigne un intervalle de $\mathbb{R}$.

Partie A – Fonctions BV

Définition 8.24.

Une fonction $f : I \to \mathbb{R}$ est BV si elle se décompose comme somme d'une fonction croissante et d'une fonction décroissante.

1. Montrer qu'une fonction monotone est BV.
2. Montrer qu'une somme de fonctions BV est BV.
   Soit $a, b \in I$ tels que $a < b$.
3. Soit $f : I \to \mathbb{R}$ une fonction BV, somme de la fonction $g : I \to \mathbb{R}$ croissante et de la fonction $h : I \to \mathbb{R}$ décroissante. Montrer que
$$g(b) - g(a) \geq f(b) - f(a) \geq h(b) - h(a).$$
4. Montrer qu'une fonction $f : I \to \mathbb{R}$ BV est bornée sur le segment $[a, b]$.
5. a. Démontrer qu'une fonction croissante admet des limites finies à gauche et à droite en tout point intérieur à son intervalle de définition.
   b. En déduire qu'une fonction $f : I \to \mathbb{R}$ BV admet des limites finies à gauche et à droite en tout point intérieur à I.
6. Soit $f : \mathbb{R} \to \mathbb{R}$ T-périodique telle qu'il existe $g : [0, T] \to \mathbb{R}$ croissante et $h : [0, T] \to \mathbb{R}$ décroissante vérifiant
$$\forall t \in [0, T], \qquad f(t) = g(t) + h(t).$$
   Montrer que f est BV.

Partie B – Variations d'une fonction

Définition 8.25.

Soit $f : I \to \mathbb{R}$, $a, b \in I$ avec $a < b$.
▹ La variation de f entre $a < b$ est la quantité $V_a^b(f) \in \mathbb{R}_+ \cup \{+\infty\}$ définie comme la borne supérieure des sommes de la forme
$$\sum_{k=1}^{N} |f(x_k) - f(x_{k-1})|$$
avec $N \in \mathbb{N}$ et $x_0 = a < x_1 < \ldots < x_N = b$.
La fonction f est à variations bornées sur le segment $[a, b]$ si $V_a^b(f) < +\infty$.
▹ Par convention, $V_b^a(f) = -V_a^b(f)$ et $V_a^a(f) = 0$.

1. Soit $g : I \to \mathbb{R}$, $h : I \to \mathbb{R}$, $a, b \in I$ avec $a < b$. Supposons que g et h sont à variations bornées sur $[a, b]$. Montrer que $g + h$ est à variations bornées sur $[a, b]$.
2. Soit $f : I \to \mathbb{R}$, $a, b \in I$ avec $a < b$. Justifier que $|f(b) - f(a)| \leq V_a^b(f)$.
3. Soit $f : I \to \mathbb{R}$ à variations bornées sur tout segment inclus dans I, $a, b, c \in I$ avec $a < c < b$. Montrer que $V_a^b(f) = V_a^c(f) + V_c^b(f)$.
   Indication : on pourra séparer les deux inégalités.
4. Soit $f : I \to \mathbb{R}$ à variations bornées sur tout segment inclus dans I, $a, b, c \in I$. Montrer que $V_a^b(f) = V_a^c(f) + V_c^b(f)$.

Partie C – Caractérisation

1. Montrer qu'une fonction $g : I \to \mathbb{R}$ croissante est à variations bornées sur tout segment $[a, b] \subset I$. Préciser la valeur $V_a^b(g)$.
2. En déduire qu'une fonction $f : I \to \mathbb{R}$ BV est à variations bornées sur tout segment inclus dans I.

PROBLÈMES

3. Soit $f : I \to \mathbb{R}$ à variations bornées sur tout segment inclus dans I et $a \in I$. Posons $g : x \mapsto \frac{1}{2}(f(x) + V_a^x(f))$.
   a. Montrer que $g : I \to \mathbb{R}$ est croissante.
   b. En déduire que f est BV.

Partie D – Fonctions de classe $\mathscr{C}^1$

Soit $f : I \to \mathbb{R}$ de classe $\mathscr{C}^1$, a, $b \in I$ avec $a < b$.

1. Soit x, $y \in I$ tels que $x < y$. Montrer que
$$|f(y) - f(x)| \leq \int_x^y |f'(t)| \, \mathrm{d}t \, .$$
2. Montrer que f est à variations bornées sur $[a, b]$ et que
$$V_a^b(f) \leq \int_a^b |f'(t)| \, \mathrm{d}t \, .$$
3. Soit $\varepsilon > 0$.
   a. Montrer qu'il existe une subdivision $x_0 = a < x_1 < \ldots < x_N = b$ telle que
$$\forall k \in [\![1, N]\!], \ \forall x, y \in [x_{k-1}, x_k], \qquad |f'(x) - f'(y)| \leq \varepsilon.$$
   b. Justifier que, pour tout $k \in [\![1, N]\!]$, il existe $c_k \in]x_{k-1}, x_k[$ tel que
$$|f'(c_k)|(x_k - x_{k-1}) = |f(x_k) - f(x_{k-1})|.$$
   c. En déduire que
$$V_a^b(f) \geq \int_a^b |f'(t)| \, \mathrm{d}t - \varepsilon(b - a).$$
4. En déduire l'expression de $V_a^b(f)$.

Partie E – Un contre-exemple dérivable

Considérons la fonction $f : \mathbb{R} \to \mathbb{R}$ définie par
$$f : x \mapsto \begin{cases} x^2 \sin \frac{1}{x^2} & \text{si } x \neq 0, \\ 0 & \text{sinon.} \end{cases}$$

1. Justifier successivement que f est continue sur $\mathbb{R}$ puis que f est dérivable sur $\mathbb{R}$.
2. Étudier la continuité de f'.
3. Montrer que la fonction f n'est pas à variations bornées sur le segment $[-1, 1]$.
4. Déterminer tous les segments de $\mathbb{R}$ sur lesquels f est à variations bornées.

●●● Problème 2 Théorème de Denjoy

Soit $f : [a, b] \to \mathbb{C}$ continue et $g : [a, b] \to \mathbb{R}$ continue croissante. Supposons qu'il existe une suite croissante $(x_n)_{n \in I}$ (finie ou infinie) à valeurs dans $[a, b]$ telle que, pour tout $x \in [a, b] \setminus \{x_n, \ n \in I\}$,

- f admet une dérivée à droite en x, notée $f'_d(x) \in \mathbb{C}$,
- g admet une dérivée à droite en x, notée $g'_d(x) \in \mathbb{R}_+$,
- $|f'_d(x)| \leq g'_d(x)$.

Soit $\varepsilon > 0$. Notons J_ε l'ensemble des $y \in [a, b]$ tel que, pour tout $x \in [a, y]$,
$$|f(x) - f(a)| \leq g(x) - g(a) + \varepsilon(x - a) + \varepsilon \sum_{\substack{k \in I \\ x_k < x}} \frac{1}{2^k}.$$

1. Montrer que J_ε est un intervalle.

2. Vérifier que $a \in J_\varepsilon$.
3. Soit $c = \sup J_\varepsilon$. Montrer que $c \in J_\varepsilon$.
4. Supposons $c < b$.
   a. Supposons qu'il existe $n \in I$ tel que $c = x_n$. Montrer qu'il existe un voisinage à droite de c sur lequel
   $$|f(x) - f(c)| \leq \varepsilon \frac{1}{2^n}.$$
   En déduire une contradiction avec la définition de c.
   b. Supposons qu'il n'existe pas de $n \in I$ tel que $c = x_n$. Montrer qu'il existe un voisinage à droite de c sur lequel
   $$|f(x) - f(c)| \leq g(x) - g(c) + \varepsilon(x - c).$$
   En déduire une contradiction avec la définition de c.
5. Déduire des questions précédentes que
$$|f(b) - f(a)| \leq g(b) - g(a).$$

En guise d'application, remarquons qu'une fonction à valeurs complexes, admettant une dérivée à droite nulle (sauf sur un ensemble au plus dénombrable) est constante.

PROBLÈMES

●●○ **Problème 3 Fonctions de Takagi-Landsberg**

Partie A – Distance à $\mathbb{Z}$

Considérons la fonction $g : x \mapsto \left|x - \left\lfloor x + \frac{1}{2} \right\rfloor\right|$.

1. Montrer que g est 1-périodique, paire et continue sur $\mathbb{R}$.
2. Définissons, pour tout $n \in \mathbb{N}$, $g_n : x \mapsto \frac{g(2^n x)}{2^n}$.
   a. Quelle transformation géométrique simple envoie la courbe représentative de g sur celle de g_n ?
   b. Tracer les courbes représentatives de g, g_1 et g_2 sur le même graphique que celui de g.
3. Montrer que les fonctions g, g_n sont 1-lipschitziennes. Montrer que pour tout $k \in \mathbb{Z}$, g_n est affine sur l'intervalle $\left[\frac{k}{2^{n+1}}, \frac{k+1}{2^{n+1}}\right]$. Quelles valeurs peut prendre la pente ?

Partie B – Fonction de Takagi

Posons, pour tout $n \in \mathbb{N}$,
$$f_n : x \mapsto \sum_{k=0}^{n} g_k(x) = \sum_{k=0}^{n} \frac{g(2^k x)}{2^k}.$$

1. a. Soit $x \in \mathbb{R}$. Montrer que la suite $(f_n(x))_{n \geq 0}$ est convergente. On note $f(x)$ sa limite.
   b. Montrer que f est 1-périodique et paire.
   c. Montrer que, pour tout $x \in \mathbb{R}$, $|f(x) - f_n(x)| \leq \frac{1}{2^n}$.
   En déduire que, pour tous x, y tels que $|x - y| \leq \frac{1}{2^n}$, $|f(x) - f(y)| \leq \frac{n+3}{2^n}$.
   d. Conclure que f est uniformément continue sur $\mathbb{R}$.
2. Dans cette question, on fixe x et on souhaite montrer que f n'est pas dérivable en x. On définit deux suites $(a_n)_n$ et $(b_n)_n$ par :
$$a_n = \frac{\lfloor 2^{n+1} x \rfloor}{2^{n+1}}, \qquad b_n = \frac{\lfloor 2^{n+1} x \rfloor + 1}{2^{n+1}}.$$
On admettra que $\lfloor x \rfloor + \lfloor y \rfloor \leq \lfloor x + y \rfloor \leq \lfloor x \rfloor + \lfloor y \rfloor + 1$ pour tous x, $y \in \mathbb{R}$.
   a. Montrer que $(a_n)_n$ et $(b_n)_n$ sont des suites adjacentes dont on précisera la limite commune.
   b. Vérifier que $f(a_n) = f_n(a_n)$ et $f(b_n) = f_n(b_n)$, pour tout $n \in \mathbb{N}$.
   c. Justifier que
$$\frac{f_n(b_{n+1}) - f_n(a_{n+1})}{b_{n+1} - a_{n+1}} = \frac{f_n(b_n) - f_n(a_n)}{b_n - a_n}.$$

d. Posons $\Delta_n = \frac{f(b_n)-f(a_n)}{b_n-a_n}$. Vérifier que $|\Delta_{n+1}-\Delta_n| = 1$.

e. Montrer que si f est dérivable en x, alors $(\Delta_n)_n$ converge vers $f'(x)$. Conclure.

3. a. Soit $x = \frac{k}{2^n}$; posons $x_p = x + \frac{1}{2^p}$. Montrer que pour tout $p \geq 2n$, alors $f(x_p) > f(x)$.

b. En déduire que f n'est monotone sur aucun intervalle.

Partie C – Fonction de Takagi-Landsberg de paramètre $\frac{1}{4}$

Posons, pour tout $n \in \mathbb{N}$, $h_n : x \longmapsto \sum_{k=0}^{n} \frac{g(2^k x)}{4^k}$. On admet que la suite $(h_n(x))_{n\geq 0}$ est convergente pour tout x et que sa limite définit une fonction continue h.

1. Justifier rapidement que $h(0) = 0$, h est 1-périodique et que
$$\forall x \in [0, \tfrac{1}{2}], \quad h(2x) = 4(h(x) - x).$$

2. Définissons une fonction 1-périodique p par $p(x) = 2x(1-x)$ pour $x \in [0,1]$.
Montrer que
$$\forall x \in [0, \tfrac{1}{2}], \quad p(2x) = 4(p(x) - x).$$

3. Montrer que $h = p$ (et donc que h est dérivable sur $]0,1[$).
Indication : on pourra commencer par montrer que $h - p$ est nulle sur l'ensemble des nombres dyadiques $\{\frac{k}{2^n},\ k \in \mathbb{Z},\ n \in \mathbb{N}\}$.

CORRIGÉS

Corrigés des Vrai/Faux

a) Vrai par définition.

b) Vrai, la dérivée en 0 vaut 0.

c) Faux, la dérivée en 0 vaut 0 mais la dérivée n'admet pas de limite en 0.

d) Vrai, les deux premières dérivées en 0 valent 0.

e) Faux.

f) Faux, par exemple $x \mapsto |x|$ n'est pas dérivable mais est 1-lipschitzienne d'après l'inégalité triangulaire.

g) Vrai.

h) Faux, par exemple, la fonction $x \mapsto 1 - \frac{1}{x}$ admet une dérivée paire.

i) Vrai sur un intervalle.

j) Faux, par exemple $x \mapsto x^3$.

k) Faux, par exemple $x \mapsto x^2$.

l) Vrai en itérant le théorème de Rolle.

m) Vrai.

n) Faux, par exemple $f : x \mapsto x^3$ et $a = 0$.

o) Faux.

p) Faux, par exemple si a est un bord de l'intervalle de définition de f.

q) Faux, par exemple $x \mapsto |x \sin \frac{1}{x}|$ prolongée par 0 en 0.

Corrigés des exercices d'application du chapitre 8

Exercice 1

Remarquons que $f(x) = 2x^2\left(\sin \frac{1}{x}\right)^2$ donc $\lim\limits_{x\to 0} f(x) = 0$ et $\lim\limits_{x\to 0} \frac{f(x)}{x} = \lim\limits_{x\to 0} 2x\left(\sin \frac{1}{x}\right)^2 = 0$. Ainsi, f est continue et dérivable en 0.

Exercice 2

Pour tout $x \neq a$,

$$\begin{aligned}\frac{xf(a)-af(x)}{x-a} &= \frac{(x-a+a)f(a)-af(x)}{x-a}\\ &= f(a)-a\frac{f(x)-f(a)}{x-a}.\end{aligned}$$

Ainsi, avec l'hypothèse de dérivabilité de f en a, la limite recherchée est $f(a)-af'(a)$.

Exercice 3

Supposons $\ell > 0$ (si $\ell < 0$, on se ramène à ce cas en considérant l'énoncé avec $-f$ à la place de f). Par définition de la limite de f' en $+\infty$, il existe $A > 0$ tel que

$$\forall x \geq A, \qquad f'(x) \geq \frac{\ell}{2}.$$

Par conséquent, pour tout $x \geq A$,

$$f(x) = f(A) + \int_A^x f'(t)\,\mathrm{d}t \geq f(A) + \frac{\ell}{2}(x-A).$$

Ainsi, $\lim\limits_{+\infty} f = +\infty$: contradiction.

Comme on ne peut avoir ni $\ell > 0$, ni $\ell < 0$, on a $\ell = 0$.

Remarque

On notera qu'il est possible d'avoir une fonction de classe $\mathscr{C}^1$ et convergente en $+\infty$ sans que la dérivée admette une limite en $+\infty$. Par exemple, la fonction définie sur $\mathbb{R}_+^*$ par

$$x \mapsto \frac{1}{x}\cos(x^2)$$

vérifie ces hypothèses.

Exercice 4

1. Notons $g : x \mapsto (x-a)^n$ et $h : x \mapsto (x-b)^n$ et appliquons la formule de Leibniz. Pour tout $x \in \mathbb{R}$,

$$\begin{aligned}f^{(n)}(x) &= \sum_{k=0}^{n}\binom{n}{k} g^{(k)}(x)h^{(n-k)}(x)\\ &= \sum_{k=0}^{n}\binom{n}{k}\frac{n!}{(n-k)!}(x-a)^{n-k}\frac{n!}{k!}(x-b)^k\\ &= n!\sum_{k=0}^{n}\binom{n}{k}^2(x-a)^{n-k}(x-b)^k.\end{aligned}$$

2. En choisissant $a = b$, la première question donne une autre expression de $f^{(n)}(x) = \frac{(2n)!}{n!}(x-a)^n$:

$$n!\sum_{k=0}^{n}\binom{n}{k}^2(x-a)^n.$$

En divisant par $n!(x-a)^n$, on obtient

$$\sum_{k=0}^{n}\binom{n}{k}^2 = \binom{2n}{n}.$$

Exercice 5

En appliquant le théorème de Rolle à la fonction

$$x \mapsto (g(x)-g(a))(f(b)-f(a))-(f(x)-f(a))(g(b)-g(a))$$

nulle en a et b, on obtient l'existence d'un réel $c \in]a,b[$ tel que

$$g'(c)(f(b)-f(a))-f'(c)(g(b)-g(a))=0.$$

Exercice 6

1. Comme l'indication le propose, utilisons l'exercice précédent : il existe $c \in]a,x[$ tel que

$$\frac{f(x)-f(a)}{g(x)-g(a)}=\frac{f'(c)}{g'(c)}.$$

Ainsi, pour tout $x \in]a,b[$, il existe $c \in]a,x[$ tel que

$$\frac{f(x)-f(a)}{g(x)-g(a)}=\frac{f'(c)}{g'(c)} \leq \frac{f'(x)}{g'(x)}.$$

Par conséquent,

$$\left(\frac{f(x)-f(a)}{g(x)-g(a)}\right)'=\frac{f'(x)(g(x)-g(a))-g'(x)(f(x)-f(a))}{(g(x)-g(a))^2} \geq 0$$

et $x \mapsto \frac{f(x)-f(a)}{g(x)-g(a)}$ est croissante.

2. Les fonctions $f : x \mapsto x^p-1$ et $g : x \mapsto x^q-1$ vérifient les hypothèses de la question précédente puisque de classe $\mathscr{C}^1$, $g'(x)=q\,x^{q-1} \geq 0$ pour tout $x>0$ et

$$\frac{f'(x)}{g'(x)}=\frac{p\,x^{p-1}}{q\,x^{q-1}}=\frac{p}{q}x^{p-q}$$

défini une fonction croissante.
En conclusion, φ est croissante.

Exercice 7

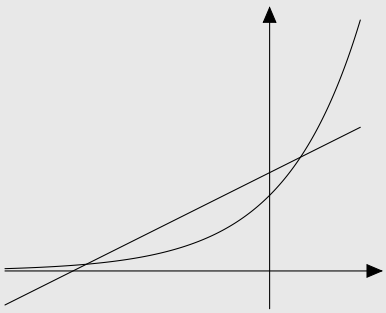
L'idée de la démonstration est très proche de celle du théorème de Rolle établi dans le cours : il suffit de montrer qu'un extremum est atteint par la fonction f.

Soit ℓ la limite de f en $+\infty$. Si f est constante égale à ℓ, le résultat est immédiat. Sinon, il existe x_0 tel que $f(x_0) \neq \ell$. Considérons y strictement entre ℓ et $f(x_0)$. D'après le théorème des valeurs intermédiaires sur $]-\infty, x_0[$ et sur $]x_0,+\infty[$, il existe deux antécédents $a<b$ à y par f. Pour conclure, il suffit ensuite d'appliquer le théorème de Rolle à f sur $[a,b]$.

Exercice 8

On retrouve dans cet exercice un résultat qui semble très intuitif pour les intersections de la courbe exponentielle avec une droite affine (c'est le cas d'un polynôme de degré au plus 1*).*

CORRIGÉS



Supposons, par l'absurde, que l'équation admette $\deg P+2$ solutions, c'est-à-dire que la fonction $f : x \mapsto e^x - P(x)$ s'annule $\deg P+2$ fois. En appliquant de manière itérée le théorème de Rolle à cette fonction de classe $\mathscr{C}^\infty$, on obtient que $f^{(\deg P+1)} : x \mapsto e^x$ s'annule : contradiction.

Exercice 9

Procédons par récurrence sur $n \in \mathbb{N} \setminus \{0\}$.

- Le résultat est évident pour f_1 car P_1 admet un nombre fini de zéros.
- Soit $n \in \mathbb{N} \setminus \{0\}$. Alors, avec des notations évidentes,

$$f_{n+1}(x) = e^{Q_{n+1}(x)}\left(P_{n+1}(x) + \sum_{k=1}^{n} P_k(x) e^{Q_k(x)-Q_{n+1}(x)}\right).$$

Si f_{n+1} admet une infinité de zéros, alors $x \mapsto P_{n+1}(x) + \sum_{k=1}^{n} P_k(x) e^{Q_k(x)-Q_{n+1}(x)}$ aussi. Quitte à appliquer le théorème de Rolle suffisamment, la dérivée d'ordre $1+\deg P_{n+1}$ de cette fonction admet une infinité de zéros : contradiction d'après l'hypothèse de récurrence.

Exercice 10

Raisonnons par l'absurde en supposant que f s'annule au moins quatre fois. En raisonnant sur chacun des trois intervalles délimités par deux de ces zéros consécutifs, on obtient par le théorème de Rolle, que f' s'annule trois fois. En recommençant sur les intervalles délimités par les zéros de f', on obtient que f'' s'annule deux fois. Or, $f'' : x \mapsto n(n-1)x^{n-2}$ admet un seul zéro : contradiction.

Exercice 11

Montrons par récurrence sur $k \in [\![0, n]\!]$ qu'il existe $x \in [a, b[$ tel que $f^{(k)}(x) = 0$.

- Pour $k = 0$, il suffit de prendre a.
- Soit $k \in [\![0, n-1]\!]$ et $x \in [a, b[$ tel que $f^{(k)}(x) = 0$. Comme $f^{(k)}$ est dérivable sur $[x, b]$ et nulle en x et b, le théorème de Rolle entraîne l'existence d'un zéro de $f^{(k+1)}$ et donc la propriété au rang $k+1$.

La propriété au rang n correspond à la propriété de l'énoncé.

Exercice 12

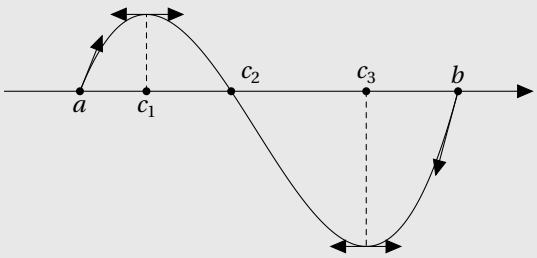
Supposons sans perte de généralité que $f(a) > 0$ et donc $f'(a) < 0$. Par définition de $f'(a)$, il existe $x_0 \in]0, a[$ tel que $f(x_0) > f(a)$ (sinon la limite $f'(a)$ serait positive). Par ailleurs, f atteint son maximum sur le segment $[0, a]$ (car continue) en un point c autre que 0 et a. En ce point la dérivée est nulle.

Exercice 13

Rappelons que l'équation de la tangente en $c \in [a,b]$ est $y = f'(c)(x-c)+f(c)$.
Fixons $x \notin [a,b]$; la tangente en c passe par le point $(x,0)$ si, et seulement si, $f'(c)(x-c)+f(c)=0$, soit encore $g'(c)=0$ pour la fonction $g : t \mapsto \frac{f(t)}{x-t}$.
Or, d'après les hypothèses de l'énoncé, g est dérivable sur $[a,b]$ et $g(a)=g(b)=0$: le théorème de Rolle assure l'existence d'un zéro de g'.

Exercice 14

Supposons sans perte de généralité que $f'(a)>0$ donc il existe $x_0 \in [a,b]$ tel que $f(x_0)>f(a)=0$. Comme $f'(b)>0$, il existe $x_1 \in [x_0,b]$ tel que $f(x_1)<f(b)=0$. D'après le théorème des valeurs intermédiaires appliqué à f sur le segment $[x_0,x_1]$, il existe $c_2 \in]x_0,x_1[$ tel que $f(c_2)=0$. Pour conclure, il suffit d'appliquer le théorème de Rolle sur $[a,c_2]$ et $[c_2,b]$.



L'introduction des réels x_0 et x_1 est nécessaire pour ce raisonnement car les hypothèses $f'(a)>0$ et $f'(b)>0$ ne garantissent pas la croissance de f aux voisinages de a et b.

Exercice 15

1. Pour tout i, dans le calcul de $P(a_i)$ tous les termes de la somme d'indice différents de i s'annulent

$$\begin{aligned} P(a_i) &= \sum_{k=1}^{n} f(a_k) \prod_{j \neq k} \frac{a_i - a_j}{a_k - a_j} \\ &= f(a_i) \prod_{j \neq i} \frac{a_i - a_j}{a_i - a_j} = f(a_i). \end{aligned}$$

2. La quantité $\frac{(x-a_1)\ldots(x-a_n)}{n!}$ est non nulle donc l'équation $f(x)-P(x)-A\frac{(x-a_1)\ldots(x-a_n)}{n!}=0$ admet bien une solution pour A.
3. Par construction, g s'annule $n+1$ fois en $a_1,\ldots,a_n$ d'après la première question et en x d'après la deuxième. En itérant le théorème de Rolle, $g^{(n)}$ s'annule en un $\xi \in [a,b]$. Par ailleurs, P est de degré au plus $n-1$ donc $P^{(n)}=0$. Ainsi

$$0 = g^{(n)}(\xi) = f^{(n)}(\xi) - A,$$

c'est-à-dire $A = f^{(n)}(\xi)$.
En reportant dans la condition $g(x)=0$, on obtient

$$f(x)-P(x) = \frac{(x-a_1)\ldots(x-a_n)}{n!} f^{(n)}(\xi).$$

Corrigés des exercices d'approfondissement du chapitre 8

Exercice A

Par hypothèse, il existe des suites $(x_n)_n$ et $(y_n)_n$ à valeurs dans I et de limite a telles que, pour tout n, $f(x_n) > g(x_n)$ et $f(y_n) < g(y_n)$. Entre a et ces points le taux d'accroissement de h s'exprime comme le taux d'accroissement de f ou de g. Plus précisément, pour tout n,

$$\frac{h(x_n)-h(a)}{x_n-a} = \frac{f(x_n)-f(a)}{x_n-a} \to f'(a)$$
$$\frac{h(y_n)-h(a)}{y_n-a} = \frac{g(y_n)-g(a)}{x_n-a} \to g'(a)$$

Or, la fonction h est dérivable en a donc ces limites sont $h'(a)$. Ainsi, $f'(a) = h'(a) = g'(a)$.

Exercice B

1. Soit F la primitive de $t \mapsto e^{t^2}$ nulle en 0. Comme F est dérivable à dérivée strictement positive, F est une bijection de classe $\mathscr{C}^1$ de $\mathbb{R}$ dans $F(\mathbb{R})$ de réciproque de classe $\mathscr{C}^1$.
Déterminons $F(\mathbb{R})$. Pour tout $x \geq 0$,

$$F(x) = \int_0^x e^{t^2}\,\mathrm{d}t \geq \int_0^x 1\,\mathrm{d}t = x,$$

d'où $\lim\limits_{+\infty} F = +\infty$. De même, pour $x \leq 0$,

$$F(x) = \int_0^x e^{t^2}\,\mathrm{d}t = -\int_{-x}^0 e^{t^2}\,\mathrm{d}t \leq -\int_0^{-x} 1\,\mathrm{d}t = x,$$

d'où $\lim\limits_{-\infty} F = -\infty$. Comme F est continue, $F(\mathbb{R}) = \mathbb{R}$.
Ainsi, pour tout $x \in \mathbb{R}$, $1 + F(x) \in \mathbb{R}$ admet un unique antécédent par F que l'on note $\varphi(x)$.
2. Avec les notations de la question précédente, $\varphi : x \mapsto F^{-1}(1 + F(x))$. Comme F et F^{-1} sont de classe $\mathscr{C}^1$, φ est, par composition, de classe $\mathscr{C}^1$.

Exercice C

Commençons par remarquer que pour tous x et $y \in \mathbb{R}$, $f(x) - f(y) \leq (x-y)^2$ et $f(y) - f(x) \leq (y-x)^2$ d'où

$$|f(y) - f(x)| \leq (y-x)^2,$$

et par conséquent, si $y \neq x$,

$$\left|\frac{f(y)-f(x)}{y-x}\right| \leq |y-x|.$$

En faisant tendre y vers x, on en déduit que f est dérivable en x et que $f'(x) = 0$. Ainsi, f est dérivable sur $\mathbb{R}$ à dérivée nulle donc constante.
Réciproquement, les fonctions constantes sont solutions.

Exercice D

1. Pour tout $n \in \mathbb{N}$,
$$u_n = \frac{f(y_n)-f(a)}{y_n-a}\,\frac{y_n-a}{y_n-x_n} + \frac{f(a)-f(x_n)}{a-x_n}\,\frac{a-x_n}{y_n-x_n}.$$
Or, $\frac{y_n-a}{y_n-x_n} > 0$ et $\frac{a-x_n}{y_n-x_n} > 0$ et
$$\frac{y_n-a}{y_n-x_n} + \frac{a-x_n}{y_n-x_n} = 1.$$
Ainsi, le réel u_n est une combinaison linéaire convexe de $\frac{f(y_n)-f(a)}{y_n-a}$ et de $\frac{f(a)-f(x_n)}{a-x_n}$ donc se situe entre ces deux réels. En passant à la limite, on obtient par encadrement que $\lim u_n = f'(a)$.
2. Si f de classe $\mathscr{C}^1$ sur un voisinage I de a, alors les suites $(x_n)_n$ et $(y_n)_n$ sont à valeurs dans I à partir d'un certain rang N. Pour tout $n \geq N$, le théorème des accroissements finis assure l'existence de z_n entre x_n et y_n tel que
$$f(y_n)-f(x_n) = f'(z_n)(y_n-x_n),$$
c'est-à-dire $u_n = f'(z_n)$. Comme, par encadrement, la suite $(z_n)_n$ converge vers a et que f' est continue sur le voisinage I, on en déduit que $\lim u_n = f'(a)$.
3. Pour illustrer les différents comportements asymptotiques possibles, considérons pour tout $\alpha > 1$ la fonction
$$f_\alpha : x \mapsto \begin{cases} x^\alpha \sin\frac{1}{x} & \text{si } x \neq 0, \\ 0 & \text{sinon.} \end{cases}$$
Les fonctions f_α sont continues sur $\mathbb{R}$, dérivables en 0 pour $\alpha > 1$ et
$$f'_\alpha(0) = \lim_{x\to 0} \frac{x^\alpha \sin\frac{1}{x} - 0}{x-0} = \lim_{x\to 0} x^{\alpha-1}\sin\frac{1}{x} = 0.$$
Posons, pour tout $n \in \mathbb{N}\setminus\{0\}$, $x_n = \frac{2}{4n\pi}$ et $y_n = \frac{2}{(4n+1)\pi}$. Les suites $(x_n)_n$ et $(y_n)_n$ convergent vers 0 et, après calculs,
$$\frac{f_\alpha(y_n)-f_\alpha(x_n)}{y_n-x_n} = -2n(4n+1)\pi\left(\frac{2}{(4n+1)\pi}\right)^\alpha.$$
Ainsi,
- si $\alpha > 2$, $\lim u_n = 0 = f'_\alpha(0)$ (dans ce cas la fonction f_α est de classe $\mathscr{C}^1$) ;
- si $\alpha = 2$, $\lim u_n = -\frac{2}{\pi} \neq f'_\alpha(0)$;
- si $\alpha \in]1,2[$, $\lim u_n = -\infty$.

Exercice E

La condition entraîne que $\varphi : x \mapsto f(x) - xf'(c)$ est injective. Comme elle est par ailleurs continue sur $\mathbb{R}$, elle est strictement monotone. Comme elle est dérivable, $\varphi' : x \mapsto f'(x) - f'(c)$ est de signe constant. Ainsi, c est un extremum de f' ce qui entraîne $f''(c) = 0$.

Exercice F

Supposons que la limite α est strictement positive. Alors, il existe $\eta > 0$ tel que pour tout $x \in]0,\eta]$ $xf'(x) \geq \frac{\alpha}{2}$. D'après la formule des accroissements finis, pour tout $x \in]0, \frac{\eta}{2}]$, il existe $c \in [x, 2x]$ tel que
$$f(2x)-f(x) = xf'(c) \geq \frac{x}{c}\,cf'(c) \geq \frac{\alpha}{2c}\,x \geq \frac{\alpha}{4}.$$
Or, $\lim\limits_{x\to 0}(f(2x)-f(x)) = 0$: contradiction.
En procédant de même pour $\alpha < 0$, on finit par obtenir $\alpha = 0$.

Corrigés des problèmes du chapitre 8

Problème 1

Partie A – Fonctions BV

1. Une fonction g croissante se décompose comme $g+0$ et $x \mapsto 0$ est décroissante donc g est BV. De même une fonction h décroissante se décompose $0+h$ donc est BV.
2. Comme une somme de fonctions croissantes (respectivement décroissantes) est croissante (respectivement décroissante), une décomposition de la somme de fonctions BV s'obtient en regroupant les composantes croissantes des différentes décompositions d'une part et les composantes décroissantes d'autre part : une somme de fonctions BV est BV.
3. Remarquons que $f(b)-f(a)=(g(b)-g(a))+(h(b)-h(a))$; dans la somme le premier terme est positif (par croissance de g), le second négatif (par décroissance de h). Ainsi,

$$g(b)-g(a) \geq f(b)-f(a) \geq h(b)-h(a).$$

4. Soit f une fonction BV, g et h des fonctions croissante et décroissante respectivement telles que $f = g+h$. Pour tout $x \in [a,b]$

$$\begin{aligned} f(x) &\leq f(a)+g(x)-g(a) \leq f(a)+g(b)-g(a) \\ f(x) &\geq f(a)+h(x)-h(a) \geq f(a)+h(b)-h(a). \end{aligned}$$

Ainsi, f est bornée sur le segment $[a,b]$.
5. a. C'est simplement le théorème de la limite monotone.
b. Le résultat ci-dessus est également vrai pour les fonctions décroissantes. Par conséquent, une fonction BV, en tant que somme de fonctions monotones, admet des limites finies à gauche et à droite en tout point intérieur à son intervalle de définition.
6. Posons $\tilde{g} : x \mapsto g(x-T\lfloor x/T\rfloor)+\lfloor x/T\rfloor(g(T)-g(0))$ et $\tilde{h} = f-\tilde{g}$.
▹ Remarquons que $\lfloor x/T\rfloor$ est constante sur un intervalle de la forme $[kT,(k+1)T[$ donc la croissance de g entraîne la croissance de $\tilde{g}$ sur ces intervalles. Par ailleurs, si $x \in [kT,(k+1)T[$ et $y \in [(k+1)T,(k+2)T[$,

$$\begin{aligned} g(y)-g(x) &= g(y-T(k+1))+(k+1)(g(T)-g(0))-g(x-Tk)-k(g(T)-g(0)) \\ &= (g(T)-g(0))-(g(x-Tk)-g(y-T(k+1))) \geq 0 \end{aligned}$$

par croissance de g. En recollant les morceaux, on obtient que $\tilde{g}$ est croissante.
▹ Comme $f_{|[0,T]} = g+h$, on obtient que

$$\tilde{h} : x \mapsto h(x-T\lfloor x/T\rfloor)+\lfloor x/T\rfloor(h(T)-h(0))$$

et on montre comme précédemment que cette fonction est décroissante.
En conclusion, f est BV.

Partie B – Variations d'une fonction

1. Pour toute subdivision $x_0 = a < x_1 < \ldots < x_N = b$, l'inégalité triangulaire entraîne

$$\begin{aligned} \sum_{k=1}^{N} |(f+g)(x_k)-(f+g)(x_{k-1})| &\leq \sum_{k=1}^{N} |f(x_k)-f(x_{k-1})| + \sum_{k=1}^{N} |g(x_k)-g(x_{k-1})| \\ &\leq V_a^b(f)+V_a^b(g). \end{aligned}$$

En conclusion, la fonction $g+h$ est à variations bornées sur $[a,b]$ (car la borne supérieure des sommes sur toutes les subdivisions est majorée par $V_a^b(f)+V_a^b(g)$).
2. Il suffit de considérer la subdivision élémentaire à deux points : $x_0 = a$, $x_1 = b$.

3. ▹ À toute subdivision $x_0 = a < x_1 < \ldots < x_N = b$, associons la subdivision

$$x_0 = a < x_1 < \ldots < x_j < c \leq x_{j+1} < \ldots < x_N = b.$$

En séparant la somme en deux

$$\begin{aligned}\sum_{k=1}^{N} |f(x_k) - f(x_{k-1})| &= \sum_{k=1}^{j} |f(x_k) - f(x_{k-1})| + |f(x_{j+1}) - f(x_j)| + \sum_{k=j+1}^{N} |f(x_k) - f(x_{k-1})| \\ &\leq \sum_{k=1}^{j} |f(x_k) - f(x_{k-1})| + |f(x_{j+1}) - f(c)| + |f(c) - f(x_j)| + \sum_{k=j+1}^{N} |f(x_k) - f(x_{k-1})| \\ &\leq V_a^c(f) + V_c^b(f).\end{aligned}$$

Ainsi, $V_a^b(f) \leq V_a^c(f) + V_c^b(f)$.
▹ Soit $\varepsilon > 0$, il existe une partition $x_0 = a < \ldots < x_N = c$ et une partition $x_N = c < \ldots < x_n = b$ telles que

$$\sum_{k=1}^{N} |f(x_k) - f(x_{k-1})| \geq V_a^c(f) - \varepsilon$$
$$\sum_{k=N}^{n} |f(x_k) - f(x_{k-1})| \geq V_c^b(f) - \varepsilon$$

Par conséquent,

$$\sum_{k=1}^{n} |f(x_k) - f(x_{k-1})| \geq V_a^c(f) + V_c^b(f) - 2\varepsilon.$$

En conclusion, $V_a^b(f) \geq V_a^c(f) + V_c^b(f) - 2\varepsilon$. Comme ce résultat est valide pour tout $\varepsilon > 0$, on obtient l'inégalité $V_a^b(f) \geq V_a^c(f) + V_c^b(f)$ donc l'égalité.
4. Il suffit de discuter selon l'ordre relatif de a, b ou c. Par exemple, si $a < b < c$, on utilise le résultat de la question précédente pour obtenir

$$V_a^c(f) = V_a^b(f) + V_b^c(f).$$

On passe alors le terme $V_b^c(f)$ de l'autre côté de l'égalité et on utilise $V_b^c(f) = -V_c^b(f)$ pour obtenir l'égalité recherchée.
On procède de même dans les autres cas.

Partie C – Caractérisation

1. Soit g croissante. Pour toute subdivision $x_0 = a < x_1 < \ldots < x_N = b$,

$$\sum_{k=1}^{N} |g(x_k) - g(x_{k-1})| = \sum_{k=1}^{N} (g(x_k) - g(x_{k-1})) = g(b) - g(a).$$

Ainsi, g est à variations bornées sur le segment $[a, b]$ et $V_a^b(g) = g(b) - g(a)$.
2. Ceci est une simple conséquence de la question précédente et de la question B1.
3. a. Soit $x < y$. D'après la question B4,

$$\begin{aligned}2g(x) &= f(x) + V_a^y(f) - V_x^y(f) \\ &\leq f(x) + V_a^y(f) - (f(x) - f(y))\end{aligned}$$

car $V_x^y(f) \geq f(x) - f(y)$ d'après la question B2. Ainsi, $2g(x) \leq 2g(y)$.
La fonction g est croissante.
b. Posons de même $h : x \mapsto \frac{1}{2}(f(x) - V_a^x(f))$. On montre que h est décroissante puis que $f = g + h$. En conclusion, f est BV.

Partie D – Fonctions de classe $\mathscr{C}^1$

1. Soit $x < y$. Alors, par inégalité sur l'intégrale,

$$|f(y)-f(x)| = \left|\int_x^y f'(t)\,\mathrm{d}t\right| \leq \int_x^y |f'(t)|\,\mathrm{d}t\,.$$

2. Pour toute subdivision $x_0 = a < x_1 < \ldots < x_N = b$,

$$\begin{aligned}\sum_{k=1}^{N} |f(x_k)-f(x_{k-1})| &\leq \int_{x_{k-1}}^{x_k} |f'(t)|\,\mathrm{d}t \\ &\leq \int_a^b |f'(t)|\,\mathrm{d}t\end{aligned}$$

en utilisant la relation de Chasles. Ainsi, toutes les sommes sur les subdivisions sont majorées par $\int_a^b |f'(t)|\,\mathrm{d}t$. Par conséquent, f est à variations bornées sur $[a,b]$ et $V_a^b(f) \leq \int_a^b |f'(t)|\,\mathrm{d}t$.

3. a. Comme f est de classe $\mathscr{C}^1$ sur I, f' est continue sur le segment $[a,b] \subset I$. D'après le théorème de Heine, f' est uniformément continue sur ce segment ; il existe $\eta > 0$ tel que

$$\forall x, y \in [a,b], \qquad |x-y| \leq \eta \Rightarrow |f'(x)-f'(y)| \leq \varepsilon.$$

Pour obtenir le résultat désiré, il suffit de considérer une subdivision $x_0 = a < x_1 < \ldots < x_N = b$ telle que

$$\forall k \in [\![1, N]\!], \qquad |x_k - x_{k-1}| \leq \eta.$$

b. Il suffit d'appliquer l'égalité des accroissements finis à f entre x_{k-1} et x_k car f est de classe $\mathscr{C}^1$ sur ce segment.

c. Remarquons tout d'abord que

$$\begin{aligned}\left|\int_{x_{k-1}}^{x_k} |f'(t)|\,\mathrm{d}t - |f'(c_k)|(x_k - x_{k-1})\right| &= \left|\int_{x_{k-1}}^{x_k} (|f'(t)| - |f'(c_k)|)\,\mathrm{d}t\right| \\ &\leq \int_{x_{k-1}}^{x_k} ||f'(t)| - |f'(c_k)||\,\mathrm{d}t \\ &\leq \int_{x_{k-1}}^{x_k} |f'(t) - f'(c_k)|\,\mathrm{d}t \\ &\leq \varepsilon(x_k - x_{k-1}).\end{aligned}$$

Ainsi,

$$|f(x_k)-f(x_{k-1})| \geq \int_{x_{k-1}}^{x_k} |f'(t)|\,\mathrm{d}t - \varepsilon(x_k - x_{k-1}).$$

soit, en additionnant,

$$V_a^b(f) \geq \sum_{k=1}^{N} |f(x_k)-f(x_{k-1})| \geq \int_a^b |f'(t)|\,\mathrm{d}t - \varepsilon(b-a).$$

4. D'après les questions 2 et 3, $V_a^b(f) = \int_a^b |f'(t)|\,\mathrm{d}t$.

Partie E – Un contre-exemple dérivable

1. ▹ La fonction f est de classe $\mathscr{C}^\infty$ sur $\mathbb{R}^*$ comme produit et composée de fonctions de classe $\mathscr{C}^\infty$.
▹ La limite de $x^2 \sin \frac{1}{x^2}$ est nulle en 0 (produit d'une fonction de limite nulle et d'une fonction bornée) donc f est continue en 0.
▹ Le taux d'accroissement de f entre 0 et x vaut $x \sin \frac{1}{x^2}$ et sa limite est nulle en 0 avec le même argument que ci-dessus. Ainsi, f est dérivable en 0.
2. On a déjà justifié que f' est continue sur $\mathbb{R}^*$. Par ailleurs, pour tout $x \neq 0$,

$$f'(x) = 2x \sin \frac{1}{x^2} - \frac{2}{x} \cos \frac{1}{x^2}.$$

On remarque que $f'(\frac{1}{\sqrt{2n\pi}}) = -2\sqrt{2n\pi} \to -\infty$ donc f' n'est pas continue en 0 (d'après la caractérisation séquentielle de limite).
3. Considérons une subdivision contenant les points suivants

$$x_{2n} = -\frac{1}{\sqrt{2n\pi}}, \qquad x_{2n+1} = -\frac{1}{\sqrt{\frac{\pi}{2} + 2n\pi}}$$

pour n entre 1 et N. Alors,

$$\sum_{k=1}^{2N} |f(x_k) - f(x_{k-1})| \geq \sum_{k=1}^{N} |f(x_{2k}) - f(x_{2k+1})| = \sum_{k=1}^{N} \frac{1}{2\pi k}.$$

La somme harmonique (à droite) diverge lorsque $N \to \infty$: l'ensemble des sommes sur les substitutions n'est donc pas majoré, c'est-à-dire la fonction f n'est pas à variations bornées sur le segment $[-1, 1]$.
4. D'après la partie D, f est à variations bornées sur tout segment ne contenant pas 0 car de classe $\mathscr{C}^1$. En revanche, sur tout segment contenant 0, on peut procéder comme à la question précédente : f n'est pas à variations bornées.

Problème 2

1. Soit $y, z \in J_\varepsilon$ tel que $y < z$. Par définition de z, l'inégalité est vérifiée pour tout $x \in [a, z]$ donc tout élément de $[y, z]$ appartient à J_ε. La partie J_ε est connexe dans $\mathbb{R}$ donc un intervalle.
2. $a \in J_\varepsilon$ puisque l'inégalité s'écrit pour $x = a$ comme $0 \leq 0$.
3. Soit $(u_n)_n$ suite de J_ε de limite c. Pour tout n,

$$\begin{aligned} |f(u_n) - f(a)| &\leq g(u_n) - g(a) + \varepsilon(u_n - a) + \varepsilon \sum_{\substack{k \in I \\ x_k < u_n}} \frac{1}{2^k} \\ &\leq g(u_n) - g(a) + \varepsilon(u_n - a) + \varepsilon \sum_{\substack{k \in I \\ x_k < c}} \frac{1}{2^k}. \end{aligned}$$

Passons à la limite en n en utilisant la continuité de f et g,

$$|f(c) - f(a)| \leq g(c) - g(a) + \varepsilon(c - a) + \varepsilon \sum_{\substack{k \in I \\ x_k < c}} \frac{1}{2^k}.$$

Ainsi, l'inégalité est vérifiée en c et sur la réunion de $[a, u_n]$ donc sur $[a, c]$: en conclusion, $c \in J_\varepsilon$.
4. a. L'inégalité est une simple conséquence de la continuité de f en c.

Par inégalité triangulaire, pour les x vérifiant l'inégalité précédente

$$\begin{aligned}|f(x)-f(a)| &\leq |f(c)-f(a)|+|f(x)-f(c)|\\ &\leq g(c)-g(a)+\varepsilon(c-a)+\varepsilon\sum_{\substack{k\in I\\ x_k<c}}\frac{1}{2^k}+\varepsilon\frac{1}{2^n}\\ &\leq g(x)-g(a)+\varepsilon(x-a)+\varepsilon\sum_{\substack{k\in I\\ x_k\leq c}}\frac{1}{2^k}\\ &\leq g(x)-g(a)+\varepsilon(x-a)+\varepsilon\sum_{\substack{k\in I\\ x_k<x}}\frac{1}{2^k}.\end{aligned}$$

On contredit la définition de c comme borne supérieure de J_ε.
b. Soit $u\in\mathbb{C}$ de module au plus 1 tel que $f_d'(c)=ug_d'(c)$. La fonction $\varphi : x\mapsto f(x)-ug(x)$ est à dérivée nulle à droite en c donc il existe $\eta>0$ tel que, pour tout $x\in]c,c+\eta[$,

$$|\varphi(x)-\varphi(c)|\leq\varepsilon|x-c|.$$

Ainsi,

$$|f(x)-f(c)|\leq|u(g(x)-g(c))|+\varepsilon|x-c|\leq g(x)-g(c)+\varepsilon(x-c).$$

Par inégalité triangulaire, pour les x vérifiant l'inégalité précédente

$$\begin{aligned}|f(x)-f(a)| &\leq |f(c)-f(a)|+|f(x)-f(c)|\\ &\leq g(c)-g(a)+\varepsilon(c-a)+\varepsilon\sum_{\substack{k\in I\\ x_k<c}}\frac{1}{2^k}+g(x)-g(c)+\varepsilon(x-c)\\ &\leq g(x)-g(a)+\varepsilon(x-a)+\varepsilon\sum_{\substack{k\in I\\ x_k<c}}\frac{1}{2^k}\\ &\leq g(x)-g(a)+\varepsilon(x-a)+\varepsilon\sum_{\substack{k\in I\\ x_k<x}}\frac{1}{2^k}.\end{aligned}$$

On contredit la définition de c comme borne supérieure de J_ε.
5. D'après la question 4, $c=b$ et donc, d'après la question 3, $b\in J_\varepsilon$. D'où

$$|f(b)-f(a)|\leq g(b)-g(a)+\varepsilon(b-a)+\varepsilon\sum_{\substack{k\in I\\ x_k<b}}\frac{1}{2^k}.$$

En faisant tendre $\varepsilon\to 0$, on obtient que

$$|f(b)-f(a)|\leq g(b)-g(a).$$

Problème 3

Partie A – Distance à $\mathbb{Z}$

1. La périodicité et la parité se déduisent des propriétés suivantes de la partie entière :

$$\lfloor x+1\rfloor=\lfloor x\rfloor+1,\qquad \lfloor -x\rfloor=\begin{cases}-\lfloor x\rfloor-1 & \text{si } x\notin\mathbb{Z}\\ -\lfloor x\rfloor & \text{sinon.}\end{cases}$$

Pour montrer que g est continue sur $\mathbb{R}$, on vérifie simplement que g est continue sur $[-\frac{1}{2},\frac{1}{2}]$ et que $g(-\frac{1}{2})=g(\frac{1}{2})=\frac{1}{2}$.

2. Un point (x,y) appartient à la courbe représentative de g_n si, et seulement si, $(2^n x, 2^n y)$ appartient à la courbe représentative de g. Par conséquent, la courbe représentative de g_n se déduit de celle de g par une homothétie de centre O et de rapport $\frac{1}{2^n}$.
Pour le graphique, il suffit de remarquer que g coïncide avec l'identité sur $[0,\frac{1}{2}]$.
3. g est clairement 1-lipschitzienne sur les intervalles de la forme $[k,k+\frac{1}{2}]$ et $[k+\frac{1}{2},k+1]$ (car affine de pente 1 ou -1). Pour tous $x<y\in\mathbb{R}$, il suffit de subdiviser l'intervalle $[x,y]$ aux points entiers ou demi-entiers $x=a_0<a_1<\ldots<a_{N-1}<a_N=y$ et d'appliquer l'inégalité triangulaire :

$$|g(x)-g(y)|\leq\sum_{k=0}^{N-1}|g(a_k)-g(a_{k+1})|\leq\sum_{k=0}^{N-1}(a_k-a_{k+1})=y-x.$$

D'où g (et de manière évidente g_n) est 1-lipschitzienne sur $\mathbb{R}$.
Pour tout $k\in\mathbb{Z}$, Si $x,y\in\left[\frac{k}{2^{n+1}},\frac{k+1}{2^{n+1}}\right]$, alors $2^n x$ et $2^n y$ sont tous les deux dans l'intervalle $[0,\frac{1}{2}]$ ou dans l'intervalle $[\frac{1}{2},1]$. Or, g est affine de pente ±1 sur ces intervalles. Donc g_n est affine sur l'intervalle $\left[\frac{k}{2^{n+1}},\frac{k+1}{2^{n+1}}\right]$ de pente ±1.

Partie B – Fonction de Takagi

1. a. La suite $(f_n(x))_{n\geq0}$ est croissante et majorée par $\sum_{k=1}^{\infty}2^{-k}=1$ (car g est bornée par $\frac{1}{2}$) donc convergente.
b. On déduit le résultat pour f_n (puis pour f en passant à la limite) des propriétés correspondantes pour g.
c. Pour tout $p\geq n$,

$$|f_p(x)-f_n(x)|=f_p(x)-f_n(x)=\sum_{k=n}^{p}\frac{1}{2^k}g(2^k x)\leq\frac{1}{2}\sum_{k=n}^{p}\frac{1}{2^k}\leq\frac{1}{2}\frac{1}{2^n}\leq\frac{1}{2^n}.$$

D'où le résultat en passant à la limite en p.
Pour tous $x,y\in\mathbb{R}$ tels que $|x-y|\leq\frac{1}{2^n}$,

$$\begin{aligned}|f(x)-f(y)|&\leq|f(x)-f_n(x)|+|f_n(x)-f_n(y)|+|f_n(y)-f(y)|\\&\leq\frac{1}{2^n}+\frac{n+1}{2^n}+\frac{1}{2^n}=\frac{n+3}{2^n},\end{aligned}$$

car $|f_n(x)-f_n(y)|\leq\sum_{k=0}^{n}|g_k(x)-g_k(y)|\leq\sum_{k=0}^{n}|x-y|\leq\frac{n+1}{2^n}$ (puisque les fonctions g_k sont 1-lipschitziennes).
d. Pour tout $\varepsilon>0$, il existe $n\in\mathbb{N}$ tel que $\frac{n+1}{2^n}<\varepsilon$ (car $\lim\frac{n+1}{2^n}=0$). En posant $\eta=\frac{1}{2^n}$, on retrouve la propriété d'uniforme continuité de f sur $\mathbb{R}$.
2. a. Pour tout $n\in\mathbb{N}$, avec la propriété d'encadrement de l'énoncé,

$$\begin{aligned}a_{n+1}-a_n&=\frac{1}{2^{n+1}}\left(\lfloor 2.2^{n+1}x\rfloor-2\lfloor 2^{n+1}x\rfloor\right)\geq0,\\b_{n+1}-b_n&=\frac{1}{2^{n+1}}\left(\lfloor 2.2^{n+1}x\rfloor-2\lfloor 2^{n+1}x\rfloor-1\right)\leq0,\\b_n-a_n&=\frac{1}{2^{n+1}}.\end{aligned}$$

D'où le caractère adjacent de ces suites $(a_n)_n$ et $(b_n)_n$. Leur limite commune est x.
b. Pour avoir $f(a_n)=f_n(a_n)$ et $f(b_n)=f_n(b_n)$, il suffit de remarquer que $g_p(a_n)=g_p(b_n)=0$ pour tout $p>n$.
c. D'après les remarques de la partie A, f_n est affine sur $[a_n,b_n]$.
Or, $[a_{n+1},b_{n+1}]\subset[a_n,b_n]$ d'où

$$\frac{f_n(b_{n+1})-f_n(a_{n+1})}{b_{n+1}-a_{n+1}}=\frac{f_n(b_n)-f_n(a_n)}{b_n-a_n}.$$

d. Pour tout $n \in \mathbb{N}$,

$$\begin{aligned}
\Delta_{n+1} - \Delta_n &= \frac{f_{n+1}(b_{n+1}) - f_{n+1}(a_{n+1})}{b_{n+1} - a_{n+1}} - \frac{f_n(b_n) - f_n(a_n)}{b_n - a_n} \\
&= \frac{f_{n+1}(b_{n+1}) - f_{n+1}(a_{n+1}) - f_n(b_{n+1}) + f_n(a_{n+1})}{b_{n+1} - a_{n+1}} \\
&= \frac{g_{n+1}(b_{n+1}) - g_{n+1}(a_{n+1})}{b_{n+1} - a_{n+1}} \\
&= g\left(\frac{2^{n+1}\lfloor 2^{n+2}x \rfloor + 2^{n+1}}{2^{n+2}}\right) - g\left(\frac{2^{n+1}\lfloor 2^{n+2}x \rfloor}{2^{n+2}}\right) \\
&= (-1)^{\lfloor 2^{n+2}x \rfloor}.
\end{aligned}$$

e. Si f est dérivable en x, alors on écrit

$$\Delta_n = \frac{b_n - x}{b_n - a_n}\frac{f(b_n) - f(x)}{b_n - x} + \frac{x - a_n}{b_n - a_n}\frac{f(x) - f(a_n)}{x - a_n},$$

donc $(\Delta_n)_n$ converge vers $f'(x)$. Or, $(\Delta_n)_n$ n'est pas convergente car elle n'est pas de Cauchy d'après la question précédente.

3. a. Pour tout $p \geq 2n$, $f(x_p) = f_p(x_p)$ et $f(x) = f_p(x)$. Or, f_p est la somme de p fonctions affines sur $[x, x_p]$ de pentes ± 1 dont $p - n$ au moins (les g_k pour $k \geq n$ car $[x, x_p] \subset [\frac{2^{n-p}k}{2^p}, \frac{2^{n-p}k+1}{2^p}]$) ont pour pente 1. Par conséquent, f_p est affine de pente positive et $f_p(x_p) > f_p(x)$.

b. Par densité des nombres dyadiques dans $\mathbb{R}$, f n'est décroissante sur aucun intervalle de $\mathbb{R}$. Par parité, f n'est monotone sur aucun intervalle.

Partie C – Fonction de Takagi-Landsberg de paramètre $\frac{1}{4}$

1. Il suffit d'appliquer la Partie B.
2. Le résultat s'obtient par simple vérification.
3. Posons $\varphi = h - p$. D'après les questions précédentes, on a $\varphi(x) = \frac{1}{4}\varphi(2x)$ pour tout $x \in [0, \frac{1}{2}]$ et φ s'annule sur $\mathbb{Z}$. Montrons, par récurrence sur $n \geq 1$, la propriété $\mathscr{P}_n$:" $\varphi(\frac{k}{2^n}) = 0$ pour tout $k \in \mathbb{Z}$".

- $\mathscr{P}_0$ est vraie d'après la remarque précédente.
- Soit $n > 0$. Supposons $\mathscr{P}_n$ vraie et montrons $\mathscr{P}_{n+1}$. Avec la parité et la périodicité, il suffit de considérer $k \in [\![0, 2^n]\!]$, $\varphi(\frac{k}{2^n}) = \frac{1}{4}\varphi(\frac{k}{2^{n-1}}) = 0$ d'où le résultat.

La fonction φ est nulle sur les nombres dyadiques (qui forment un ensemble dense dans $\mathbb{R}$) et est continue donc est identiquement nulle sur $\mathbb{R}$.

Chapitre 9

Études locales et asymptotiques

1. Relations de comparaison — **2.** Développements limités — **3.** Applications

Objectifs et compétences du programme

- Exprimer le comportement asymptotique d'une suite avec les relations de comparaison.
- Exprimer le comportement local ou asymptotique d'une fonction avec les relations de comparaison.
- Connaître les développements limités de références.
- Calculer un développement limité à partir des propriétés opératoires.

Brook Taylor, mathématicien britannique (1685-1731)
Il a laissé son nom aux multiples formules de Taylor (qui lui sont parfois abusivement attribuées). On trouve la première trace de cette formule dans l'ouvrage *Methodus Incrementorum Directa et Inversa* (1715) même si son importance est longtemps restée incomprise (essentiellement jusqu'à Lagrange). Dans ce même livre apparaît la formule d'intégration par parties et quelques autres éléments de calcul différentiel devenus classiques.
Il a également rédigé un second livre *Linear Perspective* et lancé des défis aux mathématiciens du continent (notamment les Bernoulli) qui ont permis de faire avancer les mathématiques.

COURS

Pour lever les indéterminations dans les calculs de limites, il apparaît nécessaire de savoir exprimer plus précisément le comportement d'une fonction au voisinage de $+\infty$ ou d'un réel. Un outil efficace pour cela est celui de développement limité. À partir de développements de références (souvent obtenus par la formule de Taylor-Young), on peut calculer ceux de nombreuses fonctions.

1. Relations de comparaison

1.1. Relations de comparaison pour les suites (rappels)

Rappelons la définition des relations de comparaison pour les suites que l'on cherche à étendre aux fonctions.

Définition 9.1.

Soit $(u_n)_n$ et $(v_n)_n$ deux suites ne s'annulant pas à partir d'un certain rang.
▹ La suite $(u_n)_n$ est dominée par $(v_n)_n$ si la suite $\left(\frac{u_n}{v_n}\right)_n$ est bornée. On note $u_n = O(v_n)$.
▹ La suite $(u_n)_n$ est négligeable devant $(v_n)_n$ si la suite $\left(\frac{u_n}{v_n}\right)_n$ est de limite nulle. On note $u_n = o(v_n)$.
▹ La suite $(u_n)_n$ est équivalente à $(v_n)_n$ si la suite $\left(\frac{u_n}{v_n}\right)_n$ est de limite 1. On note $u_n \sim v_n$.

Ces définitions s'étendent en s'affranchissant de l'hypothèse de non annulation et les propriétés de manipulations ont déjà été énoncées au chapitre 6.
Voici un exemple de manipulation un peu plus élaborée pour obtenir un développement asymptotique et non plus seulement un équivalent.

Exemple

Notons, pour tout $n \in \mathbb{N}$, x_n l'unique réel de $]-\frac{\pi}{2}+n\pi, \frac{\pi}{2}+n\pi[$ tel que $\tan x_n = x_n$ (son existence découle directement de l'étude de la fonction $x \mapsto \tan x - x$). Étudions le comportement asymptotique de la suite $(x_n)_n$.
▹ Remarquons tout d'abord que, par définition, $-\frac{\pi}{2}+n\pi < x_n < \frac{\pi}{2}+n\pi$ donc $x_n \sim n\pi$ par le lemme d'encadrement (après avoir divisé par $n\pi$).
▹ Posons, pour tout n, $y_n = x_n - n\pi$. Alors, en remplaçant dans la relation de x_n, $\tan y_n = n\pi + y_n$ et $y_n \in]-\frac{\pi}{2}, \frac{\pi}{2}[$ donc $\tan y_n \to +\infty$ et par suite (avec la limite de arctan en $+\infty$), $y_n = (\arctan(\tan y_n) \to \frac{\pi}{2}$.
▹ Posons, pour tout n, $z_n = x_n - n\pi - \frac{\pi}{2}$. Alors, en remplaçant à nouveau dans la relation de x_n, $\tan(z_n + \frac{\pi}{2}) = n\pi + \frac{\pi}{2} + z_n$ soit encore

$$-\frac{1}{\tan z_n} = n\pi + \frac{\pi}{2} + z_n.$$

Comme $z_n \to 0$, $\tan z_n \sim z_n$ et donc en passant à l'inverse

$$z_n \sim -\frac{1}{n\pi + \frac{\pi}{2} + z_n}.$$

D'où $z_n \sim -\frac{1}{n\pi}$ soit $z_n = -\frac{1}{n\pi} + o\left(\frac{1}{n}\right)$.
En remplaçant dans l'expression de x_n, on obtient $x_n = n\pi + \frac{\pi}{2} - \frac{1}{n\pi} + o\left(\frac{1}{n}\right)$.

Conseils méthodologiques

Pour obtenir un développement asymptotique d'une suite $(u_n)_n$ dont on connaît un équivalent simple, on cherche un équivalent de la suite $(v_n)_n$ définie comme la différence de u_n à cet équivalent (et on exploite que v_n est négligeable par rapport à cet équivalent).

Désormais on souhaite pouvoir généraliser ce procédé pour étudier des fonctions au voisinage d'un réel ou de $\pm\infty$. Remarquons que cela permettra notamment de revenir sur les suites de la forme $(f(n))_n$ où $f : \mathbb{R} \to \mathbb{R}$ est donnée.

1.2. Relations de comparaison pour les fonctions

Précisons les relations de comparaison pour les fonctions. On commencera par le cas où les fonctions ne s'annulent pas qui permet de comprendre intuitivement les différentes notions puis on traitera le cas général en éliminant cette hypothèse surnuméraire.

Définition 9.2.

Soit $f, g : I \to \mathbb{R}$ et $a \in \mathbb{R} \cup \{\pm\infty\}$ une borne ou un point de l'intervalle I. On suppose que g ne s'annule pas sur un voisinage de a.

▹ La fonction g domine f en a si la fonction quotient $\frac{f}{g}$ est bornée au voisinage de a.

▹ La fonction f est négligeable par rapport à g en a si la fonction quotient $\frac{f}{g}$ tend vers 0 en a.

▹ Les fonctions f et g sont équivalentes en a si la fonction quotient $\frac{f}{g}$ tend vers 1 en a.

En s'affranchissant de l'hypothèse sur g, on obtient les définitions complètes suivantes en $a \in \mathbb{R}$.

Définition 9.3.

Soit $f, g : I \to \mathbb{R}$ et $a \in \mathbb{R}$ une borne ou un point de l'intervalle I.

▹ La fonction g domine f en a si

$$\exists C \geq 0,\ \exists \eta > 0,\ \forall x \in I \cap]a-\eta, a+\eta[, \qquad |f(x)| \leq C|g(x)|.$$

On note $f(x) \underset{a}{=} O(g(x))$.

▹ La fonction f est négligeable par rapport à g en a si

$$\forall \varepsilon > 0,\ \exists \eta > 0,\ \forall x \in I \cap]a-\eta, a+\eta[, \qquad |f(x)| \leq \varepsilon|g(x)|.$$

On note $f(x) \underset{a}{=} o(g(x))$.

▹ Les fonctions f et g sont équivalentes en a si $f(x) - g(x) \underset{a}{=} o(g(x))$ ou de manière équivalente

$$f(x) \underset{a}{=} g(x) + o(g(x)).$$

On note $f(x) \underset{a}{\sim} g(x)$.

On peut, bien entendu, définir ces relations avec des phrases quantifiées en $\pm\infty$.

Exemple

Soit $c \neq 0$. Traduisons les relations de comparaison d'une fonction f avec la fonction constante égale à c.

▹ $f(x) \underset{a}{=} O(c)$ si, et seulement si, f est bornée au voisinage de a.

▹ $f(x) \underset{a}{=} o(c)$ si, et seulement si, f tend vers 0 en a.

▹ $f(x) \underset{a}{\sim} c$ si, et seulement si, f tend vers c en a.
On remarque en particulier qu'il est équivalent de dire $f(x) \underset{a}{=} O(c)$ ou $f(x) \underset{a}{=} O(1)$, $f(x) \underset{a}{=} o(c)$ ou $f(x) \underset{a}{=} o(1)$ mais que les propositions $f(x) \underset{a}{\sim} c$ et $f(x) \underset{a}{\sim} 1$ sont bien différentes pour $c \neq 1$.

Plus généralement, on obtient les résultats suivants sur les limites dans les relations de comparaison.

Proposition 9.4.

Soit $f, g : I \to \mathbb{R}$ et a une borne ou un point de l'intervalle I où g admet une limite finie ℓ.
▹ Si $f(x) \underset{a}{=} O(g(x))$, alors f est bornée au voisinage de a.
▹ Si $f(x) \underset{a}{=} o(g(x))$, alors f tend vers 0 en a.
▹ Si $f(x) \underset{a}{\sim} g(x)$, alors f tend vers ℓ en a.

Remarque

Soit $f, g : I \to \mathbb{R}$ et a une borne ou un point de l'intervalle I où f et g admettent la même limite finie $\ell \neq 0$. Alors, $\frac{f}{g}$ tend vers $\frac{\ell}{\ell} = 1$ en A donc $f(x) \underset{a}{\sim} g(x)$.
Cette réciproque est en revanche fausse en général pour $\ell = 0$, $\ell = +\infty$ ou $\ell = -\infty$.

On peut étendre toutes les propriétés opératoires étudiées pour les suites (avec des démonstrations analogues). En voici une par exemple.

Proposition 9.5.

Soit $f, \tilde{f}, g, \tilde{g} : I \to \mathbb{R}$ et a une borne ou un point de l'intervalle I. Supposons que $f(x) \underset{a}{\sim} \tilde{f}(x)$ et $g(x) \underset{a}{\sim} \tilde{g}(x)$. Alors, $f(x)g(x) \underset{a}{\sim} \tilde{f}(x)\tilde{g}(x)$ et $\frac{1}{f(x)} \underset{a}{\sim} \frac{1}{\tilde{f}(x)}$.

Remarque

Précisons deux usages afin d'éviter des erreurs de manipulations.
▹ Il n'y a pas de règle générale indiquant une composition des équivalents : il convient de justifier soigneusement lorsque l'on peut le faire.
▹ Un équivalent indiquant un « ordre de grandeur », il est inutile d'écrire des équivalents à plusieurs termes dont certains sont négligeables par rapport aux autres.

1.3. Exemples

La définition de dérivée en a donne immédiatement le résultat suivant.

Proposition 9.6.

Soit $f : I \to \mathbb{R}$ et $a \in I$ où f est dérivable avec $f'(a) \neq 0$. Alors,

$$f(x) - f(a) \underset{a}{\sim} f'(a)(x-a).$$

Exemple

Ainsi, pour tout $\alpha \in \mathbb{R}$,

$$\sin(x) \underset{0}{\sim} x, \qquad (1+x)^\alpha - 1 \underset{0}{\sim} \alpha x, \qquad e^x - 1 \underset{0}{\sim} x.$$

On peut aussi utiliser ces équivalents pour les suites par caractérisation séquentielle des limites : par exemple, si la suite $(u_n)_n$ converge vers 0 sans s'annuler à partir d'un certain rang, alors

$$(1+u_n)^\alpha - 1 \sim \alpha u_n.$$

On peut désormais réécrire les propriétés de croissances comparées avec la relation de négligeabilité.

Proposition 9.7. Croissances comparées

Pour tout $r > 0$,

$$\ln x \underset{+\infty}{=} o(x^r), \qquad x^r \underset{+\infty}{=} o(e^x).$$

$$\ln x \underset{0}{=} o(x^{-r}).$$

$$e^x \underset{-\infty}{=} o(|x|^{-r}).$$

Exemple

Soit $f, g : I \to \mathbb{R}$ et a une borne ou de l'intervalle I telles que $f(x) \underset{a}{\sim} g(x)$ et $f(x) \to +\infty$ lorsque $x \to a$. Alors $\ln(f(x)) \underset{a}{\sim} \ln(g(x))$.

En effet, sur un voisinage de a, f et g sont à valeurs strictement positives (car de limite $+\infty$. Ainsi,

$$\ln(f(x)) = \ln(g(x)) + \ln\frac{f(x)}{g(x)} \underset{a}{=} \ln(g(x)) + o(1).$$

Comme $\ln(f(x)) \to +\infty$, on obtient l'équivalent annoncé.

2. Développements limités

2.1. Définitions

Définition 9.8.

Une fonction f admet un développement limité (en abrégé DL) à l'ordre n en un point a de son domaine de définition s'il existe des réels $a_0, \ldots, a_n$ tels que

$$f(x) \underset{a}{=} \sum_{k=0}^{n} a_k (x-a)^k + o((x-a)^n),$$

ou, de manière équivalente,

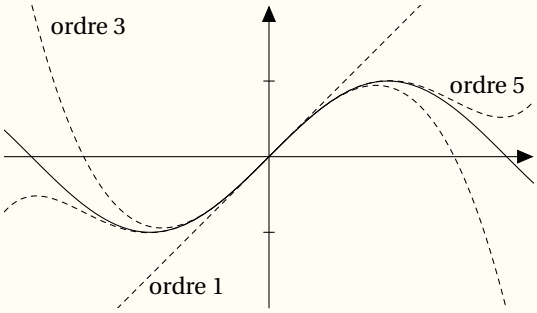
$$f(a+h) \underset{0}{=} \sum_{k=0}^{n} a_k h^k + o(h^n).$$

La partie régulière de ce développement limité est le polynôme

$$\sum_{k=0}^{n} a_k (X-a)^k.$$

Intrinsèquement, un développement limité est une approximation de la fonction au voisinage du point et l'ordre de celui-ci indique la précision de l'approximation.

Voici par exemple, les courbes représentatives de la fonction sin et des parties régulières de ses développements limités en 0 d'ordres 1, 3 et 5.



Conseils méthodologiques

Pour ne pas commettre d'erreur, il convient de conserver l'ordre d'écriture du développement limité du terme le plus important vers le terme le plus petit (pour la relation de négligeabilité) et de ne pas développer abusivement les termes $(x-a)^k$ lors d'un développement en a sous peine de perdre cette progression en précision dans la lecture du DL. Par exemple, l'écriture suivante est pertinente

$$f(x) \underset{a}{=} 1+2(x-a)^2+o((x-a)^2),$$

alors que celle-ci ne l'est pas

$$f(x) \underset{a}{=} 1+2a^2-4ax+2x^2+o((x-a)^2).$$

Proposition 9.9.

La partie régulière d'un développement limité est unique.

Démonstration

Quitte à translater, effectuons la démonstration en 0.

S'il existe P et Q, polynômes de degré au plus n tels que $P(x) \underset{0}{=} Q(x)+o(x^n)$, alors $\frac{P(x)-Q(x)}{x^n} \to 0$ quand $x \to 0$ ce qui n'est possible que si $P-Q=0$. ■

Remarque

Une fonction f admet un développement limité à l'ordre 0 (respectivement 1) en a si, et seulement si, elle est continue (respectivement dérivable) en a.

Remarque

Si f admet un développement limité à l'ordre n en a, alors f admet un développement limité à l'ordre p en a pour tout $p \leq n$ et la partie régulière est obtenue en tronquant la somme à l'exposant p (car pour tout $k > p$, $(x-a)^k \underset{a}{=} o((x-a)^p)$).

Par exemple, si $f(x) \underset{0}{=} 1+2x+3x^2+4x^3+o(x^3)$, alors $f(x) \underset{0}{=} 1+2x+3x^2+o(x^2)$.

En particulier, si l'on dispose déjà d'un développement limité à l'ordre p et que l'on cherche un développement limité à l'ordre n, il faut déterminer les éventuels coefficients d'indice entre $p+1$ et n.

Exemple

Considérons la fonction $f : x \mapsto x^{\frac{11}{3}}$.
▹ Elle admet un développement limité en 0 à l'ordre 3 puisque $f(x) \underset{0}{=} o(x^3)$.
▹ Montrons par l'absurde que f n'admet pas de développement limité en 0 d'ordre 4 ; en effet, si f admettait un DL d'ordre 4 en 0, il serait de la forme $f(x) \underset{0}{=} a x^4 + o(x^4)$ car la troncature de la partie régulière à l'ordre 3 est nulle. Par conséquent, on aurait $\frac{f(x)}{x^4} \to a$ quand $x \to 0$. Or, $x \mapsto \frac{x^{\frac{11}{3}}}{x^4} = x^{\frac{-1}{3}}$ n'admet pas de limite finie en 0 : contradiction.

Proposition 9.10.

Si f admet un développement limité en 0 et que f est paire (respectivement impaire), alors la partie régulière du développement est paire (respectivement impaire).

Démonstration
Traitons le cas où f est paire de développement limité

$$f(x) \underset{0}{=} \sum_{k=0}^{n} a_k x^k + o(x^n).$$

Alors,

$$f(x) = f(-x) \underset{0}{=} \sum_{k=0}^{n} a_k (-1)^k x^k + o(x^n).$$

Comparons : l'unicité de la partie régulière entraîne que $a_k = 0$ pour tout k impair. ■

Enfin, voici une dernière proposition essentielle car elle permet de se ramener à des développements limités en 0 (et donc de limiter le nombre de développements usuels à retenir).

Proposition 9.11.

Une fonction f admet un développement limité à l'ordre n en a si, et seulement si, $x \mapsto f(x+a)$ admet un développement limité à l'ordre n en 0.

Conseils méthodologiques

Très souvent, on translate la fonction pour se ramener à l'étudier en 0 et pouvoir utiliser tous les développements limités de reférence. Pour cela, on pose $x = a + h$ et on remarque que lorsque x est proche de a, h est proche de 0.

2.2. Formule de Taylor-Young

Proposition 9.12. Formule de Taylor-Young

Soit f une fonction de classe $\mathscr{C}^n$ sur I et $a \in I$. Alors, $f(x) \underset{a}{=} \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + o((x-a)^n)$.

Démonstration

Montrons cette propriété par récurrence sur $n \in \mathbb{N} \setminus \{0\}$.

▹ La propriété est vraie pour $n = 1$ par définition de la dérivabilité en a.

▹ Soit $n \geq 1$ tel que la propriété soit vraie au rang n ; considérons f une fonction $n+1$ fois dérivable sur I. Appliquons l'hypothèse de récurrence à f' et posons ε la fonction continue de limite nulle en a définie par

$$\varepsilon(x) = \frac{1}{(x-a)^n}\left(f'(x) - \sum_{k=0}^{n} \frac{f^{(k+1)}(a)}{k!}(x-a)^k\right).$$

Intégrons cette relation entre a et x. Pour tout $\varepsilon > 0$, il existe $\eta > 0$ tel que

$$\forall x \in I, \qquad |x-a| \leq \eta \quad \Rightarrow \quad |\varepsilon(x)| \leq \varepsilon.$$

Ainsi, pour tout $x \in I$ tel que $|x-a| \leq \eta$,

$$\left|\frac{1}{(x-a)^{n+1}} \int_a^x \varepsilon(t)(t-a)^n \,\mathrm{d}t\right| \leq \frac{1}{n+1}\varepsilon,$$

d'où le résultat. ■

Remarque

D'après cette démonstration, le résultat demeure vrai avec l'hypothèse plus faible d'une fonction f n fois dérivable sur I.

Remarque

Avec davantage de connaissance de la formule de Taylor avec reste intégral à l'ordre $n-1$, on peut obtenir une preuve plus simple :

$$f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k = \int_a^x \frac{(x-t)^{n-1}}{(n-1)!} f^{(n)}(t)\,\mathrm{d}t - \frac{f^{(n)}(a)}{n!}(x-a)^n.$$

Le second membre se réécrit $\int_a^x \frac{(x-t)^{n-1}}{(n-1)!}(f^{(n)}(t) - f^{(n)}(a))\,\mathrm{d}t$. Fixons $\varepsilon > 0$ et utilisons dorénavant la continuité de $f^{(n)}$ en a : il existe $\eta > 0$ tel que

$$\forall x \in I, \qquad |x-a| \leq \eta \quad \Rightarrow \quad |f^{(n)}(x) - f^{(n)}(a)| \leq \varepsilon.$$

Ainsi, pour tout $x \in I$ vérifiant $a \leq x \leq a + \eta$,

$$\left|\int_a^x \frac{(x-t)^{n-1}}{(n-1)!}(f^{(n)}(t) - f^{(n)}(a))\,\mathrm{d}t\right| \leq \varepsilon \frac{(x-a)^n}{n!}.$$

On conclut par encadrement que le terme de gauche est négligeable devant $(x-a)^n$.

La formule de Taylor-Young donne directement l'existence de développement limité.

Corollaire 9.13.

Une fonction de classe $\mathscr{C}^n$ en a admet un développement limité à l'ordre n en a.

En revanche, la réciproque est fausse : il existe des fonctions admettant un développement limité à l'ordre $n \geq 2$ sans être de classe $\mathscr{C}^n$.

Exemple

La fonction f définie par

$$f : x \mapsto \begin{cases} 0 & \text{si } x = 0 \\ x^3 \sin \frac{1}{x} & \text{sinon} \end{cases}$$

n'est pas de classe $\mathscr{C}^2$ en 0 et pourtant elle admet un développement limité d'ordre 2 en 0 car $f(x) \underset{0}{=} o(x^2)$.

Exemple

Les fonctions exp, cos et sin sont de classe $\mathscr{C}^\infty$. La formule de Taylor-Young en 0 appliquée à ces fonctions donne

$$\exp(x) \underset{0}{=} \sum_{k=0}^{n} \frac{1}{k!} x^k + o(x^n),$$
$$\sin(x) \underset{0}{=} \sum_{k=0}^{n} \frac{(-1)^k}{(2k+1)!} x^{2k+1} + o(x^{2n+2}),$$
$$\cos(x) \underset{0}{=} \sum_{k=0}^{n} \frac{(-1)^k}{(2k)!} x^{2k} + o(x^{2n+1}),$$

en utilisant les calculs suivants des dérivées $n^{\text{ièmes}}$

$$\cos^{(n)} : x \mapsto \cos\left(x + n\frac{\pi}{2}\right), \qquad \sin^{(n)} : x \mapsto \sin\left(x + n\frac{\pi}{2}\right), \qquad \exp^{(n)} : x \mapsto \exp(x).$$

Exemple

La fonction $f : x \mapsto (1+x)^\alpha$ est de classe $\mathscr{C}^\infty$ sur un voisinage de 0. Sa dérivée $n^{\text{ième}}$ est $f^{(n)} : x \mapsto \alpha(\alpha-1)\dots(\alpha-n+1)(1+x)^{\alpha-n}$. Par conséquent, la formule de Taylor-Young s'écrit

$$(1+x)^\alpha \underset{0}{=} 1 + \sum_{k=1}^{n} \frac{\alpha(\alpha-1)\dots(\alpha-k+1)}{k!} x^k + o(x^n).$$

Exemple

Calculons le développement limité de tan en $\frac{\pi}{4}$ en appliquant la formule de Taylor-Young. Commençons par calculer les dérivées successives

$$\tan' = \tan^2 + 1$$
$$\tan'' = 2(\tan^2+1)\tan$$
$$\tan''' = 6(\tan^2+1)\tan^2 + 2(\tan^2+1).$$

Ainsi, $\tan\frac{\pi}{4} = 1$, $\tan'\frac{\pi}{4} = 2$, $\tan''\frac{\pi}{4} = 4$ et $\tan'''\frac{\pi}{4} = 16$ donc

$$\tan x \underset{\pi/4}{=} \frac{1}{0!} + \frac{2}{1!}\left(x - \frac{\pi}{4}\right) + \frac{4}{2!}\left(x - \frac{\pi}{4}\right)^2 + \frac{16}{3!}\left(x - \frac{\pi}{4}\right)^3 + o\left(\left(x - \frac{\pi}{4}\right)^3\right)$$
$$\underset{\pi/4}{=} 1 + \left(x - \frac{\pi}{4}\right) + 2\left(x - \frac{\pi}{4}\right)^2 + \frac{8}{3}\left(x - \frac{\pi}{4}\right)^3 + o\left(\left(x - \frac{\pi}{4}\right)^3\right)$$

Exemple

Soit f une fonction deux fois dérivable en $x \in I$. Alors, pour tout $h \neq 0$ tel que $]x-h, x+h[\subset I$,

$$\begin{aligned}\frac{f(x+h)+f(x-h)-2f(x)}{h^2} &= f''(x) + \frac{f(x+h)-f(x)-hf'(x)-\frac{h^2}{2}f''(x)}{h^2} \\ &\quad + \frac{f(x-h)-f(x)+hf'(x)-\frac{h^2}{2}f''(x)}{h^2} \\ &\underset{0}{=} f''(x) + \frac{o(h^2)}{h^2} + \frac{o(h^2)}{h^2} \underset{0}{=} f''(x) + o(1),\end{aligned}$$

d'après la formule de Taylor-Young pour la fonction f deux fois dérivable, soit encore

$$\lim_{h \to 0} \frac{f(x+h)+f(x-h)-2f(x)}{h^2} = f''(x).$$

2.3. Manipulations

On sait donc obtenir des développements limités par la formule de Taylor-Young. Cette section permet de voir comment utiliser quelques développements simples pour en obtenir d'autres.
Pour simplifier la rédaction des propriétés suivantes, on adopte la définition technique suivante.

Définition 9.14.

La troncature (en 0) à l'ordre $p \leq n$ du polynôme $\sum_{k=0}^{n} a_k X^k$ est le polynôme $\sum_{k=0}^{p} a_k X^k$ qui ne contient que les monômes de degré au plus p.

Proposition 9.15.

▹ Soit f, g deux fonctions admettant un développement limité à l'ordre n en a de parties régulières P_n et Q_n et $\lambda \in \mathbb{R}$.
Alors, $f + \lambda g$ admet un développement limité à l'ordre n en a de partie régulière $P_n + \lambda Q_n$.
▹ Soit f, g deux fonctions admettant un développement limité à l'ordre n en 0 de parties régulières P_n et Q_n.
Alors, $f.g$ admet un développement limité à l'ordre n en 0 de partie régulière la troncature à l'ordre n de $P_n.Q_n$.

On fera attention que le résultat pour les produits est énoncé en 0 pour éviter les erreurs de troncature.

Démonstration

▹ On trouve directement en faisant la combinaison linéaire des développements limités :

$$(f+\lambda g)(x) \underset{a}{=} (P_n + \lambda Q_n)(x) + o((x-a)^n).$$

On conclut par unicité de la partie régulière.
▹ Écrivons les développements limités sous la forme $f(x) = P_n(x) + x^n \varepsilon(x)$ et $g(x) = Q_n(x) + x^n \vartheta(x)$ où ε et ϑ sont deux fonctions de limite nulle en 0. Alors,

$$f(x)g(x) = P_n(x)Q_n(x) + x^n(P_n(x)\vartheta(x) + Q_n(x)\varepsilon(x) + x^n \vartheta(x)\varepsilon(x)).$$

Or, $x \mapsto P_n(x)\vartheta(x) + Q_n(x)\varepsilon(x) + x^n \vartheta(x)\varepsilon(x)$ est de limite nulle en 0. D'où le résultat. ■

Le résultat suivant décrit les calculs de développement limité par composition. On se place en 0 pour simplifier l'énoncé (en pratique, on se ramènera toujours à l'étude en 0).

Proposition 9.16.

Soit f, g deux fonctions admettant un développement limité à l'ordre n en 0 de parties régulières P_n et Q_n. On suppose, de plus, que $g(0)=0$. Alors, $f\circ g$ admet un développement limité à l'ordre n en 0 de partie régulière la troncature à l'ordre n de $P_n\circ Q_n$.

Démonstration

Écrivons, à nouveau, les développements limités de forme $f(x)=P_n(x)+x^n\varepsilon(x)$ et $g(x)=Q_n(x)+x^n\vartheta(x)$ où ε et ϑ sont deux fonctions de limite nulle en 0. Alors, pour tout $k\in\mathbb{N}$, la formule du binôme est

$$\begin{aligned}(g(x))^k &= (Q_n(x)+x^n\vartheta(x))^k\\ &\underset{0}{=} Q_n(x)^k+o(x^n).\end{aligned}$$

Ainsi, $P_n(g(x))\underset{0}{=}P_n(Q_n(x))+o(x^n)$. On en déduit, en utilisant $g(0)=0$ (donc $Q_n(0)=0$),

$$\begin{aligned}f\circ g(x) &= P_n(g(x))+g(x)^n\varepsilon(g(x))\\ &\underset{0}{=} P_n(Q_n(x))+o(x^n).\end{aligned}$$

■

Exemple

Calculons le développement limité à l'ordre 4 en 0 de la fonction composée $x\mapsto e^{\cos(x)}$. On connaît les développements limités en 0 de cos et exp ; malheureusement $\cos(0)\neq 0$ donc on ne peut appliquer directement le résultat ci-dessus. On écrit donc

$$\begin{aligned}e^{\cos(x)} &= e.e^{\cos(x)-1}\\ &\underset{0}{=} e.e^{-\frac{1}{2}x^2+\frac{1}{24}x^4+o(x^4)}\\ &\underset{0}{=} e.\left[1+\left(-\frac{1}{2}x^2+\frac{1}{24}x^4+o(x^4)\right)+\frac{1}{2}\left(-\frac{1}{2}x^2+\frac{1}{24}x^4+o(x^4)\right)^2+o(x^4)\right]\\ &\underset{0}{=} e-\frac{e}{2}x^2+\frac{e}{6}x^4+o(x^4).\end{aligned}$$

Exemple

Soit $f:\mathbb{R}\to\mathbb{R}$ une fonction bijective de classe $\mathscr{C}^2$ telle que $f(x)\underset{0}{=}a_1x+a_2x^2+o(x^2)$ avec $a_1\neq 0$. Alors, f^{-1} est de classe $\mathscr{C}^2$ au voisinage de 0 (car $f'(0)\neq 0$) donc admet un développement limité à l'ordre 2 en $f(0)=0$. Notons $f^{-1}(x)\underset{0}{=}b_1x+b_2x^2+o(x^2)$. Déterminons ce développement en écrivant que $x=f\circ f^{-1}(x)$:

$$\begin{aligned}x &\underset{0}{=} a_1\left(b_1x+b_2x^2+o(x^2)\right)+a_2\left(b_1x+b_2x^2+o(x^2)\right)^2+o(x^2)\\ &\underset{0}{=} a_1b_1x+(a_1b_2+a_2b_1^2)x^2+o(x^2).\end{aligned}$$

En utilisant l'unicité de la partie régulière, $b_1=\frac{1}{a_1}$ et $b_2=-\frac{a_2}{a_1^3}$.

Proposition 9.17.

Soit f une fonction admettant un développement limité à l'ordre n en 0 telle que $f(0)\neq 0$.
Alors, $\frac{1}{f}$ admet un développement limité à l'ordre n en 0.

Démonstration
Écrivons $\frac{1}{f(x)}=\frac{1}{f(0)}\frac{1}{1+\frac{f(x)-f(0)}{f(0)}}$ et appliquons le résultat de composition précédemment établi aux fonctions $x\longmapsto\frac{f(x)-f(0)}{f(0)}$ (qui s'annule en 0) et $x\longmapsto\frac{1}{1+x}$. ■

Exemple

Calculons successivement les développements en 0 à l'ordre 5 de $\frac{1}{\cos}$ et tan.

$$\begin{aligned}\frac{1}{\cos(x)}&\underset{0}{=}\frac{1}{1-\left(\frac{1}{2}x^2-\frac{1}{24}x^4+o(x^5)\right)}\\&\underset{0}{=}1+\left(\frac{1}{2}x^2-\frac{1}{24}x^4+o(x^5)\right)+\left(\frac{1}{2}x^2-\frac{1}{24}x^4+o(x^5)\right)^2+o(x^5)\\&\underset{0}{=}1+\frac{1}{2}x^2+\frac{5}{24}x^4+o(x^5);\end{aligned}$$

d'où, par produit,

$$\begin{aligned}\tan(x)&\underset{0}{=}\left(x-\frac{1}{6}x^3+\frac{1}{120}x^5+o(x^5)\right)\left(1+\frac{1}{2}x^2+\frac{5}{24}x^4+o(x^5)\right)\\&\underset{0}{=}x+\frac{1}{3}x^3+\frac{2}{15}x^5+o(x^5).\end{aligned}$$

Proposition 9.18.

Soit f une fonction admettant un développement limité à l'ordre n en a de partie régulière P_n.
▹ Toute primitive F de f admet un développement limité à l'ordre $n+1$ en a de partie régulière la primitive de P_n dont la valeur en a est $F(a)$.
▹ Si f' admet un développement limité à l'ordre $n-1$ en a, alors la partie régulière de ce développement limité est P_n'.

On notera bien l'asymétrie des deux résultats d'intégration et de dérivation (on connaît d'autres tels résultats : on peut par exemple intégrer une inégalité mais pas la dériver).

Démonstration
▹ Par définition du terme $o((x-a)^n)$, il existe, pour tout $\varepsilon>0$, un voisinage de a tel que :

$$-\varepsilon(x-a)^n\leq f(x)-P_n(x)\leq\varepsilon(x-a)^n.$$

En intégrant ces inégalités, on obtient le résultat énoncé.
▹ Si f' admet un développement limité à l'ordre $n-1$ en a de partie régulière Q_n, alors d'après le point précédent $P_n(x)-P_n(a)=\int_a^x Q_n(t)\,\mathrm{d}t$ et donc $Q_n(x)=P_n'(x)$.

■

Exemple

On peut appliquer ce résultat aux primitives de $x \mapsto (1+x)^\alpha$ et aux fonctions associées, on obtient

$$\ln(1+x) \underset{0}{=} \sum_{k=1}^{n} \frac{(-1)^{k+1}}{k} x^k + o(x^n),$$
$$\arctan(x) \underset{0}{=} \sum_{k=0}^{n} \frac{(-1)^k}{2k+1} x^{2k+1} + o(x^{2n+2}),$$
$$\arcsin(x) \underset{0}{=} \sum_{k=0}^{n} \frac{(2k)!}{(2k+1)(2^k.k!)^2} x^{2k+1} + o(x^{2n+2}).$$

Exemple

Remarquons que $\tan(x) \underset{0}{=} o(1)$ et, par conséquent, que $\tan'(x) = 1 + \tan^2(x) \underset{0}{=} 1 + o(1)$.
Ainsi, en intégrant, $\tan(x) \underset{0}{=} x + o(x)$; par ailleurs, tan admet un développement à l'ordre 2 (comme fonction de classe $\mathscr{C}^2$) et le terme d'ordre 2 est nul car tan est impair :

$$\tan(x) \underset{0}{=} x + o(x^2).$$

En réutilisant la relation entre tan′ et tan, on en déduit le développement limité de la fonction tan à l'ordre que l'on souhaite.

3. Applications

3.1. Calculs de limite

Voici quelques exemples d'utilisation des développements limités principalement pour calculer une limite (en levant une indétermination).

Exemple

Cherchons la limite en 0 de $x \mapsto \left(\frac{e^{\alpha x}+e^{\beta x}}{2}\right)^{\frac{1}{x}}$.
Le calcul du développement limité donne

$$\begin{aligned}\left(\frac{e^{\alpha x}+e^{\beta x}}{2}\right)^{\frac{1}{x}} &\underset{0}{=} \left(1+\frac{\alpha+\beta}{2}x+o(x)\right)^{\frac{1}{x}}\\ &\underset{0}{=} e^{\frac{1}{x}\ln\left(1+\frac{\alpha+\beta}{2}x+o(x)\right)}\\ &\underset{0}{=} e^{\frac{\alpha+\beta}{2}+o(1)}.\end{aligned}$$

La limite recherchée est donc $e^{\frac{\alpha+\beta}{2}}$.

Exemple

Cherchons un équivalent en 0 de $x \mapsto (\mathrm{ch}(x))^x - (\cos(x))^x$. En effectuant le calcul du développe-

ment limité de $(\mathrm{ch}(x))^x$, on trouve

$$\begin{aligned}
(\mathrm{ch}(x))^x &= e^{x\ln(\mathrm{ch}(x))}\\
&\underset{0}{=} e^{x\ln\left(1+\frac{1}{2}x^2+\frac{1}{24}x^4+o(x^5)\right)}\\
&\underset{0}{=} e^{x\left(\left[\frac{1}{2}x^2+\frac{1}{24}x^4+o(x^5)\right]-\frac{1}{2}\left[\frac{1}{2}x^2+\frac{1}{24}x^4+o(x^5)\right]^2+o(x^5)\right)}\\
&\underset{0}{=} e^{\frac{1}{2}x^3-\frac{1}{12}x^5+o(x^6)}\\
&\underset{0}{=} 1+\left[\frac{1}{2}x^3-\frac{1}{12}x^5+o(x^6)\right]+\frac{1}{2}\left[\frac{1}{2}x^3-\frac{1}{12}x^5+o(x^6)\right]^2+o(x^6)\\
&\underset{0}{=} 1+\frac{1}{2}x^3-\frac{1}{12}x^5+\frac{1}{8}x^6+o(x^6).
\end{aligned}$$

De même,

$$(\cos(x))^x \underset{0}{=} 1-\frac{1}{2}x^3-\frac{1}{12}x^5+\frac{1}{8}x^6+o(x^6).$$

Par conséquent, $(\mathrm{ch}(x))^x-(\cos(x))^x \underset{0}{=} x^3+o(x^6)$.

3.2. Calculs d'asymptotes

Pour étudier une fonction f au voisinage de $\pm\infty$ et notamment d'éventuelles asymptotes à sa courbe représentative, on introduit la fonction auxiliaire $t\mapsto f\left(\frac{1}{t}\right)$ que l'on développe au voisinage de 0 (voire en 0 par valeurs positives si on se limite à $+\infty$).

On peut ainsi obtenir à la fois les asymptotes et la position par rapport à ces asymptotes en considérant les termes suivants.

Exemple

Étudions les asymptotes de la courbe représentative de $x\mapsto x^2\ln\frac{x}{x-1}$.
Il suffit de développer la fonction au voisinage de $\pm\infty$ en posant $x=\frac{1}{t}$.

$$\begin{aligned}
x^2\ln\frac{x}{x-1} &= t^{-2}\ln\frac{1}{1-t} = -t^{-2}\ln(1-t)\\
&\underset{0}{=} -t^{-2}(-t-\frac{t^2}{2}-\frac{t^3}{3}+o(t^3))\\
&\underset{0}{=} \frac{1}{t}+\frac{1}{2}+\frac{1}{3}t+o(t) \underset{\infty}{=} x+\frac{1}{2}+\frac{1}{3}\frac{1}{x}+o(\frac{1}{x}).
\end{aligned}$$

On en déduit que l'asymptote est $y=x+\frac{1}{2}$ et qu'au voisinage de $+\infty$ la courbe est au-dessus de l'asymptote (car $\frac{1}{3}\frac{1}{x}>0$).

Exemple

Étudions les asymptotes de la courbe représentative de $x\mapsto x\exp\left(\frac{x-1}{x^2+1}\right)$.

Il suffit de développer la fonction au voisinage de $\pm\infty$ en posant $x = \frac{1}{t}$:

$$\begin{aligned} x\exp\left(\frac{x-1}{x^2+1}\right) &= \frac{1}{t}\exp\left(\frac{t-t^2}{1+t^2}\right) \\ &\underset{0}{=} \frac{1}{t}\exp\left(t-t^2+o(t^3)\right) \\ &\underset{0}{=} \frac{1}{t}\left(1+t-\frac{t^2}{2}+o(t^2)\right) \\ &\underset{0}{=} \frac{1}{t}+1-\frac{1}{2}t+o(t) \underset{\infty}{=} x+1-\frac{1}{2}\frac{1}{x}+o(\frac{1}{x}). \end{aligned}$$

On en déduit que l'asymptote est $y = x+1$ et qu'au voisinage de $+\infty$ la courbe est en dessous de l'asymptote (car $-\frac{1}{2}\frac{1}{x} < 0$).

3.3. Développements généralisés

Il arrive que l'on cherche à étudier une fonction au voisinage d'un point où la fonction admet une limite infinie. Pour remplacer les techniques de développements limités qui sont inopérantes dans ce cas, on introduit les développements généralisés.

Définition 9.19.

Soit f une fonction définie sur I et a une borne ou un point de l'intervalle I. f admet un développement généralisé en a s'il existe $r \in \mathbb{R}$ tel que $x \mapsto (x-a)^r f(x)$ admette un développement limité en a.

Exemple

Par exemple,

$$\begin{aligned} \frac{1}{e^x-1} &\underset{0}{=} \frac{1}{x+\frac{x^2}{2}+\frac{x^3}{6}+o(x^3)} \\ &\underset{0}{=} \frac{1}{x}\left(\frac{1}{1+\frac{x}{2}+\frac{x^2}{6}+o(x^2)}\right) \\ &\underset{0}{=} \frac{1}{x}(1-\frac{x}{2}-\frac{x^2}{6}+\frac{x^2}{4}+o(x^2)) \\ &\underset{0}{=} \frac{1}{x}-\frac{1}{2}+\frac{1}{12}x+o(x). \end{aligned}$$

FICHE

SYNTHÈSE

L'objectif de ce chapitre est l'étude locale ou asymptotique des fonctions. Pour cela, on introduit trois relations de comparaison (définies de manière identique aux suites) pour les fonctions. Les définitions simples ci-dessous sont maniables mais requièrent une hypothèse technique de non-annulation.

Relations de comparaison

▹ Soit $(u_n)_n$ et $(v_n)_n$ deux suites ne s'annulant pas à partir d'un certain rang.

- La suite $(u_n)_n$ est dominée par $(v_n)_n$ si la suite $\left(\frac{u_n}{v_n}\right)_n$ est bornée.
- La suite $(u_n)_n$ est négligeable devant $(v_n)_n$ si la suite $\left(\frac{u_n}{v_n}\right)_n$ est de limite nulle.
- La suite $(u_n)_n$ est équivalente à $(v_n)_n$ si la suite $\left(\frac{u_n}{v_n}\right)_n$ est de limite 1.

▹ Soit $f, g : I \to \mathbb{R}$ et $a \in \mathbb{R} \cup \{\pm\infty\}$ une borne ou un point de l'intervalle I. On suppose que g ne s'annule pas sur un voisinage de a.

- g domine f en a si la fonction quotient $\frac{f}{g}$ est bornée au voisinage de a.
- f est négligeable par rapport à g en a si la fonction quotient $\frac{f}{g}$ tend vers 0 en a.
- f et g sont équivalentes en a si la fonction quotient $\frac{f}{g}$ tend vers 1 en a, ou de manière équivalente si $f(x) - g(x) \underset{a}{=} o(g(x))$.

On connaît quelques comparaisons usuelles.

Croissances comparées

Pour tout $r > 0$,

$$\ln x \underset{+\infty}{=} o(x^r), \qquad x^r \underset{+\infty}{=} o(e^x), \quad \ln x \underset{0}{=} o(x^{-r}), \quad e^x \underset{-\infty}{=} o(|x|^{-r}).$$

Poussant plus loin les relations de comparaison, on cherche à obtenir des développements limités.

Développement limité

Une fonction f admet un développement limité à l'ordre n en un point a de son domaine de définition (en abrégé $\mathrm{DL}_n(a)$) s'il existe des réels $a_0, \ldots, a_n$ tels que

$$f(x) \underset{a}{=} \sum_{k=0}^{n} a_k (x-a)^k + o((x-a)^n).$$

La partie régulière de ce développement limité est le polynôme (défini de manière unique)

$$\sum_{k=0}^{n} a_k (X-a)^k.$$

Sous de fortes hypothèses de régularité, la formule de Taylor-Young donne ces développements.

Formule de Taylor-Young

Soit f une fonction de classe $\mathscr{C}^n$ sur I et $a \in I$. Alors,

$$f(x) \underset{a}{=} \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + o((x-a)^n).$$

À partir de cette formule générale et des règles de manipulations des développements limités, on obtient deux familles de DL à connaître : ceux qui se déduisent du développement de la fonction exp et ceux qui se déduisent de la fonction $x \mapsto (1+x)^\alpha$.

Développements usuels en 0

	$DL_n(0)$	$DL_5(0)$
$\exp(x)$	$\sum_{k=0}^{n} \frac{x^k}{k!} + o(x^n)$	$1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + o(x^5)$
$\sin(x)$	$\sum_{k=0}^{n} (-1)^k \frac{x^{2k+1}}{(2k+1)!} + o(x^{2n+2})$	$x - \frac{x^3}{3!} + \frac{x^5}{5!} + o(x^5)$
$\cos(x)$	$\sum_{k=0}^{n} (-1)^k \frac{x^{2k}}{(2k)!} + o(x^{2n+1})$	$1 - \frac{x^2}{2!} + \frac{x^4}{4!} + o(x^5)$
$\mathrm{sh}(x)$	$\sum_{k=0}^{n} \frac{x^{2k+1}}{(2k+1)!} + o(x^{2n+2})$	$x + \frac{x^3}{3!} + \frac{x^5}{5!} + o(x^5)$
$\mathrm{ch}(x)$	$\sum_{k=0}^{n} \frac{x^{2k}}{(2k)!} + o(x^{2n+1})$	$1 + \frac{x^2}{2!} + \frac{x^4}{4!} + o(x^5)$

	$DL_n(0)$	$DL_5(0)$
$(1+x)^\alpha$	$1 + \sum_{k=1}^{n} \frac{\alpha(\alpha-1)\cdots(\alpha-k+1)}{k!} x^k + o(x^n)$	
$\frac{1}{1-x}$	$\sum_{k=0}^{n} x^k + o(x^n)$	$1 + x + x^2 + x^3 + x^4 + x^5 + o(x^5)$
$\ln(1-x)$	$-\sum_{k=1}^{n} \frac{1}{k} x^k + o(x^n)$	$-x - \frac{x^2}{2} - \frac{x^3}{3} - \frac{x^4}{4} - \frac{x^5}{5} + o(x^5)$
$\arctan(x)$	$\sum_{k=0}^{n} (-1)^k \frac{x^{2k+1}}{2k+1} + o(x^{2n+2})$	$x - \frac{x^3}{3} + \frac{x^5}{5} + o(x^5)$
$\arcsin(x)$	$\sum_{k=0}^{n} \frac{(2k)!}{(2k+1)(2^k.k!)^2} x^{2k+1} + o(x^{2n+2})$	$x + \frac{x^3}{6} + \frac{3.x^5}{40} + o(x^5)$

EXERCICES

Vrai ou faux ?

		Vrai	Faux
a)	$\ln(x^2+x) \underset{+\infty}{\sim} 2\ln(x)$.	☐	☐
b)	$\exp(x^2+x) \underset{+\infty}{\sim} \exp(x^2)$.	☐	☐
c)	$\sin\left(\dfrac{1}{x+2}\right) \underset{+\infty}{\sim} \dfrac{1}{x}$.	☐	☐
d)	$\sin\left(\dfrac{1}{x+2}\right) \underset{+\infty}{\sim} \dfrac{1}{x}+\dfrac{1}{x^3}$.	☐	☐
e)	$\cos\left(\dfrac{1}{x+2}\right) \underset{+\infty}{\sim} \dfrac{1}{x}$.	☐	☐
f)	Si $f(x) \underset{+\infty}{\sim} \ln(x)$, alors $f(x) \underset{+\infty}{=} \ln(x)+O(1)$.	☐	☐
g)	Deux fonctions qui admettent la même limite finie en a sont équivalentes en a.	☐	☐
h)	Deux fonctions équivalentes en a et qui admettent des limites en a, admettent la même limite.	☐	☐
i)	$\operatorname{sh} x \underset{+\infty}{\sim} \operatorname{ch} x$.	☐	☐
j)	La partie régulière du $\mathrm{DL}_n(0)$ d'une fonction paire est paire.	☐	☐
k)	La partie régulière du $\mathrm{DL}_n(0)$ d'une fonction périodique est périodique.	☐	☐
l)	Une fonction dérivable en 0 admet un $\mathrm{DL}_1(0)$.	☐	☐
m)	Une fonction qui admet un $\mathrm{DL}_2(0)$ est deux fois dérivable en 0.	☐	☐
n)	Si f est deux fois dérivable en 0 et $f(x) \underset{0}{=} 1+ax+bx^2+o(x^2)$, alors $f''(0)=b$.	☐	☐
o)	Si f admet un $\mathrm{DL}_n(0)$, alors $f(x^2)$ admet un $\mathrm{DL}_{2n}(0)$.	☐	☐
p)	$\ln(1+x) \underset{0}{=} \sum_{k=1}^{n} \dfrac{(-1)^k}{k}x^k+o(x^n)$.	☐	☐

Exercices d'application du chapitre 9

●○○ **Exercice 1**

Soit, pour tout $n \in \mathbb{N}$, $u_n = \cos(\pi\sqrt{n^2+n+2})$. Déterminer un équivalent de u_n.

●○○ **Exercice 2**

Déterminer la limite de la suite (u_n) définie par $u_n = \left(e - \left(1+\frac{1}{n}\right)^n\right)^{\sqrt{n^2-1}-n}$.

●●○ **Exercice 3**

1. Montrer que, pour tout $n \in \mathbb{N} \setminus \{0\}$, l'équation $e^x = n - x$ admet une unique solution positive que l'on notera x_n.
2. Déterminer deux termes du développement asymptotique de x_n.

●●○ **Exercice 4**

Soit (u_n) une suite réelle telle que, pour tout $n \in \mathbb{N}$,

$$u_n^5 + n\,u_n = 1.$$

Étudier cette suite et donner un développement asymptotique à deux termes de u_n.

●○○ **Exercice 5**

Calculer le DL en 0 des expressions suivantes

1. $\dfrac{\ln(1+x)}{\sin x}$ à l'ordre 3,
2. $\int_0^x \frac{\sin t}{t}\,\mathrm{d}t$ à l'ordre 5.

●●○ **Exercice 6**

Calculer le coefficient devant $(x-1)^{1789}$ dans le DL en 1 de $x \mapsto \ln(1+\frac{1}{x})$.

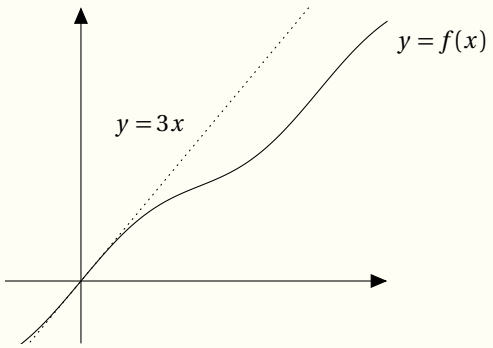
●●○ **Exercice 7**

1. Montrer que la fonction $f : x \mapsto x e^{x^2}$ réalise une bijection de $\mathbb{R}$ dans $\mathbb{R}$.
2. Déterminer un développement limité à l'ordre 5 en 0 de f^{-1}.

●●○ **Exercice 8**

Soit f la fonction de $\mathbb{R}$ dans $\mathbb{R}$ définie par $f(x) = 2x + \sin x$.

1. Montrer que f est une bijection de classe $\mathscr{C}^\infty$ de $\mathbb{R}$ dans $\mathbb{R}$.
2. Calculer un DL à l'ordre 3 en 0 de f^{-1}.



EXERCICES

●●● **Exercice 9**

Déterminer un équivalent en $+\infty$ de $\int_2^x \frac{\mathrm{d}t}{\ln t}$.

●○○ **Exercice 10**

Soit $f : \mathbb{R} \to \mathbb{R}$ de classe $\mathscr{C}^3$. Calculer, pour chaque $x \in \mathbb{R}$,

$$\lim_{h \to 0} \frac{f(x+3h) - 3f(x+2h) + 3f(x+h) - f(x)}{h^3}.$$

●●○ **Exercice 11**

Déterminer une condition nécessaire et suffisante sur $a_1, \dots, a_n \in \mathbb{R}$ pour que $f : x \mapsto \sum_{k=1}^n \frac{a_k}{\tan(kx)}$ admette une limite finie en 0.

●○○ **Exercice 12**

Soit $f : \mathbb{R} \to \mathbb{R}_+^*$ de classe $\mathscr{C}^2$ telle que $f(0) = 1$, $f'(0) = 0$ et $f''(0) = -1$. Montrer que

$$\forall a \in \mathbb{R}, \qquad \lim_{x \to +\infty} f\left(\frac{a}{\sqrt{x}}\right)^x = e^{-\frac{a^2}{2}}.$$

●○○ **Exercice 13**

Déterminer un développement asymptotique de $\arctan x$ en $+\infty$ à la précision $o(\frac{1}{x^4})$.

●●○ **Exercice 14**

Déterminer l'asymptote en $+\infty$ et étudier la position relative de la courbe par rapport à l'asymptote pour chacune des courbes données par les équations suivantes :

1. $y = \sqrt{x(x+1)}$,
2. $y = x e^{\frac{1}{x}} \arctan x$.

Exercices d'approfondissement du chapitre 9

●●● **Exercice A**

Montrer qu'il n'existe pas de réel $\alpha > 0$ tel que $\lfloor \log(x) \rfloor ! \underset{+\infty}{=} o(x^\alpha)$.

●○○ **Exercice B**

Déterminer un équivalent de $\arccos x$ en 1.

●○○ **Exercice C**

Soit f continue sur $\mathbb{R}$ telle que $\lim\limits_{\pm\infty} f = 0$ et

$$g : \left\{ \begin{array}{ccc}]-1,1[& \to & \mathbb{R} \\ x & \mapsto & f\left(\frac{x}{1-|x|}\right) \end{array} \right.$$

1. Montrer que g est continue sur $]-1,1[$ et prolongeable par continuité en une fonction encore notée g sur $[-1,1]$.
2. Montrer que g est dérivable en 1 avec $g'(1) = 0$ si, et seulement si, $f(x) \underset{+\infty}{=} o(\frac{1}{x})$.

●●○ **Exercice D**

Déterminer le DL en 0 à l'ordre 1000 de $x \mapsto \ln\left(\sum_{k=0}^{999} \frac{x^k}{k!}\right)$.

●●○ **Exercice E**

1. Établir une relation entre les expressions $\arcsin\sqrt{x}$ et $\frac{\pi}{4}+\frac{1}{2}\arcsin(2x-1)$.
2. En déduire les trois premiers termes du développement asymptotique de $\arcsin x$ en -1.

Problème du chapitre 9

●○○ **Problème 1 Théorème de Glaeser**

Dans ce court problème, on cherche une condition nécessaire et suffisante pour qu'une fonction de classe $\mathscr{C}^2$ et positive admettent une racine de classe $\mathscr{C}^1$.

1. Exhiber une fonction $f : \mathbb{R} \to \mathbb{R}_+$ de classe $\mathscr{C}^2$ telle que la fonction $\sqrt{f}$ ne soit pas de classe $\mathscr{C}^1$ sur $\mathbb{R}$.
2. Posons, dans cette question, $f : x \mapsto (\sin x)^2$.
   a. Montrer que la fonction f est positive et de classe $\mathscr{C}^2$.
   b. La fonction $\sqrt{f}$ est-elle de classe $\mathscr{C}^1$ sur $\mathbb{R}$?
3. Soit $f : \mathbb{R} \to \mathbb{R}_+$ une fonction de classe $\mathscr{C}^2$.
   a. Soit $x_0 \in \mathbb{R}$ tel que $f(x_0) \neq 0$. Montrer que la fonction $\sqrt{f}$ est dérivable en x_0.
   b. Soit $x_0 \in \mathbb{R}$ tel que $f(x_0) = 0$. En utilisant la formule de Taylor-Young, montrer que $\sqrt{f}$ admet des dérivées à droite et à gauche en x_0. Préciser les valeurs de ces dérivées à droite et à gauche en x_0.
   c. En déduire une condition nécessaire et suffisante pour que la fonction $\sqrt{f}$ soit dérivable sur $\mathbb{R}$.
4. Soit $f : \mathbb{R} \to \mathbb{R}_+$ une fonction de classe $\mathscr{C}^2$.
   a. Soit $x_0 \in \mathbb{R}$ tel que $f(x_0) = f''(x_0) = 0$. Pour tout $\eta > 0$, notons
   $$M_2(\eta) = \sup\{|f''(x)|, x \in [x_0 - 2\eta, x_0 + 2\eta]\}.$$
   On admet la formule de Taylor-Lagrange sous la forme suivante : pour tout $x \in [x_0 - \eta, x_0 + \eta]$ et tout $|h| < \eta$,
   $$|f(x+h) - f(x) - hf'(x)| \leq \frac{h^2}{2} M_2(\eta).$$
      i. Justifier que, pour tout $x \in [x_0 - \eta, x_0 + \eta]$ et tout $|h| < \eta$,
      $$f(x) + hf'(x) + \frac{h^2}{2} M_2(\eta) \geq 0.$$
      ii. Montrer que, pour tout $x \in [x_0 - \eta, x_0 + \eta]$, $|f'(x)| \leq \sqrt{2f(x)M_2(\eta)}$.
   b. En déduire une condition nécessaire et suffisante pour que la fonction $\sqrt{f}$ est $\mathscr{C}^1$ sur $\mathbb{R}$.

PROBLÈME

CORRIGÉS

Corrigés des Vrai/Faux

a) Vrai.

b) Faux.

c) Vrai, $\sin\left(\dfrac{1}{x+2}\right) \underset{+\infty}{\sim} \dfrac{1}{x+2} \sim \dfrac{1}{x}$.

d) Vrai, car $\dfrac{1}{x} + \dfrac{1}{x^3} \underset{+\infty}{\sim} \frac{1}{x}$.

e) Faux, car le membre de gauche tend vers 1 alors que celui de droite tend vers 0.

f) Faux, par exemple avec $f(x) = \ln(x) + \sqrt{\ln(x)}$.

g) Faux dans le cas d'une limite nulle : par exemple, les suites $x \mapsto \frac{1}{x}$ et $x \mapsto \frac{1}{x^2}$ ne sont pas équivalentes mais admettent la même limite.

h) Vrai.

i) Vrai, $\operatorname{sh} x \underset{+\infty}{\sim} \frac{1}{2} e^x$.

j) Vrai.

k) Faux, les seuls polynômes périodiques sont les polynômes constants et la fonction sin admet une partie régulière non constante dans son DL à l'ordre 1 en 0.

l) Vrai.

m) Faux, $x \mapsto x^3 \sin \frac{1}{x}$.

n) Faux, d'après la formule de Taylor-Young, $b = \frac{1}{2} f''(0)$.

o) Vrai.

p) Faux, l'exposant de (-1) est faux.

Corrigés des exercices d'application du chapitre 9

Exercice 1

Commençons par effectuer un développement asymptotique de l'argument du cosinus. Pour tout $n \in \mathbb{N} \setminus \{0\}$,

$$\begin{aligned}\sqrt{n^2+n+2} &= n\sqrt{1+\frac{1}{n}+\frac{2}{n^2}} = n\left(1+\frac{1}{2}\left(\frac{1}{n}+\frac{2}{n^2}\right)-\frac{1}{8}\left(\frac{1}{n}+\frac{2}{n^2}\right)^2+o\left(\frac{1}{n^2}\right)\right)\\ &= n\left(1+\frac{1}{2n}+\frac{7}{8n^2}+o\left(\frac{1}{n^2}\right)\right) = n+\frac{1}{2}+\frac{7}{8n}+o\left(\frac{1}{n}\right).\end{aligned}$$

Par conséquent,

$$\begin{aligned} u_n &= \cos(n\pi + \frac{\pi}{2} + \frac{7\pi}{8n} + o\left(\frac{1}{n}\right)) \\ &= (-1)^{n+1}\sin(\frac{7\pi}{8n} + o\left(\frac{1}{n}\right)) \\ &\sim (-1)^{n+1}\frac{7\pi}{8n}. \end{aligned}$$

Exercice 2

Commençons par développer le terme entre parenthèses. Pour tout $n \in \mathbb{N} \setminus \{0\}$,

$$\begin{aligned} e - \left(1+\frac{1}{n}\right)^n &= e - e^{n\ln\left(1+\frac{1}{n}\right)} \\ &= e - e^{n\left(\frac{1}{n} - \frac{1}{2n^2} + o\left(\frac{1}{n^2}\right)\right)} \\ &= e\left(1 - e^{-\frac{1}{2n} + o\left(\frac{1}{n}\right)}\right) \\ &= \frac{e}{2n} + o(\frac{1}{n}). \end{aligned}$$

Par ailleurs, l'exposant vérifie

$$\sqrt{n^2-1} - n = n\left(\sqrt{1-\frac{1}{n^2}} - 1\right) = -\frac{1}{2n} + o(\frac{1}{n}).$$

En combinant ces deux résultats, on obtient $u_n = e^{\left(-\frac{1}{2n} + o\left(\frac{1}{n}\right)\right)\ln\left(\frac{e}{2n} + o\left(\frac{1}{n}\right)\right)}$.
Or, $\lim \frac{1}{n}\ln\frac{1}{n} = 0$ d'où, par continuité de la fonction exp, $\lim u_n = 1$.

Exercice 3

1. Considérons la fonction $f : x \mapsto x + e^x$. Elle est de classe $\mathscr{C}^\infty$ et à dérivée, $x \mapsto 1 + e^x$, strictement positive : elle réalise donc une bijection de $\mathbb{R}_+$ dans $f(\mathbb{R}_+) = [1, +\infty[$ (ce dernier point provient de $f(0) = 1$ et $\lim\limits_{+\infty} f = +\infty$). Ainsi, il existe un unique antécédent dans $\mathbb{R}_+$ à l'élément n.
2. Remarquons tout d'abord que, pour tout $n \in \mathbb{N}$, $f(x_n) = n < n+1 = f(x_{n+1})$ donc par stricte croissance de f puis de f^{-1}, $x_n < x_{n+1}$. Ainsi, la suite $(x_n)_n$ est strictement croissante.
▹ Si elle admet une limite $\ell \in \mathbb{R}$, alors par continuité de f, $n = f(x_n) \to f(\ell)$ ce qui est absurde. La suite $(x_n)_n$ est donc croissante et divergente donc tend vers $+\infty$.
▹ Ainsi, $x_n = o(e^{x_n})$ et donc $n \sim e^{x_n}$. Écrivons alors

$$x_n = \ln(e^{x_n}) = \ln(n + o(n)) = \ln n + \ln(1 + o(1)) = \ln n + o(1).$$

▹ Pour calculer le terme suivant, posons $y_n = x_n - \ln n$. D'après le point précédent, $y_n = o(1)$. Or, la relation définissant x_n se réécrit en remplaçant

$$n e^{y_n} + y_n + \ln n = n,$$

soit encore

$$n + n y_n + n o(y_n) + y_n + \ln n = n.$$

En simplifiant, on obtient $y_n + o(y_n) = -\frac{\ln n}{n}$, c'est-à-dire $y_n \sim -\frac{\ln n}{n}$. Par conséquent,

$$x_n = \ln n - \frac{\ln n}{n} + o(\frac{\ln n}{n}).$$

En reportant, on pourrait obtenir les termes suivants de la même manière mais avec une technicité croissante dans les calculs.

Exercice 4

Procédons par étapes.

▹ Montrons tout d'abord que la suite (u_n) est bien définie. Pour cela, étudions, pour tout $n \in \mathbb{N}$, la fonction $f_n : x \mapsto x^5 + nx$: elle est continue, strictement croissante de limite $-\infty$ en $-\infty$ et $+\infty$ en $+\infty$ donc réalise une bijection de $\mathbb{R}$ dans $\mathbb{R}$. Par conséquent, il existe un unique réel u_n tel que $f_n(u_n) = 1$. De plus, $f_n(0) = 0$ donc $u_n > 0$.

▹ Étudions ensuite les variations de (u_n). Pour tout $n \in \mathbb{N}$,

$$\begin{aligned} f_n(u_{n+1}) &= u_{n+1}^5 + n\,u_{n+1} = 1-(n+1)u_{n+1} + n\,u_{n+1} \\ &= 1 - u_{n+1} < 1 = f_n(u_n). \end{aligned}$$

Comme les fonctions f_n sont strictement croissantes, on en déduit que $u_{n+1} < u_n$ pour tout $n \in \mathbb{N}$: la suite (u_n) est strictement décroissante.

▹ Comme la suite (u_n) est décroissante, minorée par 0, elle converge vers une limite $\ell \geq 0$. Si $\ell > 0$, alors $\lim n\,u_n = +\infty$ et $\lim 1-u_n^5 = 1-\ell^5$; ces deux limites devraient être égales donc on tient une contradiction et $\ell = 0$.

▹ Passons enfin à l'étude asymptotique.

Remarquons que $n\,u_n = 1 - u_n^5 \sim 1$ car $\lim u_n = 0$ donc $u_n \sim \frac{1}{n}$. Pour étudier le terme suivant, posons $v_n = u_n - \frac{1}{n}$ pour tout $n \in \mathbb{N} \setminus \{0\}$. La relation de récurrence sur (u_n) devient, pour tout $n \in \mathbb{N}$,

$$\Big(v_n + \frac{1}{n}\Big)^5 + n\Big(v_n + \frac{1}{n}\Big) = 1.$$

D'où $n\,v_n + \frac{1}{n^5} + o\big(\frac{1}{n^5}\big) = 0$ en utilisant que $v_n = o\big(\frac{1}{n}\big)$; ainsi $v_n \sim -\frac{1}{n^6}$.

En conclusion,

$$u_n = \frac{1}{n} - \frac{1}{n^6} + o\Big(\frac{1}{n^6}\Big).$$

Exercice 5

1. D'après les développements usuels,

$$\begin{aligned} \frac{\ln(1+x)}{\sin x} &\underset{0}{=} \frac{x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \frac{1}{4}x^4 + o(x^4)}{x - \frac{1}{6}x^3 + o(x^4)} \\ &\underset{0}{=} \frac{1 - \frac{1}{2}x + \frac{1}{3}x^2 - \frac{1}{4}x^3 + o(x^3)}{1 - \frac{1}{6}x^2 + o(x^3)} \\ &\underset{0}{=} \Big(1 - \frac{1}{2}x + \frac{1}{3}x^2 - \frac{1}{4}x^3 + o(x^3)\Big)\Big(1 + \frac{1}{6}x^2 + o(x^3)\Big) \\ &\underset{0}{=} 1 - \frac{1}{2}x + \frac{1}{2}x^2 - \frac{1}{3}x^3 + o(x^3). \end{aligned}$$

On remarque qu'il a fallu effectuer le calcul à l'ordre 4 pour que le résultat final soit à l'ordre 3.

2. Commençons par rappeler que

$$\frac{\sin t}{t} \underset{0}{=} 1 - \frac{1}{6}t^2 + \frac{1}{120}t^4 + o(t^4).$$

Par intégration des développements limités,

$$\int_0^x \frac{\sin t}{t}\,\mathrm{d}t \underset{0}{=} 0 + x - \frac{1}{18}x^3 + \frac{1}{600}x^5 + o(x^5).$$

Exercice 6

Écrivons $x = 1+h$. Alors,

$$\begin{aligned}\ln(1+\frac{1}{x}) &= \ln(2+h)-\ln(1+h)\\ &= \ln 2+\ln(1+\frac{h}{2})-\ln(1+h)\\ &\underset{h\to 0}{=} \ln 2+\sum_{k=1}^{n}\frac{(-1)^{k+1}}{k}2^{-k}h^k-\sum_{k=1}^{n}\frac{(-1)^{k+1}}{k}h^k+o(h^n)\\ &\underset{x\to 1}{=} \ln 2+\sum_{k=1}^{n}\frac{(-1)^{k+1}}{k}(2^{-k}-1)(x-1)^k+o((x-1)^n).\end{aligned}$$

Le coefficient recherché est donc $\frac{1}{1789}\left(2^{-1789}-1\right)$.

Exercice 7

1. La fonction f est dérivable sur $\mathbb{R}$ et sa dérivée est $x \mapsto (2x^2+1)e^{x^2}$. En particulier, f' est strictement positive donc f réalise une bijection de $\mathbb{R}$ dans $f(\mathbb{R})$. Comme $\lim\limits_{-\infty} f = -\infty$ et $\lim\limits_{+\infty} f = +\infty$, $f(\mathbb{R}) = \mathbb{R}$ et donc f réalise une bijection de $\mathbb{R}$ dans $\mathbb{R}$.
2. Notons que f est de classe $\mathscr{C}^\infty$ comme composée et produit de fonctions de classe $\mathscr{C}^\infty$ et sa dérivée ne s'annule pas ; par conséquent, f^{-1} est de classe $\mathscr{C}^\infty$ et admet des DL à tout ordre en 0 avec la formule de Taylor-Young. De plus, f est impaire donc f^{-1} est aussi impaire. Notons $ax+bx^3+cx^5$ la partie régulière de DL à l'ordre 5 de f^{-1}. Remarquons enfin que

$$f(x) \underset{0}{=} x(1+x^2+\frac{x^4}{2}+o(x^4)) = x+x^3+\frac{1}{2}x^5+o(x^5).$$

En écrivant que $x = f^{-1}(f(x))$, on obtient

$$\begin{aligned}x &\underset{0}{=} a\left(x+x^3+\frac{1}{2}x^5\right)+b(x^3+3x^5)+cx^5+o(x^5)\\ &\underset{0}{=} ax+(a+b)x^3+\left(\frac{a}{2}+3b+c\right)x^5+o(x^5).\end{aligned}$$

En identifiant les parties régulières, on conclut que $a=1$, $b=-1$ et $c=\frac{5}{2}$ donc

$$f^{-1}(x) \underset{0}{=} x-x^3+\frac{5}{2}x^5+o(x^5).$$

Exercice 8

1. La fonction f est de classe $\mathscr{C}^\infty$ et sa dérivée, $x \mapsto 2+\cos x$, est strictement positive. Par conséquent, f réalise une bijection de $\mathbb{R}$ dans $f(\mathbb{R}) = \mathbb{R}$ (car $\lim\limits_{-\infty} f = -\infty$ et $\lim\limits_{+\infty} f = +\infty$).
2. D'après la première question, f^{-1} est de classe $\mathscr{C}^\infty$ (car la dérivée ne s'annule pas) donc admet des DL de tout ordre en 0. Comme par ailleurs, f est impaire, f^{-1} l'est aussi. Notons $f^{-1}(x) \underset{0}{=} ax+bx^3+o(x^3)$ le DL à l'ordre 3 en 0 de f^{-1}. Comme $f(x) \underset{0}{=} 3x-\frac{x^3}{6}+o(x^3)$, la relation $f^{-1}(f(x)) = x$ entraîne

$$\begin{aligned}x &\underset{0}{=} a\left(3x-\frac{x^3}{6}+o(x^3)\right)+b\left(3x-\frac{x^3}{6}+o(x^3)\right)^3+o(x^3)\\ &\underset{0}{=} 3ax+\left(-\frac{a}{6}+27b\right)x^3+o(x^3).\end{aligned}$$

En identifiant les parties régulières, on trouve $a=\frac{1}{3}$ et $b=\frac{1}{486}$, donc

$$f^{-1}(x) \underset{0}{=} \frac{1}{3}x+\frac{1}{486}x^3+o(x^3).$$

Exercice 9

Procédons à une intégration par parties (en intégrant la constante 1). Pour tout $x > 2$,

$$\int_2^x \frac{\mathrm{d}t}{\ln t} = \left[\frac{t}{\ln t}\right]_2^x + \int_2^x \frac{\mathrm{d}t}{(\ln t)^2}.$$

▹ Comme $\ln t \leq t$, $\int_2^x \frac{\mathrm{d}t}{\ln t} \geq \int_2^x \frac{\mathrm{d}t}{t} = \ln x - \ln 2$. Ainsi, $\int_2^x \frac{\mathrm{d}t}{\ln t} \to +\infty$.

▹ Montrons que $\int_2^x \frac{\mathrm{d}t}{(\ln t)^2} = o\left(\int_2^x \frac{\mathrm{d}t}{\ln t}\right)$. Soit $\varepsilon > 0$ et A tel que $\frac{1}{\ln A} \leq \varepsilon$ Alors, pour $x \geq A$,

$$\int_A^x \frac{\mathrm{d}t}{(\ln t)^2} \leq \frac{1}{\ln A}\int_A^x \frac{\mathrm{d}t}{\ln t} \leq \varepsilon \int_2^x \frac{\mathrm{d}t}{\ln t}.$$

Par ailleurs, pour x suffisamment grand

$$\int_2^A \frac{\mathrm{d}t}{(\ln t)^2} \leq \varepsilon \int_2^x \frac{\mathrm{d}t}{\ln t},$$

le premier membre est une constante et le second tend vers $+\infty$ en $+\infty$. Nous avons établi que pour tout $\varepsilon > 0$, il existe A tel que pour tout $x \geq A$,

$$\int_2^x \frac{\mathrm{d}t}{(\ln t)^2} \leq 2\varepsilon \int_2^x \frac{\mathrm{d}t}{\ln t},$$

c'est-à-dire, avec la définition de limite, $\int_2^x \frac{\mathrm{d}t}{(\ln t)^2} = o\left(\int_2^x \frac{\mathrm{d}t}{\ln t}\right)$.
En conclusion,

$$\int_2^x \frac{\mathrm{d}t}{\ln t} \underset{+\infty}{\sim} \frac{x}{\ln x}.$$

Exercice 10

Appliquons la formule de Taylor-Young à f en x à l'ordre 3 pour estimer la quantité $\Delta(h) = f(x+3h) - 3f(x+2h) + 3f(x+h) - f(x)$. Alors, pour $x \in \mathbb{R}$,

$$\begin{aligned}
\Delta(h) &\underset{h\to 0}{=} (f(x) + 3f'(x)h + \frac{9}{2}f''(x)h^2 + \frac{9}{2}f'''(x)h^3) \\
&\quad -3(f(x) + 2f'(x)h + 2f''(x)h^2 + \frac{4}{3}f'''(x)h^3) \\
&\quad +3(f(x) + f'(x)h + \frac{1}{2}f''(x)h^3 + \frac{1}{6}f'''(x)h^3) - f(x) + o(h^3) \\
&\underset{h\to 0}{=} f'''(x)h^3 + o(h^3).
\end{aligned}$$

Par conséquent, $\lim\limits_{h\to 0} \frac{\Delta(h)}{h^3} = f'''(x)$.

Exercice 11

Fixons k et étudions $x \mapsto \frac{1}{\tan kx}$ au voisinage de 0.

$$\frac{1}{\tan kx} \underset{0}{=} \frac{1}{kx + o(x^2)} = \frac{1}{kx}\frac{1}{1+o(x)} \underset{0}{=} \frac{1}{kx}(1+o(x)) = \frac{1}{kx} + o(1).$$

Par conséquent,

$$f(x) \underset{0}{=} \left(\sum_{k=1}^{n} \frac{a_k}{k}\right)\frac{1}{x} + o(1).$$

Ainsi, f admet une limite finie en 0 si, et seulement si, $\sum_{k=1}^{n} \frac{a_k}{k} = 0$.

Exercice 12

D'après la formule de Taylor-Young appliquée à f à l'ordre 2 en 0, $f(x) \underset{0}{=} 1 - \frac{1}{2}x^2 + o(x^2)$. Par conséquent, pour $a \in \mathbb{R}$,

$$\begin{aligned} f\left(\frac{a}{\sqrt{x}}\right)^x &= \exp(x \ln f\left(\frac{a}{\sqrt{x}}\right)) \\ &\underset{+\infty}{=} \exp(x \ln\left(1 - \frac{a^2}{2x} + o\left(\frac{1}{x}\right)\right)) \\ &\underset{+\infty}{=} \exp(x\left(-\frac{a^2}{2x} + o\left(\frac{1}{x}\right)\right)) \\ &\underset{+\infty}{=} \exp(-\frac{a^2}{2} + o(1)). \end{aligned}$$

En conclusion, $\lim_{x\to+\infty} f\left(\frac{a}{\sqrt{x}}\right)^x = e^{-\frac{a^2}{2}}$.

Exercice 13

Commençons par rappeler que pour tout $x > 0$, $\arctan x = \frac{\pi}{2} - \arctan \frac{1}{x}$. Cette formule permet ainsi de ramener l'étude de $\arctan x$ quand x est au voisinage de $+\infty$ à celle de $\arctan \frac{1}{x}$ avec $\frac{1}{x}$ au voisinage de 0. Or, $\arctan u \underset{0}{=} u - \frac{1}{3}u^3 + o(u^4)$. En remplaçant, on obtient donc $\arctan x = \frac{\pi}{2} - \frac{1}{x} + \frac{1}{3x^3} + o\left(\frac{1}{x^4}\right)$.

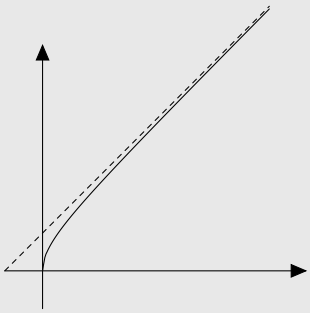
Exercice 14

Pour chaque exemple, posons $x = \frac{1}{t}$ puis effectuons un DL en $t = 0$.

1. En mettant en facteur $\frac{1}{t^2}$ puis en utilisant le développement en o de $u \mapsto \sqrt{1+u}$, on obtient

$$\begin{aligned} \sqrt{x(x+1)} &= \left(\frac{1}{t^2} + \frac{1}{t}\right)^{\frac{1}{2}} \\ &\underset{t\to 0}{=} \frac{1}{t}(1 + \frac{1}{2}t - \frac{1}{8}t^2 + o(t^2)) \\ &\underset{t\to 0}{=} \frac{1}{t} + \frac{1}{2} - \frac{1}{8}t + o(t) \\ &\underset{x\to+\infty}{=} x + \frac{1}{2} - \frac{1}{8}\frac{1}{x} + o\left(\frac{1}{x}\right). \end{aligned}$$

L'asymptote recherchée admet donc pour équation $y = x + \frac{1}{2}$, l'écart à cette asymptote est équivalent à $-\frac{1}{8}\frac{1}{x}$ donc négatif au voisinage de $+\infty$: la courbe est sous l'asymptote.



2. Utilisons que pour tout $u>0$, $\arctan\frac{1}{u}=\frac{\pi}{2}-\arctan u$:

$$\begin{aligned}
x\,e^{\frac{1}{x}}\arctan x &= \frac{1}{t}e^{t}\arctan\frac{1}{t}=\frac{1}{t}e^{t}\Big(\frac{\pi}{2}-\arctan t\Big)\\
&\underset{t\to 0}{=}\frac{1}{t}\Big(1+t+\frac{1}{2}t^2+o(t^2)\Big)\Big(\frac{\pi}{2}-t+o(t^2)\Big)\\
&\underset{t\to 0}{=}\frac{\pi}{2t}+\left(\frac{\pi}{2}-1\right)+\left(\frac{\pi}{4}-1\right)t+o(t)\\
&\underset{x\to+\infty}{=}\frac{\pi}{2}x+\left(\frac{\pi}{2}-1\right)+\left(\frac{\pi}{4}-1\right)\frac{1}{x}+o\Big(\frac{1}{x}\Big).
\end{aligned}$$

L'asymptote recherchée admet donc pour équation $y=\frac{\pi}{2}x+\left(\frac{\pi}{2}-1\right)$, l'écart à cette asymptote est équivalent à $\left(\frac{\pi}{4}-1\right)\frac{1}{x}$ donc négatif au voisinage de $+\infty$: la courbe est sous l'asymptote.

Corrigés des exercices d'approfondissement du chapitre 9

Exercice A

Supposons qu'il existe $\alpha>0$ tel que $|\log(x)|!\underset{+\infty}{=}o(x^{\alpha})$. Alors, en considérant pour x les entiers de la forme 10^n, on obtient que $\frac{n!}{10^{\alpha n}}\to 0$ ce qui est absurde. En effet, en posant $u_n=\frac{n!}{10^{\alpha n}}$, on obtient que $\frac{u_{n+1}}{u_n}=\frac{n+1}{10^{\alpha}}\to+\infty$ donc $u_n\to+\infty$ d'après le critère de D'Alembert.

Exercice B

Remarquons tout d'abord que la fonction arccos est continue en 1 (et y prend la valeur 0) mais qu'elle n'y est pas dérivable.
Utilisons que $\arccos x\to 0$ quand $x\to 1$ pour écrire $\arccos x\underset{1}{\sim}\sin(\arccos x)$. Comme $\arccos x\in[0,\pi]$, $\sin(\arccos x)\geq 0$ donc $\sin(\arccos x)=\sqrt{1-\cos^2(\arccos x)}=\sqrt{1-x^2}$. En conclusion, $\arccos x\underset{1}{\sim}\sqrt{1-x^2}$ soit encore $\arccos x\underset{1}{\sim}2\sqrt{1-x}$ en utilisant que $\sqrt{1+x}\underset{1}{\sim}2$.

Exercice C

1. La fonction g est continue sur $]-1,1[$ et, d'après les limites sur f, g tend vers 0 en ± 1, d'où le prolongement par continuité en posant $g(\pm 1)=0$.
2. Pour tout $x\in|0,1[$,

$$\frac{g(x)-g(1)}{x-1}=-\frac{x}{1-x}f\left(\frac{x}{1-|x|}\right)-f\left(\frac{x}{1-x}\right).$$

Par conséquent, g est dérivable en 1 avec $g'(1) = 0$ si, et seulement si, $\lim\limits_{x\to 1} \frac{x}{1-x} f\left(\frac{x}{1-x}\right) = 0$. Comme la fonction $\varphi : x \mapsto \frac{x}{1-x}$ est une bijection continue de $[0,1[$ dans $\mathbb{R}_+$ qui vérifie $\lim\limits_{1^-} \varphi = +\infty$, la condition sur f est $\lim\limits_{t\to\infty} t f(t) = 0$ ce qui est bien le résultat recherché.

Exercice D

Remarquons que l'argument du logarithme est la partie régulière du DL à l'ordre 999 en 0 de l'exponentielle. Par conséquent,

$$\begin{aligned}
\ln\left(\sum_{k=0}^{999} \frac{x^k}{k!}\right) &\underset{0}{=} \ln(e^x - \frac{x^{1000}}{1000!} + o(x^{1000})) \\
&\underset{0}{=} x + \ln(1 + e^{-x}\frac{x^{1000}}{1000!} + o(e^{-x}x^{1000})) \\
&\underset{0}{=} x + \ln(1 + (1+o(1))\frac{x^{1000}}{1000!} + o(x^{1000})) \\
&\underset{0}{=} x + \ln(1 + \frac{x^{1000}}{1000!} + o(x^{1000})) \\
&\underset{0}{=} x - \frac{x^{1000}}{1000!} + o(x^{1000}).
\end{aligned}$$

Exercice E

1. Les deux expressions définissent des fonctions dérivables de dérivées $x \mapsto \frac{1}{2\sqrt{x(1-x)}}$ et nulles en 0 : elles sont donc égales.
2. Pour calculer ce développement asymptotique, posons $x = -1 + h$ puis utilisons la question précédente et le DL de arcsin en 0

$$\begin{aligned}
\arcsin x = \arcsin(-1+h) &= -\frac{\pi}{2} + 2\arcsin\sqrt{\frac{h}{2}} \\
&\underset{h\to 0}{=} -\frac{\pi}{2} + 2\left(\sqrt{\frac{h}{2}} + \frac{1}{6}\sqrt{\frac{h}{2}}^3 + o(h^{\frac{3}{2}})\right) \\
&\underset{x\to -1}{=} -\frac{\pi}{2} + \sqrt{2}(x+1)^{\frac{1}{2}} + \frac{1}{6\sqrt{2}}(x+1)^{\frac{3}{2}} + o((x+1)^{\frac{3}{2}}).
\end{aligned}$$

Corrigés des problèmes du chapitre 9

Problème 1

1. Par exemple, la fonction $f : x \mapsto x^2$ convient car elle est de classe $\mathscr{C}^\infty$ et $\sqrt{f} : x \mapsto |x|$ n'est pas dérivable en 0.
2. a. La fonction f est positive comme carré d'une fonction réelle et de classe $\mathscr{C}^2$ comme composées de fonctions de classe $\mathscr{C}^2$.
b. La fonction $\sqrt{f}$ n'est pas de classe $\mathscr{C}^1$ sur $\mathbb{R}$, car elle ne l'est pas pour les éléments de $\pi\mathbb{Z}$ (où elle s'annule). En effet, $\sqrt{f}$ est π-périodique et

$$\lim_{x\to 0^-}\left(\sqrt{f(x)}\right)' = \lim_{x\to 0^-} -\cos x = -1, \qquad \lim_{x\to 0^+}\left(\sqrt{f(x)}\right)' = \lim_{x\to 0^+} \cos x = 1.$$

3. a. La fonction f est de classe $\mathscr{C}^1$ sur $\mathbb{R}$ donc en x_0. La fonction $y \mapsto \sqrt{y}$ est de classe $\mathscr{C}^1$ sur $\mathbb{R}_+^*$ donc, par composition (car $f(x_0) \in \mathbb{R}_+^*$), la fonction $\sqrt{f}$ est de classe $\mathscr{C}^1$ en x_0.
b. Comme f est de classe $\mathscr{C}^2$, la formule de Taylor-Young à l'ordre 2 en x_0 donne

$$f(x) \underset{x_0}{=} f(x_0) + f'(x_0)(x-x_0) + \frac{1}{2} f''(x_0)(x-x_0)^2 + o((x-x_0)^2).$$

Par ailleurs, comme f est positive et $f(x_0) = 0$, f admet un minimum en x_0 donc $f'(x_0) = 0$.
Par conséquent,

$$f(x) \underset{x_0}{\sim} \frac{1}{2} f''(x_0)(x-x_0)^2.$$

Ainsi, le taux d'accroissement en x_0 de $\sqrt{f}$ devient

$$\frac{\sqrt{f(x)} - \sqrt{f(x_0)}}{x-x_0} = \frac{\sqrt{f(x)}}{x-x_0} \underset{x_0}{\sim} \frac{|x-x_0|}{x-x_0} \sqrt{\frac{1}{2} f''(x_0)}$$

Ainsi, en x_0, la fonction $\sqrt{f}$ admet pour dérivées à droite $-\sqrt{\frac{1}{2} f''(x_0)}$ et à gauche $\sqrt{\frac{1}{2} f''(x_0)}$.
c. ▹ En $x_0 \notin f^{-1}(\{0\})$, $\sqrt{f}$ est de classe $\mathscr{C}^1$.
▹ En $x_0 \in f^{-1}(\{0\})$, $\sqrt{f}$ est dérivable en x_0 si, et seulement si, les dérivées à droite à gauche coïncident c'est-à-dire si, et seulement si, $f''(x_0) = 0$.
Une condition nécessaire et suffisante pour que la fonction $\sqrt{f}$ soit dérivable sur $\mathbb{R}$ est que, pour tout $x_0 \in f^{-1}(\{0\})$, $f''(x_0) = 0$, autrement dit f'' s'annule partout où f s'annule.
4. a. i. Pour tout $x \in [x_0 - \eta, x_0 + \eta]$ et tout $|h| \leq \eta$, la formule de Taylor-Lagrange donne

$$f(x) + h f'(x) + \frac{h^2}{2} M_2(\eta) \geq f(x+h) \geq 0.$$

ii. ▹ Remarquons tout d'abord que si $M_2(\eta) = 0$, la formule de Taylor-Lagrange entraîne que la fonction f est affine sur $[x_0 - 2\eta, x_0 + 2\eta]$ donc $f = 0$ puisque $f(x_0) = f'(x_0) = 0$.
▹ Supposons désormais $M_2(\eta) > 0$ et notons $P(h) = f(x) + h f'(x) + \frac{h^2}{2} M_2(\eta)$.
Ce polynôme admet son minimum en $-\frac{f'(x)}{M_2(\eta)}$; or, l'inégalité des accroissements finis pour la fonction f' de classe $\mathscr{C}^1$ donne

$$|f'(x)| = |f'(x) - f'(x_0)| \leq M_2(\eta)|x - x_0| \leq M_2(\eta)\eta.$$

Ainsi, P atteint son minimum en $-\frac{f'(x)}{M_2(\eta)} \in [-\eta, \eta]$. Or, P est positif sur cet intervalle donc en particulier en ce minimum; par conséquent, P est positif sur tout $\mathbb{R}$. D'après l'étude du signe d'un polynôme de degré 2, cela se traduit par la négativité du discriminant de P soit

$$\Delta = f'(x)^2 - 2f(x)M_2(\eta) \leq 0.$$

En conclusion, pour tout $x \in [x_0 - \eta, x_0 + \eta]$, $|f'(x)| \leq \sqrt{2f(x)M_2(\eta)}$.
b. Comme dans la question précédente, $\sqrt{f}$ est de classe $\mathscr{C}^1$ en tout $x_0 \notin f^{-1}(\{0\})$.
Montrons que f est de classe $\mathscr{C}^2$ en $x_0 \in f^{-1}(\{0\})$ si, et seulement si, $f''(x_0) = 0$.
($\Rightarrow$) On a déjà établi que si $\sqrt{f}$ est dérivable en $x_0 \in f^{-1}(\{0\})$, alors $f''(x_0) = 0$.
($\Leftarrow$) Soit $x_0 \in f^{-1}(\{0\})$ tel que $f''(x_0) = 0$. Alors, d'après la question précédente, pour tout $\eta > 0$ et $x \in [x_0 - \eta, x_0 + \eta]$, $|f'(x)| \leq \sqrt{2f(x)M_2(\eta)}$ donc

$$\frac{f'(x)}{\sqrt{f(x)}} \leq \sqrt{2M_2(\eta)}.$$

Comme $M_2(\eta)$ tend vers 0 lorsque η vers 0, on obtient que $\left(\sqrt{f}\right)'(x) = \frac{f'(x)}{\sqrt{f(x)}}$ tend vers $0 = \left(\sqrt{f}\right)'(x_0)$ lorsque x tend vers x_0 : la fonction $\sqrt{f}$ est de classe $\mathscr{C}^1$ en x_0.
Une condition nécessaire et suffisante pour que la fonction $\sqrt{f}$ soit $\mathscr{C}^1$ sur $\mathbb{R}$ est que, pour tout $x_0 \in f^{-1}(\{0\})$, $f''(x_0) = 0$.

Chapitre 10

Structures algébriques

1. Groupes — **2.** Anneaux et corps

Objectifs et compétences du programme

- Reconnaître une structure algébrique.
- Utiliser une caractérisation de sous-structure.
- Manipuler les éléments d'un groupe.
- Effectuer des calculs dans un anneau (pour des éléments qui commutent).

Emmy Noether, mathématicienne allemande (1882-1935)
Elle démarre ses études à l'université d'Erlangen (Bavière) où, en tant que femme, elle devait demander à chaque professeur la permission d'assister aux cours. Mathématicienne de génie, elle a laissé une immense trace dans deux domaines mathématiques distincts :

- en Physique Mathématique, elle a démontré l'un des « théorèmes de Noether » détaillant l'équivalence entre les lois de conservation au sein d'un système physique et l'invariance pour certaines symétries de l'opérateur lagrangien associé au système
- en algèbre générale, elle a entre autres étudié les propriétés des anneaux : dans un célèbre article *Idealtheorie in Ringbereichen*, elle introduit une condition de finitude qui est désormais la définition des anneaux noetheriens.

COURS

Ce court chapitre introduit du vocabulaire pour permettre une présentation unifiée de plusieurs résultats mathématiques. On commence tout d'abord avec la structure de groupe (une seule loi de composition interne) puis on aborde anneaux et corps (avec deux lois de composition internes compatibles entre elles). La théorie complète sous-jacente est, conformément au programme, à peine abordée.
Pour simplifier la lecture, on rappelle les propriétés des lois de composition interne.

Définition 10.1.

Soit un ensemble E muni d'une loi de composition interne $\star$.
▹ La loi $\star$ est commutative si
$$\forall (x,y)\in E^2, \qquad x\star y = y\star x.$$

▹ La loi $\star$ est associative si
$$\forall (x,y,z)\in E^3, \qquad x\star(y\star z)=(x\star y)\star z.$$

▹ La loi $\star$ admet un élément neutre si
$$\exists e\in E, \quad \forall x\in E, \qquad x\star e = e\star x = x.$$

▹ Si la loi $\star$ admet l'élément neutre e, un élément $x\in E$ admet un symétrique (dans E) pour $\star$ si
$$\exists y\in E, \qquad x\star y = y\star x = e.$$

Lorsqu'il y a plusieurs lois de composition interne, on étudie la distributivité.

Définition 10.2.

Soit E un ensemble muni de deux lois de composition interne $\diamond$ et $\star$. La loi $\star$ est distributive sur la loi $\diamond$ si elle vérifie les propriétés de distributivité à gauche et à droite suivantes
$$\forall (x,y,z)\in E^3, \qquad x\star(y\diamond z)=(x\star y)\diamond(x\star z),$$
$$\forall (x,y,z)\in E^3, \qquad (y\diamond z)\star x=(y\star x)\diamond(z\star x).$$

1. Groupes

1.1. Définitions des structures

Définition 10.3.

▹ Un ensemble G est un groupe s'il est muni d'une loi de composition interne $\star$
- associative,
- qui admet un élément neutre $e\in G$,
- telle que tout élément de G admet un symétrique pour cette loi.

▹ Si, de plus, $\star$ est commutative, alors G est abélien (ou commutatif).

Exemple

Voici quelques exemples de parties qui ne sont pas des groupes :
- $(\mathbb{N},+)$ car les éléments non nuls pas de symétrique dans l'ensemble,
- $(\mathbb{Z},\times)$ car les éléments différents de 1 ou -1 n'admettent pas de symétriques dans $\mathbb{Z}$,
- $(\mathbb{R},\times)$ car 0 n'admet pas de symétrique.

Exemple

Voici quelques exemples de groupes :
- $(\mathbb{Z},+)$, $(\mathbb{Q},+)$, $(\mathbb{R},+)$, $(\mathbb{C},+)$,
- $(\mathbb{R}^*,\times)$, $(\mathbb{C}^*,\times)$, $(\mathbb{U},\times)$, $(\mathbb{U}_n,\times)$,
- l'ensemble des similitudes directes muni de la composition, l'ensemble D_n des isométries préservant les n sommets d'un polygone régulier à n sommets lui aussi muni de la composition.

Exemple

Soit $(G,\times)$ et $(G',\otimes)$ deux groupes. On obtient un nouveau groupe en munissant le produit cartésien $G\times G'$ de la loi $\boxtimes$ définie par

$$\forall g,h\in G,\ \forall g',h'\in G',\qquad (g,g')\boxtimes(h,h')=(g\times h,g'\otimes h').$$

On vérifie facilement tous les axiomes (l'élément neutre est (e,e') et le symétrique d'un couple (g,g') est le couple formé des symétriques (g^{-1},g'^{-1})).
De manière informelle, on munit donc l'ensemble des couples de l'opération « terme à terme ».

Proposition 10.4.

Un groupe $(G,\star)$ admet un unique élément neutre et tout élément de G admet un unique symétrique pour $\star$.

Démonstration

▹ Soit e et e' deux éléments neutres de G. Alors, $e=e\star e'=e'$ en utilisant successivement les propriétés d'élément neutre de e' et de e : d'où l'unicité de l'élément neutre.

▹ Soit $x\in G$ et $y\in G$ tel que $x\star y=e$. En multipliant à gauche par le symétrique x^{-1} de x, on obtient $y=x^{-1}$ d'où l'unicité du symétrique. ■

On obtient alors des règles de calcul dans un groupe.

Proposition 10.5.

Soit $(G,\star)$ un groupe. Alors, pour tous $x,\ y\in G$,

$$(x^{-1})^{-1}=x,\qquad (x\star y)^{-1}=y^{-1}\star x^{-1}.$$

Démonstration

▹ Soit $x \in G$. Comme $x \star x^{-1} = x^{-1} \star x = e$ (car x^{-1} est le symétrique de x), on obtient que x est l'élément de G qui, multiplié par x^{-1}, donne e, c'est-à-dire le symétrique de x^{-1}.

▹ Soit $x, y \in G$. Pour obtenir le second point, il suffit d'écrire avec l'associativité

$$\begin{aligned}(x \star y) \star (y^{-1} \star x^{-1}) &= x \star (y \star y^{-1}) \star x^{-1} \\ &= x \star e \star x^{-1} \\ &= x \star x^{-1} = e\end{aligned}$$

et de procéder de même pour le calcul de $(y^{-1} \star x^{-1}) \star (x \star y)$. ■

Notation

▹ Si la loi est notée additivement c'est-à-dire $+$, le symétrique d'un élément x est appelé opposé et noté $-x$; l'élément neutre est noté 0_G. Pour tout $n \in \mathbb{N}$, $n.x$ désigne l'élément défini comme

$$\underbrace{x + x + \ldots + x}_{n \text{ fois}}.$$

On définit de même $n.x$ pour un entier $n < 0$ comme $(-n)(-x)$ où $-x$ est le symétrique de x.

▹ Si la loi est notée multiplicativement, c'est-à-dire $\times$ ou $\circ$ (ou une simple apposition sans symbole), le symétrique d'un élément x est noté x^{-1} et l'élément neutre 1_G ou Id_G. Pour tout $n \in \mathbb{N}$, x^n désigne l'élément défini comme

$$\underbrace{x \times x \times \ldots \times x}_{n \text{ fois}} \quad \text{ou} \quad \underbrace{x \circ x \circ \ldots \circ x}_{n \text{ fois}}.$$

On définit de même x^n pour un entier $n < 0$ comme $(x^{-1})^{-n}$ où x^{-1} est le symétrique de x.

Définition 10.6.

L'ordre d'un groupe G fini est son cardinal ; on le note $|G|$.

Exemple

D'après le cours sur les nombres complexes, le groupe $\mathbb{U}_n$ des racines $n^{\text{ièmes}}$ de l'unité est d'ordre n.

1.2. Table de Cayley d'un groupe

Un groupe est essentiellement compris au travers de sa loi de composition interne (on pourrait changer les noms des éléments et la notation de la loi, cela ne change pas intrinsèquement le groupe).

Décrivons les lois de groupe sur des groupes $(G, \star)$ d'ordre inférieur à 4 à l'aide de leurs tables de Cayley, c'est-à-dire le tableau à deux entrées indexées par les éléments de G. À l'intersection de la ligne $a \in G$ et de la colonne $b \in G$ de la table, on trouve l'élément $a \star b$. Autrement dit, la table de Cayley est la « table de multiplication » du groupe.

1.2.1. Ordre 2

Il y a une seule possibilité pour les groupes d'ordre 2, que l'on peut interpréter comme $\{\pm 1\} = \mathbb{U}_2$ avec le produit ou l'ensemble d'une réflexion et de l'identité muni de la composition :

$\star$	e	x
e	e	x
x	x	e

En effet, la première ligne et la première colonne sont imposées d'après la propriété d'élément neutre e ; la dernière case est aussi imposée par l'existence d'un élément symétrique de x (qui est donc x lui-même).

1.2.2. Ordre 3

Il n'y a également qu'une seule possibilité pour les groupes d'ordre 3 (que l'on peut donc voir comme $\mathbb{U}_3 = \{1, j, j^2\}$ et qui est donc commutatif) et voici la table de sa loi :

$\star$	e	x	x^2
e	e	x	x^2
x	x	x^2	e
x^2	x^2	e	x

En effet, il n'y a que les trois cases du coin en bas à droite pour lesquelles un argument est nécessaire : on complète la case en position $(2,3)$ (deuxième ligne, troisième colonne) et celle en position $(3,2)$ d'après l'existence d'un symétrique à x (qui est donc x^2). La case en position $(3,3)$ ne peut contenir ni e, ni x^2 par unicité du symétrique et de l'élément neutre : elle contient donc x.

On remarque la propriété suivante qui est générale.

Proposition 10.7. Carré latin

Dans la table de Cayley d'un groupe fini, chaque élément apparaît une et une seule fois sur chaque ligne et chaque colonne.

On peut aussi reformuler ainsi cette propriété.

Proposition 10.8.

Pour tout $g \in G$, les applications $\left\{\begin{array}{ccc} G & \to & G \\ x & \mapsto & x \star g \end{array}\right.$ et $\left\{\begin{array}{ccc} G & \to & G \\ x & \mapsto & g \star x \end{array}\right.$ sont des bijections.

Démonstration

Pour trouver les bijections réciproques, il suffit de considérer les applications similaires définies en remplaçant g par g^{-1}. ■

1.2.3. Ordre 4

Forts de cette propriété, on trouve deux tables possibles pour les groupes à 4 éléments :

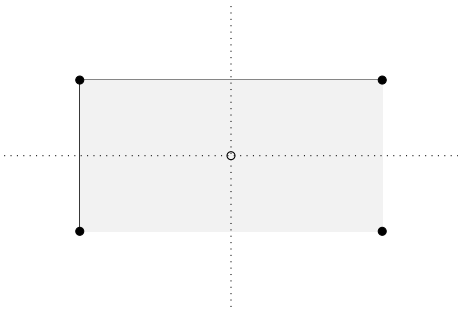
$\star$	e	a	b	c
e	e	a	b	c
a	a	b	c	e
b	b	c	e	a
c	c	e	a	b

$\star$	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

La première correspond par exemple au groupe $(\mathbb{U}_4, \times)$ (avec $e = 1$, $a = i$, $b = -1$ et $c = -i$) ; la seconde au groupe $(\mathscr{P}(\{0,1\}), \Delta)$ (avec $e = \varnothing$, $a = \{0\}$, $b = \{1\}$ et $c = \{0,1\}$).

Exemple

Le groupe $(V_4, \circ)$ (pour *Vierergruppe*) des isométries du plan laissant stable un rectangle (non carré) est composé de e l'identité du plan, a et b les deux réflexions par rapport aux droites passant par les milieux des côtés opposés du rectangle et c la symétrie par rapport au centre du rectangle.



On vérifie que ce groupe admet bien la seconde table de Cayley exhibée ci-dessus.

1.3. Sous-groupes

On cherche à voir comment une partie d'un groupe peut hériter de la structure de groupe.

Définition 10.9.

Soit $(G, \star)$ un groupe. Une partie H de G est un sous-groupe de $(G, \star)$ si

- H est stable pour $\star$, c'est-à-dire
$$\forall x, y \in H, \qquad x \star y \in H.$$
- $(H, \star)$ est un groupe.

Remarque

En particulier, l'élément neutre de G appartient à tous ses sous-groupes. De plus, si H est un sous-groupe de G, alors, pour tout $x \in H$, $x^{-1} \in H$.

La définition de sous-groupe est peu pratique puisqu'elle requiert de redémontrer des propriétés déjà acquises (comme, par exemple, l'associativité de la loi). La proposition suivante est la caractérisation pratique des sous-groupes.

Proposition 10.10.

Soit $(G,\star)$ un groupe. Une partie non vide H de G est un sous-groupe de G si, et seulement si,

$$\forall x, y \in H, \qquad x \star y^{-1} \in H.$$

Remarque

En notation additive, cela signifie qu'une partie non vide H de G est un sous-groupe de $(G,+)$ si, et seulement si,

$$\forall x, y \in H, \qquad x - y \in H.$$

Conseils méthodologiques

On note bien l'importance du caractère non vide de H dans cette caractérisation : pour l'établir, il suffit en général de remarquer que l'élément neutre de G appartient à H.

Démonstration

Le sens direct est évident. Il suffit donc de montrer le sens retour (qui est effectivement la caractérisation des sous-groupes). $\star$ est bien associative sur H car associative sur G. L'élément neutre de $\star$ appartient bien à H car $e = x \star x^{-1} \in H$ avec $x \in H$. Pour tout $x \in H$, $x^{-1} = e \star x^{-1} \in H$ donc tout élément de H admet un symétrique dans H. Il suffit donc de vérifier que $\star$ est une loi de composition interne. Or, pour tous $x, y \in H$, $y^{-1} \in H$ donc $x \star y = x \star (y^{-1})^{-1} \in H$. ■

Exemple

Pour tout $n \in \mathbb{Z}$, $n\mathbb{Z} = \{kn,\ k \in \mathbb{Z}\}$ est un sous-groupe de $(\mathbb{Z},+)$ car $0 \in n\mathbb{Z}$ et la différence de deux multiples de n est un multiple de n.

Exemple

Pour tout $n \in \mathbb{N}$, $\mathbb{U}_n$ est un sous-groupe de $(\mathbb{U},\times)$ car $1 \in \mathbb{U}_n$ et le quotient de deux racines $n^{\text{ièmes}}$ de l'unité est une racine $n^{\text{ième}}$ de l'unité.

Exemple

Pour tous $m, n \in \mathbb{N}$, $\mathbb{U}_m$ est un sous-groupe de $\mathbb{U}_n$ si, et seulement si, $\mathbb{U}_m \subset \mathbb{U}_n$ c'est-à-dire si m divise n.

Précisons les résultats généraux suivants déjà établis.

Proposition 10.11.

- Les sous-groupes de $(\mathbb{Z},+)$ sont les parties de la forme $n\mathbb{Z}$ avec $n \in \mathbb{N}$.
- Les sous-groupes de $(\mathbb{R},+)$ sont soit les parties de la forme $a\mathbb{Z}$ avec $a \in \mathbb{R}$, soit denses.

Exemple

Le centre du groupe $(G,\star)$ est l'ensemble $Z(G)$ des éléments de G qui commutent avec tous les autres éléments, c'est-à-dire

$$Z(G)=\{x\in G,\quad \forall y\in G,\quad x\star y=y\star x\}.$$

En particulier, $Z(G)=G$ si G est abélien.
Montrons que $Z(G)$ est un sous-groupe de G.
Tout d'abord, $Z(G)$ est non vide car $e\in Z(G)$.
Ensuite, pour tous $x,\ y\in Z(G)$ et tout $z\in G$,

$$\begin{aligned}(x\star y^{-1})\star z &= x\star(y^{-1}\star z)=x\star(z\star y^{-1})\\ &=(x\star z)\star y^{-1}=(z\star x)\star y^{-1}=z\star(x\star y^{-1}).\end{aligned}$$

On a utilisé respectivement l'associativité, le fait que $y^{-1}\in Z(G)$, l'associativité à nouveau, le fait que $x\in Z(G)$ et encore l'associativité. Ainsi, $x\star y^{-1}\in Z(G)$ ce qui termine la caractérisation de sous-groupe de G.

La caractérisation des sous-groupes en terme de stabilité permet de déduire directement la proposition suivante.

Proposition 10.12.

Une intersection de sous-groupes de $(G,\star)$ est encore un sous-groupe de G.

En revanche, la réunion de deux sous-groupes est rarement un sous-groupe.

Proposition 10.13.

Soit H et K deux sous-groupes de G ; alors $H\cup K$ est un sous-groupe de G si, et seulement si, $H\subset K$ ou $K\subset H$.

Démonstration

▹ Le sens retour est évident car, alors $H\cup K=H$ ou $H\cup K=K$ donc $H\cup K$ est un sous-groupe.

▹ Réciproquement supposons que $H\cup K$ est un sous-groupe, que $H\not\subset K$; considérons $h\in H\setminus K$ et $k\in K$. L'élément $h\star k$ appartient à $H\cup K$ (car il est le produit de deux éléments de $H\cup K$). S'il appartient à K, alors $h=(h\star k)\star k^{-1}\in K$ ce qui est contraire à l'hypothèse sur h. Ainsi, $h\star k$ appartient à H et $k=h^{-1}\star(h\star k)\in H$. En conclusion, $K\subset H$.

■

Comme une intersection de sous-groupes est encore un sous-groupe, la notion de plus petit sous-groupe contenant une partie A a bien un sens, comme intersection de tous les sous-groupes contenant A.

Définition 10.14.

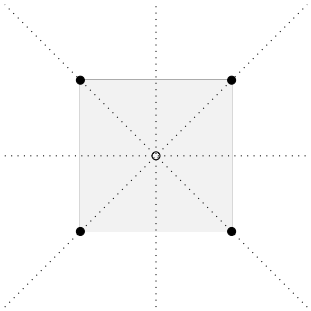
▹ Le sous-groupe engendré par une partie A est le plus petit sous-groupe contenant cette partie A. On le note souvent $\langle A\rangle$.
▹ Un groupe est monogène s'il est engendré par un unique élément.
▹ Un groupe est cyclique s'il est à la fois monogène et de cardinal fini.
▹ L'ordre d'un élément est l'ordre du sous-groupe engendré par cet élément, c'est-à-dire son cardinal si ce groupe est fini et $+\infty$ sinon.

Exemple

Le sous-groupe de $\mathbb{U}_4$ engendré par i est $\mathbb{U}_4$ (de même pour $-i$) ; le sous-groupe engendré par -1 est $\mathbb{U}_2 = \{\pm 1\}$.

Exemple

Considérons le groupe D_4 des isométries laissant stable un carré.



Ce groupe se compose des 8 éléments suivants :

- 4 isométries directes : l'identité Id, la rotation ρ d'angle $\frac{\pi}{2}$, la symétrie centrale ρ^2 et la rotation ρ^3 d'angle $\frac{3\pi}{2}$.
- 4 isométries indirectes : les réflexions s_1 et s_2 par rapport aux diagonales du carré, les réflexions σ_1 et σ_2 par rapport aux droites passant par les milieux des côtés opposés.

Le groupe engendré par ρ est d'ordre 4 et composé de Id, ρ, ρ^2 et ρ^3 : il contient obligatoirement ces éléments et l'ensemble $\{\mathrm{Id}, \rho, \rho^2, \rho^3\}$ est bien un sous-groupe.
Le groupe engendré par ρ^2 est d'ordre 2 (avec Id) comme tous les groupes engendrés par une réflexion. Le groupe engendré par ρ et s_1 est égal à D_4 car $\rho \circ s_1 = \sigma_1$, $\rho^2 \circ s_1 = s_2$, $\rho^3 \circ s_1 = \sigma_2$: il contient donc tous les éléments de D_4. Plus généralement, le groupe engendré par ρ et une réflexion est égal à D_4 pour les mêmes raisons.

Proposition 10.15.

En notation multiplicative (respectivement additive), le sous-groupe engendré par un élément $g \in G$ est $\{g^k,\ k \in \mathbb{Z}\}$ (respectivement $\{kg,\ k \in \mathbb{Z}\}$).

Démonstration

Cet ensemble est bien un sous-groupe de G contenant g et il est contenu par tous les sous-groupes de G contenant g. ■

Corollaire 10.16.

L'ordre d'un élément g est alors le plus petit entier $k \in \mathbb{N} \setminus \{0\}$ ($+\infty$ s'il n'en existe pas) tel que $g^k = e$ (en notation multiplicative) ou $k.g = e$ (en notation additive).

Remarque

Un groupe fini d'ordre n est cyclique si, et seulement s'il existe un élément d'ordre n.

Remarque
Un groupe monogène est toujours abélien car, pour tout g et tous $m, n \in \mathbb{N}$, (en notation multiplicative)

$$g^m \star g^n = g^{m+n} = g^n \star g^m.$$

Exemple

Les sous-groupes $n\mathbb{Z}$ sont monogènes car engendrés par n (ou $-n$).
On peut vérifier également que les sous-groupes $\mathbb{U}_n$ sont cycliques.

Exemple

Reprenons les deux tables de Cayley des groupes d'ordre 4 déjà détaillées.
▹ Pour un groupe qui admet la première table, il y a un élément d'ordre 1 (l'élément neutre), un élément d'ordre 2, et deux éléments d'ordre 4 : ce groupe est cyclique.
▹ Pour un groupe qui admet la seconde table, il y a un élément d'ordre 1 (l'élément neutre), et trois éléments d'ordre 2 : ce groupe n'est pas cyclique puisqu'il est d'ordre 4 et qu'il n'y a aucun élément d'ordre 4.

1.4. Sous-groupes d'un groupe fini

On remarque que si un groupe G est fini, alors ses sous-groupes H sont aussi d'ordre fini (et d'ordre inférieur à celui de G). En fait, on peut dire beaucoup mieux avec le théorème suivant qui n'est pas au programme de première année mais qui s'avère utile pour appréhender les différents exemples que l'on peut rencontrer.

Théorème 10.17. Lagrange

Soit G un groupe fini et H un sous-groupe de G. Alors, l'ordre de H divise l'ordre de G.

Démonstration
Notons, pour tout $x \in G$, $xH = \{x \star h,\ h \in H\}$ avec $\star$ la loi de G.
▹ Pour tout $x \in G$, l'application

$$\left\{\begin{array}{ccc} H & \to & xH \\ h & \mapsto & x \star h \end{array}\right.$$

est une bijection de bijection réciproque

$$\left\{\begin{array}{ccc} xH & \to & H \\ g & \mapsto & x^{-1} \star g \end{array}\right.$$

Par conséquent, $|xH| = |H|$.
▹ Pour tous $x, y \in G$ tels que $xH \neq yH$, alors

$$xH \cap yH = \varnothing.$$

En effet, si $xH \cap yH$ est non vide, il existe $h, h' \in H$ tel que $x \star h = y \star h'$ donc

$$x = y \star h' \star h^{-1} \in yH;$$

par conséquent $xH \subset yH$ et d'après le cardinal, $xH = yH$: contradiction.
▹ Les éléments de $\{xH,\ x \in G\}$ forment donc une partition de G en ensembles $x_1H, \ldots, x_pH$ de cardinal $|H|$ donc

$$|G| = \sum_{k=1}^{p} |x_kH| = p|H|;$$

ainsi, $|H|$ divise $|G|$. ■

Corollaire 10.18.

Un groupe G d'ordre premier p est cyclique.

Démonstration

Soit $x \in G$ distinct de l'élément neutre. D'après le théorème de Lagrange, l'ordre du sous-groupe engendré par x divise p donc est égal à 1 ou à p or ce sous-groupe contient deux éléments distincts (e et x) donc ne peut être de cardinal 1. Le sous-groupe engendré par x est égal à G et donc G est cyclique. ■

Remarque

Cet argument montre que tout élément d'un groupe d'ordre premier, distinct de l'élément neutre, est générateur du groupe.

En couplant encore le théorème de Lagrange avec un peu d'arithmétique, on obtient le corollaire suivant.

Corollaire 10.19.

Soit G un groupe, H et K deux sous-groupes finis d'ordre p et q, premiers entre eux.
Le sous-groupe $H \cap K$ est réduit à l'élément neutre.

Démonstration

La partie $H \cap K$ est un sous-groupe de H et de K ; d'après Lagrange, $|H \cap K|$ divise à la fois $|H|$ et $|K|$ donc le PGCD de ces deux entiers qui est 1. Ainsi, $|H \cap K|$ est de cardinal 1 donc réduit à l'élément neutre. ■

Exemple

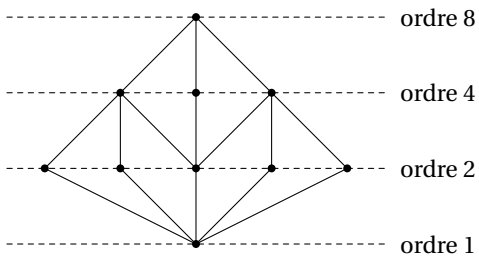
Il n'y a pas d'élément d'ordre 2 dans un groupe d'ordre impair. Par exemple, le groupe des isométries du plan engendré la rotation d'angle $\frac{2\pi}{2n+1}$ ne contient pas de réflexion.

Exemple

Déterminons tous les sous-groupes du groupe D_4. Les sous-groupes de D_4 sont d'ordre 1, 2, 4, ou 8 d'après le théorème de Lagrange.

- Il n'y a qu'un seul groupe d'ordre 1 : $\{\mathrm{Id}\}$.
- Un groupe d'ordre 2 est composé de l'identité et d'un élément d'ordre 2 ; on en obtient 5 : $\{\mathrm{Id}, s_1\}$, $\{\mathrm{Id}, s_2\}$, $\{\mathrm{Id}, \sigma_1\}$, $\{\mathrm{Id}, \sigma_2\}$ et $\{\mathrm{Id}, \rho^2\}$ (ce dernier est d'ailleurs le centre de D_4).
- Si un sous-groupe contient une rotation ρ ou ρ^3 et une réflexion, alors il contient tous les éléments de D_4. De même, si un sous-groupe contient une réflexion s_i et une réflexion σ_j, alors il contient tous les éléments de D_4. Par conséquent, il y a 3 sous-groupes d'ordre 4 : un cyclique $\{\mathrm{Id}, \rho, \rho^2, \rho^4\}$ et deux autres $\{\mathrm{Id}, \rho^2, s_1, s_2\}$ et $\{\mathrm{Id}, \rho^2, \sigma_1, \sigma_2\}$.
- Il n'y a qu'un seul groupe d'ordre 8 : D_4.

On représente quelquefois ces 10 sous-groupes sous la forme d'un graphe pour l'inclusion :



1.5. Morphismes de groupes

L'idée de cette section est d'étudier les applications entre groupes qui sont compatibles avec les structures.

Définition 10.20.

Soit $(G,\star)$ et $(G',\otimes)$ deux groupes et $f : G \to G'$.
▹ L'application f est un morphisme de groupes si

$$\forall x, y \in G, \qquad f(x \star y) = f(x) \otimes f(y).$$

▹ L'application f est un isomorphisme de groupes si f est un morphisme bijectif.
▹ Un automorphisme de G est un isomorphisme de G dans G.

Intuitivement, un morphisme « transporte » la structure initiale sur la structure du groupe d'arrivée.

Exemple

Dans le chapitre consacré aux fonctions usuelles, nous avons déjà rencontré deux isomorphismes :

$$\left\{\begin{array}{ccc} (\mathbb{R},+) & \to & (\mathbb{R}_+^*,\times) \\ x & \mapsto & e^x \end{array}\right. \qquad \left\{\begin{array}{ccc} (\mathbb{R}_+^*,\times) & \to & (\mathbb{R},+) \\ x & \mapsto & \ln(x) \end{array}\right.$$

Exemple

Soit G un groupe et $g \in G$. L'application $\varphi_g : x \mapsto g \star x \star g^{-1}$ définit un automorphisme de G appelé automorphisme intérieur.
En effet, pour tout $x, y \in G$,

$$\begin{aligned} \varphi_g(x \star y) &= g \star x \star y \star g^{-1} \\ &= g \star x \star g^{-1} \star g \star y \star g^{-1} \\ &= \varphi_g(x) \star \varphi_g(y). \end{aligned}$$

De plus, φ_g est bijective de réciproque $\varphi_{g^{-1}}$.

Proposition 10.21.

L'image de l'élément neutre (respectivement l'image du symétrique de x) par un morphisme de groupes est l'élément neutre (respectivement le symétrique de l'image de x).

Démonstration

▹ Pour tout $x \in G$, $f(x) = f(x \star e) = f(x) \otimes f(e)$ donc, en multipliant à gauche par $f(x)^{-1}$, on obtient $e' = f(e)$.
▹ De même, $f(x) \otimes f(x^{-1}) = f(x \star x^{-1}) = f(e) = e'$ d'où $f(x^{-1}) = f(x)^{-1}$. ■

Une simple utilisation de la définition nous donne :

Proposition 10.22.

La composée de deux morphismes de groupes est encore un morphisme de groupes.

Il suffit alors d'utiliser la caractérisation de sous-groupes dans le groupe des bijections de G dans G muni de la composition pour obtenir le résultat suivant.

Corollaire 10.23.

L'ensemble des automorphismes de G, noté $\mathrm{Aut}(G)$, est un groupe pour la composition.

Exemple

Vérifions que l'ensemble $\mathrm{Int}(G)$ des automorphismes intérieurs (définis ci-dessus) est bien un sous-groupe de $\mathrm{Aut}(G)$.
▹ $\mathrm{Int}(G)$ est une partie non vide car l'identité de G est un automorphisme intérieur associé à l'élément neutre.
▹ Soit $g, g' \in G$. Notons φ_g et $\varphi_{g'}$ les automorphismes intérieurs associés. Alors, $(\varphi_{g'})^{-1} = \varphi_{g'^{-1}}$ puis $\varphi_g \circ (\varphi_{g'})^{-1} = \varphi_{g \star g'^{-1}} \in \mathrm{Int}(G)$.
D'après la caractérisation des sous-groupes, $\mathrm{Int}(G)$ est un sous-groupe de $\mathrm{Aut}(G)$.

Proposition 10.24.

Soit $f : G \to G'$ un morphisme de groupes.
▹ Si H est un sous-groupe de G, alors $f(H)$ est un sous-groupe de G'.
▹ Si H' est un sous-groupe de G, alors $f^{-1}(H')$ est un sous-groupe de G.

Dans la preuve qui suit, on note pour simplifier de la même manière la loi sur G et sur G'.

Démonstration

▹ L'ensemble $f(H)$ est non vide puisqu'il contient $e_H = f(e_G)$.
Pour tous $x, y \in f(H)$, il existe $s, t \in H$ tels que $x = f(s)$ et $y = f(t)$. Alors

$$x \star y^{-1} = f(s) \star f(t)^{-1} = f(s \star t^{-1}) \in f(H).$$

Ainsi, $f(H)$ est un sous-groupe de G d'après la caractérisation des sous-groupes.
▹ L'ensemble $f^{-1}(H')$ est non vide puisqu'il contient l'élément neutre de G.
Pour tous $x, y \in f^{-1}(H')$, $f(x), f(y) \in H'$ donc

$$f(x \star y^{-1}) = f(x) \star f(y)^{-1} \in H'.$$

Ainsi, $x \star y^{-1} \in f^{-1}(H')$ et $f^{-1}(H')$ est un sous-groupe de G d'après la caractérisation précédente. ■

Définissons deux sous-groupes particuliers associés à un morphisme donné.

Définition 10.25.

Soit f un morphisme de groupes de $(G, \star)$ dans $(G', \otimes)$.
▹ Le noyau de f est le sous-groupe de G défini par :

$$\mathrm{Ker}\, f = f^{-1}(\{e'\}) = \{x \in G, \quad f(x) = e'\}.$$

▹ L'image de f est le sous-groupe de G' défini par :

$$\mathrm{Im}\, f = f(G) = \{f(x), \quad x \in G\}.$$

Proposition 10.26.

Soit $f : G \to G'$ un morphisme de groupes. Alors,
▹ f est injectif si, et seulement si, $\operatorname{Ker} f = \{e\}$.
▹ f est surjectif si, et seulement si, $\operatorname{Im} f = G'$.

Démonstration
Il suffit de montrer le premier point (le second est la réécriture de la définition de la surjectivité). Le sens direct est élémentaire (e' n'admet qu'un antécédent e). Pour le sens retour, il suffit de remarquer que $f(x) = f(y)$ si, et seulement si, $f(x \star y^{-1}) = f(x) \otimes f(y)^{-1} = e'$, c'est-à-dire $x \star y^{-1} \in \operatorname{Ker} f$. ■

1.6. Ordre d'un élément

Proposition 10.27.

Soit $(G, \star)$ un groupe et $g \in G$. L'ordre de g est

$$\min\{k \in \mathbb{N} \setminus \{0\},\ g^k = e\}$$

avec la convention $+\infty$ si l'ensemble considéré est vide.

Démonstration
Considérons le morphisme de groupes défini par

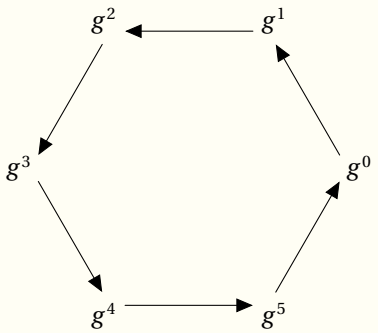
$$\varphi : \left\{ \begin{array}{ccc} \mathbb{Z} & \to & G \\ k & \mapsto & g^k \end{array} \right.$$

Le noyau $\operatorname{Ker} \varphi$ est un sous-groupe de $\mathbb{Z}$; or tous les sous-groupes de $\mathbb{Z}$ sont de la forme $n\mathbb{Z}$ avec $n \in \mathbb{N}$.
▹ Si $\operatorname{Ker} \varphi = n\mathbb{Z}$ avec $n \neq 0$, alors le groupe engendré par g est fini d'ordre n : plus précisément c'est le groupe $\{e, g, \ldots, g^{n-1}\}$.
▹ Si $\operatorname{Ker} \varphi = \{0\}$, alors le groupe engendré par g est isomorphe à $\mathbb{Z}$; en effet, φ induit un isomorphisme $\tilde{\varphi}$ entre $\mathbb{Z}$ et son image, à savoir le groupe engendré par g : surjective par définition et injective car le noyau ne contient que l'élément neutre car

$$\{0\} \subset \operatorname{Ker} \tilde{\varphi} \subset \operatorname{Ker} \varphi = \{0\}.$$

■

Remarque
On comprend un peu mieux l'appellation cyclique avec cette proposition. Par exemple, pour un groupe cyclique d'ordre 6 engendré par un élément g, la multiplication par g opère comme sur le schéma suivant



Dans un groupe en notation additive, on obtient l'analogue suivant.

Proposition 10.28.

Soit $(G,+)$ un groupe et $g \in G$. L'ordre de g est

$$\min\{k \in \mathbb{N} \setminus \{0\},\ kg = 0_G\}$$

avec la convention $+\infty$ si l'ensemble considéré est vide.

Déduisons-en quelques propriétés élémentaires sur l'ordre d'un élément.

Proposition 10.29.

Soit $(G,\star)$ un groupe et $g \in G$ d'ordre n. Alors,
▹ le symétrique g^{-1} est d'ordre n ;
▹ pour tout $h \in G$, $h \star g \star h^{-1}$ est d'ordre n.

Démonstration
▹ Il suffit de remarquer que, pour tout $k \in \mathbb{N} \setminus \{0\}$, $g^k = e$ si, et seulement si, $(g^{-1})^k = e$ donc

$$\min\{k \in \mathbb{N} \setminus \{0\},\ g^k = e\} = \min\{k \in \mathbb{N} \setminus \{0\},\ g^{-k} = e\}.$$

▹ Cette fois-ci, il suffit de remarquer que, pour tout $k \in \mathbb{N} \setminus \{0\}$, $(h \star g \star h^{-1})^k = h \star g^k \star h^{-1} = e$ si, et seulement si, $g^k = e$ donc

$$\min\{k \in \mathbb{N} \setminus \{0\},\ g^k = e\} = \min\{k \in \mathbb{N} \setminus \{0\},\ (h \star g \star h^{-1})^k = e\}.$$

■

2. Anneaux et corps

2.1. Définitions des anneaux

Définition 10.30.

Un ensemble A est un anneau s'il est muni de deux lois de composition interne $+$ et $\times$ telles que
- $(A,+)$ est un groupe abélien d'élément neutre noté 0_A,
- $\times$ est associative et admet un élément neutre $1_A \neq 0_A$,
- $\times$ est distributive sur $+$.

Si, de plus, $\times$ est commutative, alors l'anneau est commutatif.

Voici quelques exemples usuels d'anneaux.

Exemple

L'ensemble des entiers $\mathbb{Z}$ et l'ensemble des décimaux $\mathbb{D}$ sont des anneaux pour les opérations $+$ et $\times$ usuelles.

Exemple

Les ensembles de polynômes $\mathbb{R}[X]$ ou $\mathbb{C}[X]$ à coefficients dans $\mathbb{R}$ ou $\mathbb{C}$ sont des anneaux pour les opérations $+$ et $\times$ usuelles.

Exemple

Les ensembles de n-uplets $\mathbb{Z}^n$, de suites $\mathbb{R}^{\mathbb{N}}$, de fonctions $\mathbb{R}^I$ avec I une partie de $\mathbb{R}$ sont des anneaux pour les opérations $+$ et $\times$ effectuées « composante par composante ».

Exemple

Soit E un ensemble. L'ensemble $\mathscr{P}(E)$ des parties de E est un anneau pour la différence symétrique Δ et l'intersection $\cap$.

Exemple

L'ensemble des entiers de Gauss $\mathbb{Z}[i] = \{a + ib,\ \text{avec } a, b \in \mathbb{Z}\}$ est un anneau commutatif pour l'addition et la multiplication.

Remarque

▹ L'élément 0_A est « absorbant » pour $\times$; en effet,

$$\forall x \in A, \qquad x = (0_A + 1_A) \times x = 0_A \times x + x.$$

D'où, en ajoutant l'opposé $-x$ des deux côtés de l'égalité, $0_A \times x = 0_A$.

▹ Pour tout $x \in A$, $(-1_A) \times x$ est le symétrique de x pour $+$, c'est-à-dire $-x$, car

$$\begin{aligned} 0_A &= 0_A \times x \\ &= (1_A - 1_A) \times x \\ &= 1_A \times x + (-1_A) \times x \\ &= x + (-1_A) \times x. \end{aligned}$$

Remarque

Dans un anneau A, on dispose des règles de calcul suivantes qui généralisent les identités remarquables :

$$\forall a \in A,\ \forall n \in \mathbb{N}, \qquad (1_A - a) \times \left(\sum_{k=0}^{n} a^k \right) = 1_A - a^{n+1},$$

qui se généralise à son tour en la formule de Bernoulli (avec une démonstration analogue au cas complexe)

$$\forall a, b \in A,\ \forall n \in \mathbb{N}, \qquad a \times b = b \times a \quad \Rightarrow \quad (b - a) \times \left(\sum_{k=0}^{n} a^k \times b^{n-k} \right) = b^{n+1} - a^{n+1}.$$

Remarque

De même, on obtient la formule du binôme par récurrence pour deux éléments d'un anneau A qui commutent.

$$\forall a, b \in A,\ \forall n \in \mathbb{N}, \qquad a \times b = b \times a \quad \Rightarrow \quad (a + b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k \times b^{n-k}.$$

En particulier, la formule du binôme de Newton est toujours vraie dans un anneau commutatif.

Définition 10.31.

Soit $(A,+,\times)$ un anneau et $x \in A$.
▹ x est régulier à droite (respectivement à gauche) si pour tout $y \in A$, l'égalité $y \times x = 0_A$ (respectivement $x \times y = 0_A$) entraîne $y = 0_A$.
▹ x est inversible si x admet un inverse dans A.
▹ x est un diviseur de 0 si $x \neq 0_A$ et s'il existe $y \neq 0_A$ tel que $x \times y = y \times x = 0_A$.

Exemple

Dans $\mathbb{Z}$, tout élément non nul est régulier mais seuls 1 et -1 sont inversibles.

Proposition 10.32.

Un élément qui n'est pas un diviseur de zéro est régulier.

Proposition 10.33.

L'ensemble A^* des éléments inversibles d'un anneau $(A,+,\times)$ est un groupe pour la loi $\times$.

Remarque

On remarque que cette notation est cohérente avec l'usage $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$ et $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$ puisque tout élément non nul de ces ensembles est inversible. En revanche, $\mathbb{Z}^* = \{-1,1\} \neq \mathbb{Z} \setminus \{0\}$.

Démonstration

Par définition d'anneau, $\times$ est associative et son élément neutre 1_A appartient à A^*. Soit $x, y \in A^*$. Alors, $x \times y$ est inversible d'inverse $y^{-1} \times x^{-1}$ donc $x \times y \in A^*$ et $\times$ est une loi de composition interne pour A^*.
De plus, pour tout $x \in A^*$, x^{-1} est inversible (d'inverse x) donc appartient à A^*.
En conclusion, A^* est un groupe pour $\times$. ■

Exemple

Montrons que le groupe des inversibles de l'ensemble $\mathbb{D}$ des décimaux est

$$\{2^\alpha 5^\beta,\ (\alpha,\beta) \in \mathbb{Z}^2\}.$$

Soit $n \in \mathbb{Z}$ et $a \in \mathbb{N}$ tel que $\frac{n}{10^a} \in \mathbb{D}^*$; il existe $m \in \mathbb{Z}$ et $b \in \mathbb{N}$ tels que

$$\frac{n}{10^a} \cdot \frac{m}{10^b} = 1.$$

Par conséquent, n est un diviseur de 10^{a+b} et $\frac{n}{10^a} \in \{2^\alpha 5^\beta,\ (\alpha,\beta) \in \mathbb{Z}^2\}$.
Réciproquement, on vérifie que tous les éléments de la forme $2^\alpha 5^\beta$ sont des décimaux inversibles.

Exemple

Étudions le groupe des inversibles de l'anneau $\mathbb{Z}[i]$ des nombres complexes de parties réelle et imaginaire dans $\mathbb{Z}$.
À cette fin, posons $N(a+ib) = |a+ib|^2 = a^2+b^2$ pour tout $a+ib \in \mathbb{Z}[i]$. Cette fonction est, par définition, multiplicative (l'image d'un produit est le produit des images).
▹ Vérifions que $z \in \mathbb{Z}[i]$ est inversible si, et seulement si, $N(z) = 1$.
Le sens direct est évident car $N(z) \in \mathbb{N}$ et $N(z)N(z^{-1}) = N(1) = 1$.
Pour le sens retour, on remarque que si $N(z) = 1$ alors $z^{-1} = \overline{z} \in \mathbb{Z}[i]$. Ainsi, $a+ib \in \mathbb{Z}[i]^*$ si, et

seulement si, $a^2+b^2=1$. Comme a et b sont entiers, on obtient quatre possibilités qui correspondent à $1, i, -1, -i$ et le groupe des inversibles de $\mathbb{Z}[i]$ est $\mathbb{U}_4$.

Définition 10.34.

Un anneau intègre est un anneau commutatif sans diviseur de zéro.

Exemple

$\mathbb{Z}$ est un anneau intègre. En revanche, $\mathbb{Z}^2$ muni de l'addition et du produit composante par composante n'est pas un anneau intègre car il admet des diviseurs de zéro (tous les couples avec une composante nulle) comme, par exemple,

$$(1,0)\times(0,1)=(0,0).$$

2.2. Définitions des corps

Définition 10.35.

Un ensemble A est un corps s'il est muni de deux lois de composition interne $+$ et $\times$ telles que $(A,+,\times)$ est un anneau, que tous les éléments différents de 0_A sont inversibles et que $\times$ est commutative (c'est-à-dire si $(A\setminus\{0_A\},\times)$ est un groupe abélien).

Exemple

Les ensembles de nombres $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, $\mathbb{Q}[\sqrt{2}]=\{a+b\sqrt{2},\quad a,b\in\mathbb{Q}\}$ sont des corps.

Proposition 10.36.

Un corps est un anneau intègre, la réciproque est fausse.

Démonstration

Tout élément non nul d'un corps est inversible donc ne peut être un diviseur de zéro. L'ensemble des entiers $\mathbb{Z}$ est un anneau intègre sans être un corps (par exemple, 2 n'admet pas d'inverse dans $\mathbb{Z}$). ■

2.3. Sous-structures

Définition 10.37.

Une partie B d'un anneau (respectivement corps) $(A,+,\times)$ est un sous-anneau (respectivement sous-corps) si

- B est stable par $+$ et $\times$;
- $(B,+,\times)$ est un anneau (respectivement corps) avec le même élément neutre pour $\times$.

Comme pour les sous-groupes, on dispose des caractérisations suivantes en termes de stabilité.

Proposition 10.38.

Une partie B d'un anneau $(A,+,\times)$ est un sous-anneau si, et seulement si,
- $(B,+)$ est un sous-groupe de $(A,+)$;
- $1_A \in B$;
- B est stable par multiplication.

Proposition 10.39.

Une partie L d'un corps $(K,+,\times)$ est un sous-corps si, et seulement si :
- $(L,+)$ est un sous-groupe de $(K,+)$;
- L est stable par multiplication ;
- pour tout $x \in L \setminus \{0_K\}$, $x^{-1} \in L$.

Voici quelques exemples immédiats de sous-anneaux.

Exemples

$(\mathbb{Z},+,\times)$ est un sous-anneau de $(\mathbb{Q},+,\times)$. $(\mathscr{C}^\infty(I,\mathbb{R}),+,\times)$ est un sous-anneau de $(\mathbb{R}^I,+,\times)$.

Exemple

Soit $\alpha \in \mathbb{C}$ tel que $\alpha^2 = a\alpha + b$ avec $(a,b) \in \mathbb{Z}^2$. Montrons que la partie $\mathbb{Z}[\alpha] = \{x + y\alpha,\ (x,y) \in \mathbb{Z}^2\}$ est un sous-anneau de $\mathbb{C}$.
▹ Comme $\mathbb{Z}[\alpha]$ est non vide et stable par différence, c'est un sous-groupe de $\mathbb{C}$.
▹ Par ailleurs, $1 = 1 + 0\alpha \in \mathbb{Z}[\alpha]$.
▹ Il ne reste plus qu'à étudier la stabilité par produit. Pour tous x, y, x' et $y' \in \mathbb{Z}$

$$\begin{aligned}(x+y\alpha)(x'+y'\alpha) &= xx' + (xy' + x'y)\alpha + yy'\alpha^2 \\ &= (xx' + byy') + (xy' + x'y + ayy')\alpha \in \mathbb{Z}[\alpha].\end{aligned}$$

Remarque

Attention, un sous-anneau de A n'est pas seulement un anneau inclus dans A.

Remarque

Par exemple, $(\mathbb{Z}\times\{0\},+,\times)$ est un anneau mais pas un sous-anneau de $(\mathbb{Z}^2,+,\times)$ car il ne contient pas $(1,1)$ qui est l'élément neutre de $\mathbb{Z}^2$ pour $\times$.

FICHE

SYNTHÈSE

Dans ce chapitre, il y a peu de définitions et de propositions. En revanche, il est bon de connaître de nombreux exemples pour chaque structure.

Groupes et sous-groupes

▹ Un ensemble G est un groupe s'il est muni d'une loi de composition interne $\star$ associative, qui admet un élément neutre $e \in G$ et telle que tout élément de G admet un symétrique pour cette loi.
Si, de plus, $\star$ est commutative, alors G est abélien (ou commutatif).
▹ Soit $(G, \star)$ un groupe. Une partie non vide H de G est un sous-groupe de G si, et seulement si,

$$\forall x, y \in H, \qquad x \star y^{-1} \in H.$$

Avec deux lois de composition interne, on trouve une structure adaptée à l'arithmétique dont le principal exemple est $\mathbb{Z}$: les anneaux.

Anneaux et sous-anneaux

▹ Un ensemble A est un anneau s'il est muni de deux lois de composition interne $+$ et $\times$ telles que
- $(A, +)$ est un groupe abélien d'élément neutre noté 0_A,
- $\times$ est associative et admet un élément neutre $1_A \neq 0_A$,
- $\times$ est distributive sur $+$.

Si, de plus, $\times$ est commutative, alors l'anneau est commutatif.
▹ Une partie B d'un anneau $(A, +, \times)$ est un sous-anneau si, et seulement si,
- $(B, +)$ est un sous-groupe de $(A, +)$;
- $1_A \in B$;
- B est stable par multiplication.

Outre sa structure de groupe additif, un anneau cache un autre groupe : l'ensemble des éléments inversibles muni du produit. Lorsque cet ensemble contient tous les éléments non-nuls, la structure est rebaptisée corps.

Corps

Un corps est un anneau commutatif où tout élément non nul admet un symétrique pour le produit.

EXERCICES

Vrai ou faux ?

		Vrai	Faux
a)	L'ensemble des racines $n^{\text{ièmes}}$ de -1 est un groupe pour la multiplication.	☐	☐
b)	Le groupe $\mathbb{U}_d$ est un sous-groupe de $\mathbb{U}_n$ si, et seulement si, $d \mid n$.	☐	☐
c)	Les sous-groupes finis de $\mathbb{C}^*$ sont les $\mathbb{U}_n$.	☐	☐
d)	La partie vide est un sous-groupe de $\mathbb{Z}$.	☐	☐
e)	Le seul sous-groupe de $\mathbb{Z}$ contenant 1 est $\mathbb{Z}$.	☐	☐
f)	La réunion de deux sous-groupes est un sous-groupe.	☐	☐
g)	L'intersection de deux sous-groupes est un sous-groupe.	☐	☐
h)	L'anneau $\mathbb{Z}^2$ muni des opérations terme à terme est intègre.	☐	☐
i)	Soit $1 \leq p \leq n$. $\mathbb{Z}^p$ est un sous-anneau de $\mathbb{Z}^n$.	☐	☐
j)	Les éléments neutres de l'anneau $(\mathscr{P}(E), \Delta, \cup)$ sont $\varnothing$ et E.	☐	☐
k)	$\mathbb{R}$ est un sous-corps de $\mathbb{C}$.	☐	☐
l)	Un corps est un anneau intègre.	☐	☐

EXERCICES

Exercices d'application du chapitre 10

●○○ **Exercice 1**

Montrer que l'ensemble $G =]-1, 1[$ muni de la loi de composition interne définie par $x \oplus y = \frac{x+y}{1+xy}$ est un groupe.

●●○ **Exercice 2**

Soit G un ensemble non vide muni d'une loi associative $\star$ telle que, pour tous a, $b \in G$,

- il existe $x \in G$ tel que $a \star x = b$,
- il existe $y \in G$ tel que $y \star a = b$.

Montrer que G est un groupe pour $\star$.

●●○ **Exercice 3**

Soit $(G, \star)$ un groupe, et $H = \{x \in G,\ \forall g \in G,\ (x \star g)^2 = (g \star x)^2\}$. Montrer que H est un sous-groupe de G.

●●○ **Exercice 4**

Soit $(G,\times)$ un groupe et H une partie finie de G non vide et stable par $\times$. Montrer que H est un sous-groupe de G.

●●● **Exercice 5**

Soit p un nombre premier. Notons G l'ensemble de toutes les racines de l'unité dont l'ordre est une puissance de p c'est-à-dire

$$G = \bigcup_{n\in\mathbb{N}} \mathbb{U}_{p^n}.$$

Rappelons que l'ordre de $z \in G$ est le plus petit exposant $k > 0$ tel que $z^k = 1$; on admet que z est d'ordre k si, et seulement si, tout élément de $\mathbb{U}_k$ est une puissance de z.

1. Montrer que G est un sous-groupe infini de $(\mathbb{C}^*,\times)$.
2. Soit H un sous-groupe de G distinct de G, $z_0 \in G \setminus H$ d'ordre p^{n_0}.
   a. Montrer que si H contient un élément d'ordre p^n, alors $\mathbb{U}_{p^n} \subset H$.
   b. Montrer que $H \subset \mathbb{U}_{p^{n_0}}$.
   c. En déduire qu'il existe n tel que $H = \mathbb{U}_{p^n}$.

●○○ **Exercice 6**

Soit A l'ensemble des rationnels à dénominateur impair en écriture irréductible.

1. Montrer que A est un sous-anneau de $\mathbb{Q}$.
2. Préciser les éléments inversibles de A.

●●○ **Exercice 7**

Montrer que les sous-anneaux de $\mathbb{Z}^2$ muni des opérations terme à terme sont de la forme

$$A_n = \{(x,y) \in \mathbb{Z}^2,\ x = y[n]\}$$

avec $n \in \mathbb{N}$.

●●● **Exercice 8**

Considérons

$$\mathbb{Z}[\sqrt{3}] = \{x + y\sqrt{3},\ (x,y) \in \mathbb{Z}^2\}.$$

1. Montrer que $\mathbb{Z}[\sqrt{3}]$ est un sous-anneau de $\mathbb{R}$.
2. Considérons l'application $N : \left\{ \begin{array}{ccc} \mathbb{Z}[\sqrt{3}] & \to & \mathbb{Z} \\ x + y\sqrt{3} & \mapsto & |x^2 - 3y^2| \end{array} \right.$

   Établir successivement que :
   a. $\forall z, z' \in \mathbb{Z}[\sqrt{3}], \quad N(zz') = N(z)N(z')$.
   b. $\forall z \in \mathbb{Z}[\sqrt{3}], \quad N(z) = 0 \Leftrightarrow z = 0$.
   c. $\forall z \in \mathbb{Z}[\sqrt{3}], \quad N(z) = 1 \Leftrightarrow z$ inversible.
   d. L'équation $x^2 - 3y^2 = -1$ n'admet pas de racines entières.
3. a. En déduire qu'un élément inversible $x + y\sqrt{3}$ est strictement supérieur à 1 si, et seulement si, x et y sont strictement positifs.
   b. En déduire ω le plus petit élément inversible strictement supérieur à 1. $\omega = 2 + \sqrt{3}$.
4. a. Montrer que, pour tout $u \geq 0$ inversible, il existe $n \in \mathbb{Z}$ tel que

   $$\omega^n \leq u < \omega^{n+1},$$

   et montrer que $\omega^{-n} u$ est un inversible dans l'intervalle $[1, \omega[$.
   b. Expliciter l'ensemble des éléments inversibles positifs (puis tous les inversibles) de $\mathbb{Z}[\sqrt{3}]$.

●●● **Exercice 9**

Soit A un anneau tel que $x^3 = x$ pour tout $x \in A$.

1. Montrer que, pour tout $x \in A$, $6x = 0_A$.
2. Soit $B = \{x \in A,\ 2x = 0_A\}$ et $C = \{x \in A,\ 3x = 0_A\}$. Montrer que $A = B + C$.

●○○ **Exercice 10**

Soit $\alpha \in \mathbb{N}$ tel que $\sqrt{\alpha} \notin \mathbb{Q}$. Posons

$$\mathbb{Q}(\sqrt{\alpha}) = \left\{a + b\sqrt{\alpha}, \quad (a,b) \in \mathbb{Q}^2\right\}.$$

1. Soit $x \in \mathbb{Q}(\sqrt{\alpha})$. Montrer qu'il existe un unique couple $(a,b) \in \mathbb{Q}^2$ tel que $x = a + b\sqrt{\alpha}$.
2. Montrer que $\mathbb{Q}(\sqrt{\alpha})$ est un sous-corps de $(\mathbb{R}, +, \times)$.

●○○ **Exercice 11**

Déterminer tous les sous-corps de $(\mathbb{Q}, +, \times)$.

Exercices d'approfondissement du chapitre 10

EXERCICES

●○○ **Exercice A**

Soit G un groupe fini et A, B deux parties de G telles que $|A| + |B| > |G|$. Notons

$$AB = \{ab,\ a \in A,\ b \in B\}.$$

1. Montrer que, pour tout $g \in G$, $A \cap \{gb^{-1}, b \in B\}$ est non vide.
2. Montrer que $G = AB$.

●●○ **Exercice B**

Montrer que, pour tout sous-groupe H de $(\mathbb{Q}, +)$ de la forme $r_1\mathbb{Z} + r_2\mathbb{Z} + \ldots + r_N\mathbb{Z}$ avec $r_1, r_2, \ldots, r_N \in \mathbb{Q}$, il existe $r \in \mathbb{Q}$ tel que $H = r\mathbb{Z}$.

●●● **Exercice C**

Soit G un groupe d'ordre $n \geq 2$. Pour tout $g = (g_1, \ldots, g_k) \in G^k$, notons

- $k = \nu(g)$,
- $A(g)$ l'ensemble des produits $g_{i_1} \cdots g_{i_s}$ où $0 \leq s \leq k$ et $1 \leq i_1 < \cdots < i_s \leq k$,
- $a(g) = |A(g)|$.

Notons enfin E la réunion des G^k pour $k \geq 1$.

1. Soit $g \in E$ tel que $A(g) = G$. Montrer que $\nu(g) \geq \log_2 n$.
2. Soit $g \in E$. Montrer l'existence de $x \in G$ tel que l'élément $g' = (g, x)$ vérifie
$$1 - \frac{a(g')}{n} \leq \left(1 - \frac{a(g)}{n}\right)^2.$$
3. Montrer qu'il existe $g \in E$ tel que $A(g) = G$ et $\nu(g) \leq \lceil \log_2(2n \ln n) \rceil$.

●○○ **Exercice D**

Considérons un anneau $(A, +, \times)$ intègre et fini.

1. Soit $a \in A$ distinct de 0_A. Montrer que l'application
$$\varphi_a : \left\{ \begin{array}{ccc} A & \to & A \\ x & \mapsto & ax \end{array} \right.$$
est bijective
2. En déduire que A est un corps.

●●○ **Exercice E**

Soit A un anneau de cardinal n. Montrer qu'il y a exactement n^2 fonctions polynomiales sur A si, et seulement si, pour tout $x \in A$, $x^2 = x$.

●●● **Exercice F**

Soit A un anneau fini commutatif. Un élément $x \in A$ est nilpotent s'il existe un entier k tel que $x^k = 0_A$; son indice de nilpotence est alors le plus petit entier positif vérifiant cette condition.

1. Montrer que la somme et le produit de deux éléments nilpotents est encore nilpotent.
2. Montrer que le nombre d'éléments nilpotents divise le nombre de diviseurs de zéro.
   Indication : on pourra utiliser le théorème de Lagrange.

Problèmes du chapitre 10

●○○ **Problème 1 Dual d'un groupe**

Pour tout groupe $(G, \times)$, on définit l'ensemble $\widehat{G}$, appelé dual de G, des morphismes de groupes de G dans $\mathbb{C}^*$.

1. Vérifier que, pour tout groupe G, le dual $\widehat{G}$ est un groupe muni du produit des applications.
2. a. Déterminer tous les éléments de $\widehat{\mathbb{U}_3}$.
   b. Montrer que le dual d'un groupe cyclique est cyclique.
3. Soit G et H deux groupes abéliens finis. Montrer que le dual de $G \times H$ est isomorphe à $\widehat{G} \times \widehat{H}$.

●○○ **Problème 2 Théorème de Scorza**

Définition 10.40.

Soit $(G, \star)$ un groupe. Un sous-groupe H est distingué dans G si, pour tout $x \in G$ et tout $h \in H$,

$$x \star h \star x^{-1} \in H.$$

Partie A – Quotient par un sous-groupe distingué

1. Soit $(G, \star)$ un groupe et H un sous-groupe distingué de G. Notons, pour tout $x \in G$,

$$\overline{x} = xH = \{x \star h, h \in H\},$$

puis G/H l'ensemble $\{\overline{x},\ x \in G\}$ muni de la loi de composition interne définie par

$$\forall x, x' \in G, \qquad \overline{x} \star \overline{x'} = \overline{x \star x'}.$$

   a. Vérifier que $\star$ est bien définie que G/H, c'est-à-dire que si $\overline{x} = \overline{y}$ et $\overline{x'} = \overline{y'}$, alors $\overline{x \star x'} = \overline{y \star y'}$.
   b. Montrer que G/H est un groupe puis que l'application

$$\varphi : \left\{ \begin{array}{ccc} G & \to & G/H \\ x & \mapsto & \overline{x} \end{array} \right.$$

   est un morphisme de groupes.
   c. Montrer que pour tout sous-groupe K de G contenant H, $\varphi(K)$ est un sous-groupe de G/H.
   d. Montrer que pour tout sous-groupe K' de G/H, $\varphi^{-1}(K')$ est un sous-groupe de G contenant H.

Partie B – Exemples

1. Écrire les tables de Cayley des groupes $(\mathbb{Z}/2\mathbb{Z},+)$ et $(\mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/2\mathbb{Z},+)$.
2. Vérifier que $\mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/2\mathbb{Z}$ est réunion de trois sous-groupes stricts.

Partie C – Cas de deux sous-groupes

Montrer que si un groupe G est la réunion de deux sous-groupes H_1 et H_2 alors $H_1 = G$ ou $H_2 = G$.

Partie D – Cas de trois sous-groupes

1. Soit $(G,\times)$ un groupe qui est réunion de trois sous-groupes H_1, H_2 et H_3, tous distincts de G.
   On note $A_{1,2,3}$ l'intersection des trois sous-groupes H_1, H_2 et H_3 ; $A_{1,2}$ (respectivement $A_{1,3}$ et $A_{2,3}$) l'intersection des sous-groupes H_1 et H_2 (respectivement H_1 et H_3 ou H_2 et H_3) privée de $A_{1,2,3}$; A_1 (respectivement A_2 et A_3) l'ensemble H_1 (respectivement H_2 et H_3) privé de $H_2\cup H_3$ (respectivement $H_1\cup H_3$ et $H_1\cup H_2$).
   a. Justifier que les parties A_1, A_2 et A_3 sont non vides.
   b. Montrer que $A_{1,2}=\varnothing$.
      Indication : on pourra considérer le produit d'un élément de $A_{1,2}$ avec un élément de A_3 et tester s'il appartient à H_1, H_2 ou H_3.
   c. Montrer que les éléments de A_3 sont les produits d'un élément de A_1 et d'un élément de A_2 (dans cet ordre ou l'autre).
      Écrire toutes les propriétés analogues.
      On admet que le produit de deux éléments de A_1 (respectivement A_2 ou A_3) appartient à $A_{1,2,3}$.
   d. Montrer que $A_{1,2,3}$ est un sous-groupe distingué de G.
      Notons φ la projection de G dans $G/A_{1,2,3}$.
   e. Montrer que $\varphi(A_1)$, $\varphi(A_2)$ et $\varphi(A_3)$ sont trois singletons distincts.
   f. En déduire que le groupe $G/A_{1,2,3}$ est isomorphe à $\mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/2\mathbb{Z}$.
2. Soit G un groupe et H un sous-groupe distingué de G tel que le groupe G/H est isomorphe à $\mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/2\mathbb{Z}$.
   Montrer que G est la réunion de trois sous-groupes de G distincts de G.
   Indication : on pourra utiliser les parties A et B.

PROBLÈMES

Problème 3 Actions de groupe

Définition 10.41.

Une action d'un groupe $(G,.)$ sur un ensemble E est une application

$$\left\{\begin{array}{ccc} G\times E & \longrightarrow & E \\ (g,x) & \longmapsto & g\bullet x \end{array}\right.$$

telle que
- $\forall g_1,g_2\in G,\ \forall x\in E,\qquad (g_1g_2)\bullet x = g_1\bullet(g_2\bullet x)$.
- $\forall x\in E,\qquad e\bullet x = x$.

On définit alors l'orbite de $x\in E$ comme la partie de E

$$\omega(x)=\{y\in E \text{ tel qu'il existe } g\in G \text{ vérifiant } y=g\bullet x\},$$

le stabilisateur de x comme la partie de G

$$S(x)=\{g\in G \text{ tel que } g\bullet x = x\},$$

l'ensemble des points fixes de $g\in G$ est la partie de E

$$\mathrm{Fix}(g)=\{x\in E \text{ tel que } g\bullet x = x\}.$$

Un système de représentants est alors une partie R de E telle que, pour tout $x\in E$, il existe un unique élément à la fois dans $\omega(x)$ et R.

Partie A – Exemples

Déterminer les actions de groupes parmi les applications suivantes.

1. L'action de $\mathbb{Z}$ sur $\mathbb{Q}^*$ définie par
$$\left\{\begin{array}{ccc} \mathbb{Z}\times\mathbb{Q}^* & \to & \mathbb{Q}^* \\ (n,r) & \mapsto & r^n \end{array}\right.$$
2. L'action de G sur G définie par
$$\left\{\begin{array}{ccc} G\times G & \to & G \\ (g,x) & \mapsto & gxg^{-1} \end{array}\right.$$
3. L'action de G sur $\mathscr{P}_n(G)$, l'ensemble des parties de G de cardinal n, définie par
$$\left\{\begin{array}{ccc} G\times\mathscr{P}_n(G) & \to & \mathscr{P}_n(G) \\ (g,A) & \mapsto & gA=\{ga \text{ avec } a\in A\} \end{array}\right.$$

Partie B – Premières propriétés

Considérons une action d'un groupe G sur un ensemble E notée
$$\left\{\begin{array}{ccc} G\times E & \to & E \\ (g,x) & \mapsto & g\bullet x \end{array}\right.$$

1. Définissons une relation binaire sur E par $x\sim y$ si, et seulement si $y\in\omega(x)$.
   a. Vérifier que $\sim$ est une relation d'équivalence.
   b. Préciser la classe d'équivalence pour $\sim$ de $x\in E$.
2. Montrer que, pour tout $x\in E$, $S(x)$ est un sous-groupe de G.
3. Soit $x\in E$. Pour tout $g\in G$, notons $\overline{g}=gS(x)$ et considérons $C=\{\overline{g},\ g\in G\}$.
   a. Montrer que si $h\in\overline{g}$, alors $h\bullet x=g\bullet x$.
   b. Montrer que l'application
   $$\left\{\begin{array}{ccc} C & \to & \omega(x) \\ \overline{g} & \mapsto & g\bullet x \end{array}\right.$$
   est bijective.
   c. En déduire, lorsque G est fini, que $|G|=|S(x)|.|\omega(x)|$.

Partie C – Équation aux classes, formule de Burnside

Considérons une action d'un groupe fini G sur un ensemble fini E notée
$$\left\{\begin{array}{ccc} G\times E & \to & E \\ (g,x) & \mapsto & g\bullet x \end{array}\right.$$

Considérons R un système de représentants.

1. Montrer que
$$|E|=\sum_{x\in R}\frac{|G|}{|S(x)|}.$$
2. Montrer que le nombre d'orbites est
$$\sum_{g\in G}\frac{|\mathrm{Fix}(g)|}{|G|}.$$
Indication : on pourra étudier $\{(g,x)\in G\times E \text{ tel que } g\bullet x=x\}$.

CORRIGÉS

Corrigés des Vrai/Faux

a) Faux, il ne contient même pas 1.

b) Vrai.

c) Vrai. Soit z un élément d'un tel groupe, il existe $p > q$ tel que $z^p = z^q$ (car le groupe est fini) donc $z^{p-q} = 1$. Ainsi, tous les éléments sont des racines de l'unité. On conclut en considérant le plus grand ordre d'un élément de ce groupe.

d) Faux, un groupe est non vide car il contient l'élément neutre.

e) Vrai.

f) Faux, par exemple, $2\mathbb{Z}$ et $3\mathbb{Z}$ sont des sous-groupes de $\mathbb{Z}$ mais $2\mathbb{Z}\cup 3\mathbb{Z}$ n'est pas un sous-groupe de $\mathbb{Z}$: il n'est pas stable par addition : $2+3 \notin 2\mathbb{Z}\cup 3\mathbb{Z}$.

g) Vrai.

h) Faux, par exemple, $(1,0)\times(0,1)=(0,0)$.

i) Faux, il n'y a même pas inclusion.

j) Vrai.

k) Vrai.

l) Vrai.

Corrigés des exercices d'application du chapitre 10

Exercice 1

▹ Montrons d'abord que $\oplus$ est bien interne à G. Pour tous $x, y \in G$, $(1-x)(1-y)>0$ donc $1+xy > x+y$ et $(1+x)(1+y)>0$ donc $1+xy > -(x+y)$: par conséquent, $\frac{x+y}{1-xy} \in G$.

▹ Étudions les propriétés de $\oplus$.

- $0 \in G$ est l'élément neutre de $\oplus$ car, pour tout $x \in G$, $x \oplus 0 = 0 \oplus x = \frac{x+0}{1+0x} = x$.
- Pour tout $x \in G$, $-x \in G$ et

$$x \oplus (-x) = (-x) \oplus x = \frac{x-x}{1-x^2} = 0.$$

Donc tout élément de G admet un symétrique dans G.

- Soit x, y et $z \in G$.

$$x \oplus (y \oplus z) = \frac{x+(y\oplus z)}{1+x(y\oplus z)} = \frac{x+y+z+xyz}{1+xy+yz+zx},$$
$$(x\oplus y)\oplus z = \frac{(x\oplus y)+z}{1+(x\oplus y)z} = \frac{x+y+z+xyz}{1+xy+yz+zx}.$$

Les deux quantités obtenues sont égales donc $x \oplus (y \oplus z) = (x \oplus y) \oplus z$.
Ainsi, $\oplus$ est associative.
En conclusion, $(G, \oplus)$ est un groupe ; il est par ailleurs abélien.

Exercice 2

Comme $\star$ est associative, il suffit d'établir l'existence d'un élément neutre e puis de montrer que tout élément de G admet un symétrique.
▹ Soit $a \in G$ (qui existe car G est non vide). D'après le premier point, il existe $e \in G$ tel que $a \star e = a$. Pour tout $b \in G$, il existe d'après le second point, un élément $y \in G$ tel que $b = y \star a$. Alors,

$$b \star e = (y \star a) \star e = y \star (a \star e) = y \star a = b.$$

On montre de même (en inversant l'ordre d'utilisation des deux propriétés) qu'il existe un élément $e' \in G$ tel que, pour tout $b \in G$, $e' \star b = b$. Mais alors, en utilisant ces deux propriétés tour à tour,

$$e' = e' \star e = e.$$

Ainsi, e est l'élément neutre de $\star$.
▹. Soit $a \in G$. D'après les propriétés, il existe des éléments x, $y \in G$ tels que $a \star x = y \star a = e$. Alors,

$$x = e \star x = (y \star a) \star x = y \star (a \star x) = y \star e = y.$$

Cet élément $x = y$ est donc le symétrique de a pour $\star$.

Exercice 3

Remarquons tout d'abord que H contient l'élément neutre donc est non vide.
Par ailleurs, pour tous x, $y \in H$, et tout $g \in G$, $(x \star y \star g)^2 = (y \star g \star x)^2 = (g \star x \star y)^2$, en utilisant tour à tour $x \in H$ et $y \in H$. Ainsi, $x \star y \in H$.
Enfin, pour tout $x \in H$ et tout $g \in G$,

$$\begin{aligned}(x^{-1} \star g)^2 &= (x^{-1} \star g \star x^{-1} \star x)^2 \\ &= (x \star x^{-1} \star g \star x^{-1})^2 \\ &= (g \star x^{-1})^2.\end{aligned}$$

Ainsi, $x^{-1} \in H$. En conclusion, H est un sous-groupe de G.

Exercice 4

Pour tout $x \in H$, l'ensemble des puissances de x est inclus dans H donc est fini. Il existe donc $p > q$ tel que $x^p = x^q$ et, par conséquent, $x^{p-q} = 1_G$ soit encore $x \times x^{p-q-1} = x^{p-q-1} \times x = 1_G$: le symétrique de x est $x^{p-q-1} \in H$. Ainsi, H est une partie non vide de G stable par produit et par passage à l'inverse donc est un sous-groupe de G.

Exercice 5

1. L'ensemble G est une partie non vide de $\mathbb{C}^*$ car $1 \in G$. Pour tous z, $z' \in G$, il existe $n \in \mathbb{N}$ tel que $z, z' \in \mathbb{U}_{p^n}$ donc $z \times z'^{-1} \in \mathbb{U}_{p^n} \subset G$ et G est donc un sous-groupe de $\mathbb{C}^*$. G est infini car G contient $\mathbb{U}_{p^n}$, une partie de cardinal p^n pour tout n.
2. a. Si H contient un élément z d'ordre p^n, alors tout élément de $\mathbb{U}_{p^n}$ est une puissance de z donc $\mathbb{U}_{p^n} \subset H$.
b. Supposons $H \not\subset \mathbb{U}_{p^{n_0}}$, alors il existe $z \in H$ d'ordre p^n avec $n > n_0$. D'après la question précédente, $\mathbb{U}_{p^{n_0}} \subset \mathbb{U}_{p^n} \subset H$ donc $z_0 \in H$: contradiction.
c. Soit z l'élément d'ordre maximal p^n dans H (qui existe car $H \subset \mathbb{U}_{p^{n_0}}$). Par maximalité, $H \subset \mathbb{U}_{p^n}$; or, $\mathbb{U}_{p^n}$ est l'ensemble des puissances de z donc est inclus dans H. Donc $H = \mathbb{U}_{p^n}$.

Exercice 6

1. ▹ La partie $A \subset \mathbb{Q}$ contient $1 = \frac{1}{1}$ et $0 = \frac{0}{1}$.
▹ Soit a, $a' \in \mathbb{Z}$, n et $n' \in \mathbb{N}$. Alors,

$$\frac{a}{2n+1} - \frac{a'}{2n'+1} = \frac{a(2n'+1) - a'(2n+1)}{(2n+1)(2n'+1)},$$

et le dénominateur en écriture irréductible de cette fraction divise $(2n+1)(2n'+1)$ donc est impair. Ainsi, A est stable par différence.
▹ De même, pour a, $a' \in \mathbb{Z}$, n et $n' \in \mathbb{N}$,

$$\frac{a}{2n+1} \times \frac{a'}{2n'+1} = \frac{aa'}{(2n+1)(2n'+1)}.$$

D'où l'on déduit comme précédemment que A est stable par produit.
En conclusion, A est un sous-anneau de $\mathbb{Q}$.
2. Un élément de A d'écriture irréductible $\frac{a}{2n+1}$ est inversible si le rationnel d'écriture irréductible $\frac{2n+1}{a}$ est aussi dans A, c'est-à-dire si a est impair. Ainsi, les éléments inversibles sont donc les rationnels d'écriture irréductible de la forme $\frac{2m+1}{2n+1}$ (donc, en utilisant le vocabulaire du chapitre suivant, les rationnels de la forme $\frac{p}{q}$ avec la même valuation 2-adique).

Exercice 7

• Remarquons tout d'abord que A_n est bien un sous-anneau de $\mathbb{Z}^2$ car, $(1,1) \in A_n$ et si $x = y[n]$ et $x' = y'[n]$, alors $x - x' = y - y'[n]$ et $xx' = yy'[n]$ soit si (x,y) et $(x',y') \in A_n$, alors $(x,y) - (x',y') \in A_n$ et $(x,y) \times (x',y') \in A_n$.
• Soit A un sous-anneau de $\mathbb{Z}^2$ distinct de A_0. Il existe donc $(x,y) \in A$ avec $x \neq y$. Ensuite $(x-y,0) = (x,y) - y(1,1) \in A$ et $(y-x,0) \in A$. On peut donc poser $n = \min\{x \in \mathbb{N} \setminus \{0\},\ (x,0) \in A\}$. On vérifie alors que $(n,0) \in A$, $(0,n) \in A$ et donc, pour tous x, k, $l \in \mathbb{Z}$,

$$x(1,1) + k(n,0) + l(0,n) \in A.$$

Ainsi, $A_n \subset A$.
Considérons $(x,y) \in A$. Alors, $(x-y,0) \in A$ et par suite $(r,0) \in A$ où r est le reste de la division de $x-y$ par n. D'après la définition de n, $r = 0$ donc $(x,y) \in A_n$. Ainsi, $A \subset A_n$.
En conclusion, les sous-anneaux de $\mathbb{Z}^2$ sont les ensembles A_n où $n \in \mathbb{N}$.

Exercice 8

1. La partie $\mathbb{Z}[\sqrt{3}] \subset \mathbb{R}$ contient 0, 1, est stable par différence et par produit donc est un sous-anneau de $\mathbb{R}$.
2. On remarque que $N(x+y\sqrt{3}) = |(x+y\sqrt{3})(x-y\sqrt{3})|$ pour tout $(x,y) \in \mathbb{Z}^2$.
a. Pour tous (x,y) et $(x',y') \in \mathbb{Z}^2$,

$$\begin{aligned}
N((x+y\sqrt{3})(x'+y'\sqrt{3})) &= N((xx'+3yy')+(xy'+x'y)\sqrt{3}) \\
&= |(xx'+3yy')^2 - 3(xy'+x'y)^2| \\
&= |x^2x'^2 + 9y^2y'^2 - 3x^2y'^2 - 3x'^2y^2| \\
&= |x^2 - 3y^2|.|x'^2 - 3y'^2| \\
&= N(x+y\sqrt{3})N(x'+y'\sqrt{3}).
\end{aligned}$$

b. Soit $(x,y) \in \mathbb{Z}^2$ non nul tel que $N(x+y\sqrt{3}) = 0$, c'est-à-dire $x^2 = 3y^2$ ce qui n'est pas possible en considérant les valuations 3-adiques des deux membres. Par conséquent, le seul $z \in \mathbb{Z}[\sqrt{3}]$ tel que $N(z) = 0$ est $z = 0$.

c. Soit $z = x + y\sqrt{3}$ avec $(x, y) \in \mathbb{Z}^2$.
▹ Si z est inversible, il existe $z' \in \mathbb{Z}[\sqrt{3}]$ tel que $zz' = 1$ et donc $N(z)N(z') = 1$ avec le (a). Ainsi, $N(z)$ est un entier naturel divisant 1 donc $N(z) = 1$.
▹ Réciproquement, si $N(z) = 1$, alors $(x+y\sqrt{3})(x-y\sqrt{3}) = \pm 1$ d'où $x+y\sqrt{3}$ est inversible et $(x+y\sqrt{3})^{-1} = \pm(x-y\sqrt{3})$.
d. Si $(x, y) \in \mathbb{Z}^2$ vérifie $x^2 - 3y^2 = -1$, alors $x^2 = 2[3]$ ce qui est impossible : l'équation $x^2 - 3y^2 = -1$ n'admet pas de racines entières.
On a ainsi établi que $x + y\sqrt{3}$ est un élément inversible si, et seulement si, $x^2 - 3y^2 = 1$.
3. a. ▹ Soit $(x, y) \in \mathbb{Z}^2$ tel que $x + y\sqrt{3}$ est un élément inversible strictement supérieur à 1. L'inverse de cet élément est $x - \sqrt{3}y$ et il est strictement inférieur à 1 donc $x + \sqrt{3}y - (x - \sqrt{3}y) = 2\sqrt{3}y > 0$ donc $y > 0$. Ensuite, comme $x^2 = 1 + 3y^2 > 3y^2$, $x + \sqrt{3}y$ est du signe de x et donc $x > 0$.
▹ Si x et y sont strictement positifs, alors $x + y\sqrt{3} \geq 1 + \sqrt{3} > 1$.
b. Discutons selon la valeur de x.
Si $x \geq 2$ et $y \geq 1$, alors $x + y\sqrt{3} > 2 + \sqrt{3}$.
Par ailleurs, il n'y a pas d'éléments inversibles $x + y\sqrt{3} > 1$ tels que $x = 1$ (car cela entraînerait $y = 0$).
Comme $2 + \sqrt{3}$ est un inversible strictement plus grand que 1, on obtient $\omega = 2 + \sqrt{3}$.
4. a. ▹ En passant au logarithme (qui est une fonction croissante), on observe que l'inégalité équivaut à

$$n \ln \omega \leq \ln u < (n+1) \ln \omega,$$

donc que $n = \left\lfloor \frac{\ln u}{\ln \omega} \right\rfloor$.
▹ Le réel $\omega^{-n} u$ est produit d'éléments inversibles donc est inversible. L'inégalité ci-dessus traduit l'appartenance $\omega^{-n} u$ à l'intervalle $[1, \omega[$.
b. Avec les notations de la question précédente, $\omega^{-n} u = 1$ car ω est le plus petit inversible strictement supérieur à 1 donc $u = \omega^n$. Les inversibles positifs sont donc les puissances de ω. On en déduit que

$$\mathbb{Z}[\sqrt{3}]^* = \{\pm \omega^n,\ n \in \mathbb{Z}\}.$$

Exercice 9

1. Pour tout $x \in A$, la formule du binôme donne (pour x et x qui commutent)

$$x + x = (x + x)^3 = x^3 + 3x^2 x + 3x x^2 + x^3 = x + x + 6x.$$

Ainsi, $6x = 0_A$.
2. Les entiers 2 et 3 sont premiers entre eux donc il existe, d'après la propriété de Bézout, des entiers u et v tels que $1 = 2u + 3v$. Alors, pour tout $x \in A$,

$$x = 2ux + 3vx,$$

et $2ux \in C$ (car $3(2ux) = 6(ux) = 0_A$) et $3vx \in B$ (car $2(3vx) = 6(vx) = 0_A$).
En conclusion, $A \subset B + C$ donc $A = B + C$.

Exercice 10

1. Supposons qu'il existe deux couples (a, b) et $(a', b') \in \mathbb{Q}$ tels que $a + b\sqrt{\alpha} = a' + b'\sqrt{\alpha}$. Alors, $(b - b')\sqrt{\alpha} = (a - a')$. Si $b \neq b'$, alors $\sqrt{\alpha} \in \mathbb{Q}$ ce qui est contraire à la définition de α ; par conséquent, $b = b'$ et par suite, $a = a'$.
2. L'ensemble $\mathbb{Q}(\sqrt{\alpha}) \subset \mathbb{R}$ contient $0 = 0 + 0\sqrt{\alpha}$ et $1 = 1 + 0\sqrt{\alpha}$. Il est bien évidemment stable par différence

$$\forall (a, b), (a', b') \in \mathbb{Q}^2, \qquad (a + b\sqrt{\alpha}) - (a' + b'\sqrt{\alpha}) = (a - a') + (b - b')\sqrt{\alpha} \in \mathbb{Q}(\sqrt{\alpha}),$$

et par produit

$$\forall (a, b), (a', b') \in \mathbb{Q}^2, \qquad (a + b\sqrt{\alpha})(a' + b'\sqrt{\alpha}) = (aa' + \alpha bb') + (ab' + a'b)\sqrt{\alpha} \in \mathbb{Q}(\sqrt{\alpha}).$$

C'est donc un sous-anneau de $\mathbb{R}$. Il reste à établir que tout élément non nul de $\mathbb{Q}(\sqrt{\alpha})$ est inversible dans $\mathbb{Q}(\sqrt{\alpha})$. D'après la première question, un élément non nul de $\mathbb{Q}(\sqrt{\alpha})$ est de la forme $a+b\sqrt{\alpha}$ avec $(a,b)\neq(0,0)$; or, pour $(a,b)\neq(0,0)$,

$$\frac{1}{a+b\sqrt{\alpha}}=\frac{a-b\sqrt{\alpha}}{(a+b\sqrt{\alpha})(a-b\sqrt{\alpha})}=\frac{a-b\sqrt{\alpha}}{a^2-b^2\alpha}\in\mathbb{Q}(\sqrt{\alpha}).$$

En conclusion, $\mathbb{Q}(\sqrt{\alpha})$ est un sous-corps de $\mathbb{R}$.

Exercice 11

Soit K un sous-corps de $\mathbb{Q}$. Par définition de corps, $1\in\mathbb{K}$ et par stabilité par somme, $\mathbb{Z}\subset K$. Or, tout élément non nul de K admet son inverse dans K donc, pour tout $n\in\mathbb{N}\setminus\{0\}$, $\frac{1}{n}\in K$. Enfin, par stabilité par produit, pour tout $m\in\mathbb{Z}$ et tout $n\in\mathbb{N}\setminus\{0\}$, $\frac{m}{n}\in K$ donc $\mathbb{Q}=K$.
Le seul sous-corps de $\mathbb{Q}$ est $\mathbb{Q}$.

Corrigés des exercices d'approfondissement du chapitre 10

Exercice A

1. Soit $g\in G$ tel que $A\cap\{gb^{-1},b\in B\}$ soit vide. Alors,

$$\begin{aligned}|A\cup\{gb^{-1},b\in B\}|&=|A|+|\{gb^{-1},b\in B\}||\\&=|A|+|B|>|G|,\end{aligned}$$

ce qui est impossible pour une partie de G. Par conséquent, l'intersection $A\cap\{gb^{-1},b\in B\}$ est non vide pour tout $g\in G$.
2. Soit $g\in G$ et $a\in A\cap\{gb^{-1},b\in B\}$; par définition, il existe $b\in B$ tel que $a=gb^{-1}$ donc $g=ab$. On a ainsi $G\subset AB$. L'autre inclusion est immédiate car le produit est interne à G.

Exercice B

Soit H un tel sous-groupe. Notons $r_i=\frac{p_i}{q_i}$ pour $i\in[\![1,N]\!]$ puis q le PPCM des dénominateurs et enfin, $q=q_i'q_i$ pour tout $i\in[\![1,N]\!]$. Alors

$$H=\frac{1}{q}(q_1'p_1\mathbb{Z}+\ldots+q_N'p_N\mathbb{Z}).$$

Or, tous les sous-groupes de $\mathbb{Z}$ sont de la forme $a\mathbb{Z}$ avec $a\in\mathbb{Z}$: il existe $a\in\mathbb{Z}$ tel que

$$q_1'p_1\mathbb{Z}+\ldots+q_N'p_N\mathbb{Z}=a\mathbb{Z}.$$

En conclusion, H est engendré par $\frac{a}{q}$.

Exercice C

1. Majorons le nombre d'éléments de $A(g)$ par le nombre de parties de $\{1,\ldots,\nu(g)\}$. Comme $A(g)=G$, $n=|A(g)|\leq 2^{\nu(g)}$, c'est-à-dire $\nu(g)\geq\log_2(n)$.
2. Remarquons que, pour tout $x\in G$, $A(g')=A(g)x\cup A(g)$ en distinguant les cas où l'élément de $A(g')$ utilise ou non la dernière coordonnée. Par conséquent, $a(g')=2a(g)-|A(g)x\cap A(g)|$ car $|A(g)x|=|A(g)|=a(g)$.

Considérons $x \in G$ tel que le cardinal $|A(g)x \cap A(g)|$ est minimal. Alors,

$$n|A(g)x \cap A(g)| \leq \sum_{y \in G} |A(g)y \cap A(g)|.$$

Comme $A(g)y \cap A(g) \subset A(g)$ et que chaque élément $h \in A(g)$ appartient à $A(g)y \cap A(g)$ si, et seulement s'il existe $k \in A(g)$ tel que $y = k^{-1}h$, le membre de droite compte chaque élément de $A(g)$ exactement $|A(g)|$ fois. Ainsi,

$$n|A(g)x \cap A(g)| \leq a(g)^2.$$

Avec cette inégalité, on obtient

$$1 - \frac{a(g')}{n} \leq 1 - \frac{2a(g)}{n} + \frac{a(g)^2}{n^2} = \left(1 - \frac{a(g)}{n}\right)^2.$$

3. Considérons la suite $(g^{(k)})_{k \in \mathbb{N}}$ d'éléments de E obtenue par extensions successives avec la question précédente. Par une récurrence immédiate, pour tout $k \in \mathbb{N}$,

$$1 - \frac{a(g^{(k)})}{n} \leq \left(1 - \frac{1}{n}\right)^{2^k}.$$

Pour tout $k \geq \log_2(2n \ln n)$,

$$\left(1 - \frac{1}{n}\right)^{2^k} < e^{-\frac{2^k}{n}} \leq \frac{1}{n},$$

donc $a(g^{(k)}) > n - 1$, c'est-à-dire $a(g^{(k)}) = n$. On a ainsi comme attendu, que pour $k = \lceil \log_2(2n \ln n) \rceil$, $A(g^{(k)}) = G$.

Exercice D

1. L'application φ_a est injective car A est intègre : en effet, pour tous b, $c \in A$ tels que $a \times b = a \times c$, on a $a \times (b - c) = 0_A$ donc $b - c = 0_A$ par intégrité (et car $a \neq 0_A$).
Comme l'ensemble A est fini, une application injective de A dans A est bijective.
2. Soit $a \neq 0_A$. Comme φ_a est bijective, il existe $b \in A$ tel que $\varphi_a(b) = 1_A$ c'est-à-dire $a \times b = 1_A$. L'anneau étant commutatif (car intègre), on en déduit que a est inversible (d'inverse b). Ainsi, tout élément non nul de A est inversible donc A est un corps.

Exercice E

Commençons par remarquer qu'il y a au moins n^2 fonctions polynomiales $f_{a,b} : x \mapsto ax + b$ avec a, $b \in A$: en effet, si $f_{a,b} = f_{c,d}$, alors $b = f_{a,b}(0_A) = f_{c,d}(0_A) = d$ et $a = f_{a,b}(1_A) - b = f_{c,d}(1_A) - d = c$.
▹ Si $x^2 = x$ pour tout $x \in A$, alors $x^k = x$ pour tout $k \geq 1$ (simple récurrence) et

$$\sum_{k=0}^{n} a_k x^k = \left(\sum_{k=1}^{n} a_k\right) x + a_0 = f_{\sum_{k=1}^{n} a_k, a_0}(x).$$

Par conséquent, les seules fonctions polynomiales sont associées aux fonctions affines.
▹ Supposons qu'il n'y a pas d'autres fonctions polynomiales. Alors, la fonction $x \mapsto x^2$ est égale à une fonction affine : il existe a, $b \in A$ tels que, pour tout $x \in A$, $x^2 = ax + b$. En considérant $x = 0_A$ puis $x = 1_A$, on obtient $b = 0_A$ et $a = 1_A$.

Exercice F

1. Soit x et y nilpotents d'indices n et m. Alors, $(xy)^n = x^n y^n = 0$ et

$$(x+y)^{m+n-1} = \sum_{k=0}^{m+n-1} \binom{m+n-1}{k} x^k y^{m+n-1-k} = 0.$$

En effet, si $k \geq n$, $x^k = 0$; sinon $m+n-1-k \geq m$ et, dans ce cas, $y^{m+n-1-k} = 0$: tous les termes de la somme sont nuls.

2. Notons N, D les ensembles des éléments nilpotents et diviseurs de zéro de A.

▹ N est un sous-groupe (d'après la première question) de $(A,+)$ donc d'après le théorème de Lagrange, $|N|$ divise $|A|$.

▹ L'ensemble $\{1+x,\ x \in N\}$ est un sous-groupe de $(A^*, \times)$ (car, pour tout $x \in N$, $(1-x)(1+x+x^2+\ldots+x^{n-1}) = 1$; la stabilité par produit vient de la première question). D'après le théorème de Lagrange, $|N|$ divise $|A^*|$.

▹ Montrons que $D = \overline{A^*}$. Il est evident que $D \subset \overline{A^*}$. Pour l'inclusion réciproque, considérons $x \notin A^*$ et l'application

$$\varphi : \left\{ \begin{array}{ccc} A & \to & A \\ a & \mapsto & a \times x \end{array} \right.$$

φ n'est pas surjective car 1_A n'est pas atteint donc n'est pas injective (car A est fini) : soit $a \neq b$ tel que $a \times x = b \times x$ donc $(a-b) \times x = 0_A$: x appartient à D. Par conséquent, $|A^*| + |D| = |A|$.

En combinant ces trois points, on trouve que $|N|$ divise $|A| - |A^*|$ donc $|D|$.

Corrigés des problèmes du chapitre 10

Problème 1

1. Le produit des applications est bien une loi de composition interne. Cette loi de composition est clairement associative ; l'élément neutre est l'application $f : G \to \mathbb{C}^*$ constante égale à 1. Le morphisme f de G dans $\mathbb{C}^*$ admet pour symétrique l'application

$$\left\{ \begin{array}{ccc} G & \to & \mathbb{C}^* \\ x & \mapsto & \frac{1}{f(x)}. \end{array} \right.$$

Ainsi, $\widehat{G}$ est un groupe (et il est abélien).

2. a. Soit $\varphi \in \widehat{\mathbb{U}_3}$. Alors $\varphi(j)^3 = \varphi(1) = 1$ donc $\varphi(j) \in \mathbb{U}_3$. Remarquons ensuite que la valeur $\varphi(j)$ détermine uniquement φ car $\varphi(1) = 1$ et $\varphi(j^2) = \varphi(j)^2$. Il y a donc trois éléments dans $\widehat{\mathbb{U}_3}$:

$$\left\{ \begin{array}{ccc} 1 & \mapsto & 1 \\ j & \mapsto & 1 \\ j^2 & \mapsto & 1 \end{array} \right. \qquad \left\{ \begin{array}{ccc} 1 & \mapsto & 1 \\ j & \mapsto & j \\ j^2 & \mapsto & j^2 \end{array} \right. \qquad \left\{ \begin{array}{ccc} 1 & \mapsto & 1 \\ j & \mapsto & j^2 \\ j^2 & \mapsto & j \end{array} \right.$$

b. Soit $G = \langle x \rangle$ d'ordre n. Toute application $\varphi \in \widehat{G}$ est uniquement déterminée par $\varphi(x)$. Comme précédemment, $\varphi(x)^n = \varphi(x^n) = \varphi(1) = 1$ c'est-à-dire $\varphi(x) \in \mathbb{U}_n$. Notons φ_k l'application qui envoie x sur $e^{\frac{2ik\pi}{n}}$. Alors, pour tout k, $\varphi_k = (\varphi_1)^k$ donc $\widehat{G} = \langle \varphi_1 \rangle$ est cyclique.

3. Notons

$$f_1 : \left\{ \begin{array}{ccc} G & \to & G \times H \\ g & \mapsto & (g,e) \end{array} \right. , \qquad f_2 : \left\{ \begin{array}{ccc} H & \to & G \times H \\ h & \mapsto & (e,h) \end{array} \right.$$

et

$$p_1 : \left\{ \begin{array}{ccc} G \times H & \to & G \\ (g,h) & \mapsto & g \end{array} \right. , \qquad p_2 : \left\{ \begin{array}{ccc} G \times H & \to & H \\ (g,h) & \mapsto & h \end{array} \right.$$

Alors, l'application $\left\{\begin{array}{ccc} \widehat{G\times H} & \to & \widehat{G}\times\widehat{H} \\ \varphi & \mapsto & (\varphi\circ f_1, \varphi\circ f_2) \end{array}\right.$

est un isomorphisme de bijection réciproque $\left\{\begin{array}{ccc} \widehat{G}\times\widehat{H} & \to & \widehat{G\times H} \\ (\varphi_1,\varphi_2) & \mapsto & \varphi_1\circ p_1.\varphi_2\circ p_2 \end{array}\right.$

En conclusion, $\widehat{G\times H}$ est isomorphe à $\widehat{G}\times\widehat{H}$.

Problème 2

Partie A – Quotient par un sous-groupe distingué

1. a. Soit x, x', y, $y'\in G$ tels que $\overline{x}=\overline{y}$ et $\overline{x'}=\overline{y'}$. Par définition, il existe h, $h'\in H$ tels que $x=y\star h$ et $x'=y'\star h'$. Ainsi

$$x\star x'=y\star h\star y'\star h'=y\star y'\star y'^{-1}\star h\star y'\star h'.$$

Or, comme H est distingué, $y'^{-1}\star h\star y'\in H$ donc $x\star x'\in\overline{y\star y'}$. Par symétrie, $\overline{x\star x'}=\overline{y\star y'}$.
b. Vérifions les différentes propriétés de la loi de composition interne sur G/H. Elle est associative (simple vérification), d'élément neutre $\overline{e}$ et le symétrique de $\overline{x}$ (où $x\in G$) est $\overline{x^{-1}}$.
Pour tous $x,y\in G$, $\varphi(x\star y)=\varphi(x)\star\varphi(y)$ par définition de la loi sur G/H.
En conclusion, G/H est un groupe et φ est un morphisme de groupes.
c. Il suffit d'appliquer la première question avec le morphisme φ.
d. D'après la première question, pour tout sous-groupe K' de G/H, $\varphi^{-1}(K')$ est un sous-groupe de G. De plus, $\overline{e}\in K'$ donc $\varphi^{-1}(\{\overline{e}\})\subset\varphi^{-1}(K')$; or $\varphi^{-1}(\{\overline{e}\})=\{x\in G,\ \overline{x}=\overline{e}\}=H$. En conclusion, $\varphi^{-1}(K')$ est un sous-groupe de G contenant H.

Partie B – Exemples

1. Après calculs, on obtient successivement

$+$	$\overline{0}$	$\overline{1}$
$\overline{0}$	$\overline{0}$	$\overline{1}$
$\overline{1}$	$\overline{1}$	$\overline{0}$

$+$	$(\overline{0},\overline{0})$	$(\overline{0},\overline{1})$	$(\overline{1},\overline{0})$	$(\overline{1},\overline{1})$
$(\overline{0},\overline{0})$	$(\overline{0},\overline{0})$	$(\overline{0},\overline{1})$	$(\overline{1},\overline{0})$	$(\overline{1},\overline{1})$
$(\overline{0},\overline{1})$	$(\overline{0},\overline{1})$	$(\overline{0},\overline{0})$	$(\overline{1},\overline{1})$	$(\overline{1},\overline{0})$
$(\overline{1},\overline{0})$	$(\overline{1},\overline{0})$	$(\overline{1},\overline{1})$	$(\overline{0},\overline{0})$	$(\overline{0},\overline{1})$
$(\overline{1},\overline{1})$	$(\overline{1},\overline{1})$	$(\overline{1},\overline{0})$	$(\overline{0},\overline{1})$	$(\overline{0},\overline{0})$

2. D'après le théorème de Lagrange, les sous-groupes de $\mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/2\mathbb{Z}$ sont d'ordre 4 (le groupe entier), 2 (les groupes $\{(\overline{0},\overline{0}),(\overline{0},\overline{1})\}$, $\{(\overline{0},\overline{0}),(\overline{1},\overline{0})\}$, $\{(\overline{0},\overline{0}),(\overline{1},\overline{1})\}$) et le groupe d'ordre 1 (réduit à l'élément neutre). On vérifie immédiatement que $\mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/2\mathbb{Z}$ est la réunion des trois groupes d'ordre 2.

Partie C – Cas de deux sous-groupes

Supposons que $(G,\star)$ est la réunion de deux sous-groupes H_1 et H_2 et que $H_1\neq G$. Alors, il existe $x\in G\setminus H_1$. Pour tout $y\in H_1$, $x\star y\in G=H_1\cup H_2$. Si $x\star y\in H_1$, alors $x=(x\star y)\star y^{-1}\in H_1$ ce qui est contradictoire avec la définition de x. Par conséquent, $x\star y\in H_2$ donc $y=x^{-1}\star(x\star y)\in H_2$ ainsi, $H_1\subset H_2$ et $G=H_1\cup H_2=H_2$.

Partie D – Cas de trois sous-groupes

1. a. Si l'une des parties A_1, A_2 et A_3 est vide, alors G est réunion de deux sous-groupes stricts ce qui est impossible d'après la partie précédente.
b. Supposons $A_{1,2}\neq\varnothing$. Soit $x\in A_{1,2}$ et $y\in A_3$. Alors $x\star y\in G=H_1\cup H_2\cup H_3$.
- Si $x\star y\in H_1$, alors $y=x^{-1}\star(x\star y)\in H_1$ ce qui contredit $y\in A_3$.
- Si $x\star y\in H_2$, alors $y=x^{-1}\star(x\star y)\in H_2$ ce qui contredit $y\in A_3$.
- Si $x\star y\in H_3$, alors $x=(x\star y)\star y^{-1}\in H_3$ ce qui contredit $x\in A_{1,2}$.

Dans tous les cas, on aboutit à une contradiction donc $A_{1,2} = \varnothing$.
De la même manière, on obtient $A_{1,3} = A_{2,3} = \varnothing$.
c. Soit $x \in A_3$ et $y \in A_1$. En distinguant les cas comme précédemment, on trouve que le produit $x \star y^{-1}$ appartient à H_2 mais pas à $H_1 \cup H_3$ ce qui, par définition, signifie $x \star y^{-1} \in A_2$. Ainsi $x = (x \star y^{-1}) \star y \in A_2 A_1$.
De même

$$A_1 = A_2 A_3 = A_3 A_2, \qquad A_2 = A_1 A_3 = A_3 A_1, \qquad A_3 = A_1 A_2 = A_2 A_1.$$

d. Remarquons tout d'abord que $A_{1,2,3}$ est un sous-groupe comme intersections de trois sous-groupes.
Soit $x \in A_{1,2,3}$ et $y \in G$.
• Si $y \in A_{1,2,3}$, alors $y \star x \star y^{-1} \in A_{1,2,3}$ par propriété de groupe.
• Sinon, supposons sans perte de généralité que $y \in A_1$ et remarquons que $x \star y^{-1} \in H_1$ par propriété de groupe et exploitons $H_1 = A_1 \cup A_{1,2,3}$.
– Si $x \star y^{-1} \in A_{1,2,3}$, alors $y^{-1} \in A_{1,2,3}$ (en multipliant par x^{-1} donc $y \in A_{1,2,3}$: contradiction!
– Si $x \star y^{-1} \in A_1$, alors $y \star x \star y^{-1}$ est le produit de deux éléments de A_1 donc appartient à $A_{1,2,3}$.
Dans tous les cas de figure, $y \star x \star y^{-1} \in A_{1,2,3}$. En conclusion, $A_{1,2,3}$ est un sous-groupe distingué de G.
e. • Commençons par montrer que le symétrique d'un élément de $y \in A_1$ appartient à A_1 : en effet, il appartient à H_1 et ne peut appartenir au groupe $A_{1,2,3}$ car sinon $y = (y^{-1})^{-1} \in A_{1,2,3}$.
Ainsi, pour tous $x, y \in A_1$, $y^{-1} \star x \in A_{1,2,3}$ (produit de deux éléments de A_1), c'est-à-dire qu'il existe $z \in A_{1,2,3}$ tel que $x = y \star z$ donc $\overline{x} = \overline{y}$.
En conclusion, tous les éléments de A_1 ont la même image par φ : $\varphi(A_1)$ est un singleton.
• Soit $x \in A_1$ et $y \in A_2$. Si $\overline{x} = \overline{y}$, alors il existe $z \in A_{1,2,3}$ tel que $x = y \star z$. Comme $y, z \in H_2$, alors $x \in H_2$ qui est contradictoire avec $x \in A_1$.
Ainsi les images par φ d'éléments de A_1 et A_2 sont distinctes.
Par symétrie, $\varphi(A_1)$, $\varphi(A_2)$ et $\varphi(A_3)$ sont trois singletons distincts.
f. Notons e, a_1, a_2 et a_3 les images des éléments de $A_{1,2,3}$, A_1, A_2 et A_3 respectivement. En utilisant les propositions de la question (c), on obtient la table de Cayley suivante

$\star$	e	a_1	a_2	a_3
e	e	a_1	a_2	a_3
a_1	a_1	e	a_3	a_2
a_2	a_2	a_3	e	a_1
a_3	a_3	a_2	a_1	e

qui correspond à celle de $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$. Donc les deux groupes sont isomorphes.
2. Soit $f : \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \to G/H$ un isomorphisme et considérons $H_1 = \varphi^{-1}(f(\{(\overline{0},\overline{0}),(\overline{0},\overline{1})\}))$, $H_2 = \varphi^{-1}(f(\{(\overline{0},\overline{0}),(\overline{1},\overline{0})\}))$ et $H_3 = \varphi^{-1}(f(\{(\overline{0},\overline{0}),(\overline{1},\overline{1})\}))$. D'après la partie A, ces trois ensembles sont des sous-groupes de G (contenant H) comme image et image réciproque de sous-groupes par des morphismes. Comme

$$f(\{(\overline{0},\overline{0}),(\overline{0},\overline{1})\}) \cup f(\{(\overline{0},\overline{0}),(\overline{1},\overline{0})\}) \cup f(\{(\overline{0},\overline{0}),(\overline{1},\overline{1})\}) = G/H,$$

on a $H_1 \cup H_2 \cup H_3 = G$. De plus, ces trois groupes sont distincts de G car leur image par φ n'est pas G/H.

Problème 3

Partie A – Exemples

Les deux dernières applications sont des actions de groupes.
1. Soit $r \in \mathbb{Q}^*$ différent de 1. Alors, $r^{1+1} \neq (r^1)^1$.
2. ▹ $\forall g, h \in G,\ \forall x \in G, \qquad (gh)x(gh)^{-1} = g(hxh^{-1})g^{-1}$.
▹ $\forall x \in G, \qquad exe^{-1} = x$.
3. ▹ $\forall g, h \in G,\ \forall A \in \mathscr{P}_n(G), \qquad (gh)A = \{gha \text{ avec } a \in A\} = g\{ha \text{ avec } a \in A\} = g(hA)$.
▹ $\forall A \in \mathscr{P}_n(G), \qquad eA = \{ea, \text{ avec } a \in A\} = A$.

Partie B – Premières propriétés

1. a. ▹ $\sim$ est réflexive car $x \in \omega(x)$ pour tout $x \in E$.
▹ Soit $x, y \in E$.

$$\begin{aligned} y \in \omega(x) &\Leftrightarrow \exists g \in G, \quad y = g \bullet x \\ &\Leftrightarrow \exists g \in G, \quad g^{-1} \bullet y = g^{-1} \bullet (g \bullet x) \\ &\Leftrightarrow \exists g \in G, \quad g^{-1} \bullet y = (g^{-1} g) \bullet x = x \\ &\Leftrightarrow x \in \omega(y). \end{aligned}$$

D'où $\sim$ est symétrique.
▹ Soit $x, y, z \in E$ tels que $x \sim y$ et $y \sim z$; il existe $g, h \in G$ tels que $y = g \bullet x$ et $z = h \bullet y$ donc

$$z = h \bullet (g \bullet x) = (hg) \bullet x.$$

Ainsi, $z \sim x$. D'où $\sim$ est transitive. En conclusion, $\sim$ est une relation d'équivalence.
b. Par définition, la classe d'équivalence de $x \in E$ est $\omega(x)$.
2. ▹ $e \in S(x)$ donc $S(x)$ est non vide.
▹ Soit $g, h \in S(x)$. Alors $g \bullet x = x$ et $h \bullet x = x$ donc $x = h^{-1} \bullet x$. D'où

$$(gh^{-1}) \bullet x = g \bullet (h^{-1} \bullet x) = g \bullet x = x.$$

Ainsi, $gh^{-1} \in S(x)$. En conclusion, $S(x)$ est un sous-groupe de G.
3. a. Si $h \in \overline{g}$, il existe $k \in S(x)$ tel que $h = gk$. Alors

$$h \bullet x = (gk) \bullet x = g \bullet (k \bullet x) = g \bullet x \text{ car } k \in S(x).$$

b. ▹ Par définition de $\omega(x)$, l'application est surjective.
▹ Soit $g, h \in G$ tels que $g \bullet x = h \bullet x$; d'où $(h^{-1} g) \bullet x = x$, c'est-à-dire $h^{-1} g \in S(x)$ ou encore $g \in hS(x)$ soit enfin $\overline{g} = \overline{h}$. L'application est donc injective.
c. D'après ce qui précède, $|C| = |\omega(x)|$. De plus, les ensembles distincts $\overline{g}$ forment une partition de G et sont tous de même cardinal $|S(x)|$. Il y a $|C|$ tels ensembles d'où

$$|G| = |C|.|\overline{g}| = |\omega(x)|.|S(x)|.$$

Partie C – Équation aux classes, formule de Burnside

1. Les orbites $\omega(x)$ pour $x \in R$ forment une partition de E (car ce sont les classes d'équivalence pour $\sim$) d'où

$$|E| = \sum_{x \in R} |\omega(x)|.$$

En utilisant la dernière question de la partie B,

$$|E| = \sum_{x \in R} \frac{|G|}{|S(x)|}.$$

2. Soit $A = \{(g, x) \in G \times E \text{ tel que } g \bullet x = x\}$. Dénombrons A de deux manières. D'une part, en partitionnant selon la première coordonnée des couples,

$$A = \sum_{g \in G} |\mathrm{Fix}(g)|.$$

D'autre part, en partitionnant selon la deuxième coordonnée des couples,

$$A = \sum_{x \in E} |S(x)| = \sum_{x \in R} \sum_{y \in \omega(x)} |S(y)| = \sum_{x \in R} \sum_{y \in \omega(x)} \frac{|G|}{|\omega(x)|} = |G|.|R|.$$

En conclusion, le nombre d'orbites est

$$|R| = \sum_{g \in G} \frac{|\mathrm{Fix}(g)|}{|G|}.$$

Chapitre 11

Arithmétiques des entiers

1. Divisibilité — **2.** PGCD, PPCM — **3.** Nombres premiers — **4.** Systèmes congruentiels

Objectifs et compétences du programme

- Maîtriser la notion de divisibilité et les écritures en termes de congruence.
- Savoir caractériser la primalité relative par le théorème de Bézout.
- Calculer un PGCD et une relation de Bézout.
- Exploiter la décomposition en facteurs premiers.
- Résoudre des systèmes linéaires congruentiels.

Sophie Germain, mathématicienne, physicienne et philosophe française (1776-1831).
Davantage connue pour sa légende que pour ses résultats, sa condition de femme l'empêcha d'accéder aux études mathématiques (les cours lui sont interdits) et elle dut usurper une identité d'homme pour se livrer à sa passion.
Ses travaux ont suscité (et suscitent encore) un grand intérêt :
- d'une part, ses recherches arithmétiques sur le théorème de Fermat (alors conjecture) l'ont amené à traiter un cas particulier impressionnant (celui des exposants n premiers impairs tels que $2n+1$ est également premier) ;
- d'autre part, son analyse des figures obtenues sur les plaques vibrantes, dites figures de Chladni, lui ont valu le Grand Prix de l'Académie des Sciences en 1816.

COURS

L'arithmétique ou théorie des nombres est sûrement la partie la plus fascinante des mathématiques en ce sens qu'elle permet d'énoncer très simplement des questions très difficiles. Dans ce chapitre, l'objectif n'est pas d'aborder les grandes conjectures mais de mettre en avant les éléments de base : divisibilité entre entiers, primalité relative (et calcul de PGCD), primalité.
On remarquera que les nombres premiers arrivent en fin de chapitre et que les autres notions sont donc définies indépendamment de ceux-ci.

1. Divisibilité

1.1. Définitions et premiers exemples

Définition 11.1.

Soit a, $b \in \mathbb{Z}$.
▹ a est un multiple de b s'il existe $q \in \mathbb{Z}$ tel que $a = qb$.
▹ a est un diviseur de b s'il existe $q \in \mathbb{Z}$ tel que $qa = b$.

Ainsi, un entier a est un diviseur d'un entier b si, et seulement si, b est un multiple de a. On note alors $a|b$ pour l'affirmation a divise b.

Définition 11.2.

Soit $n \in \mathbb{N} \setminus \{0\}$. La congruence modulo n est la relation binaire sur $\mathbb{Z}$ définie par la divisibilité de la différence par n.
Plus précisément, pour tous a, $b \in \mathbb{Z}$, $a = b[n]$ si $a - b \in n\mathbb{Z}$.

La relation de congruence est également notée $\equiv$ même si on lui a préféré le symbole d'égalité dans ce cours.

Proposition 11.3.

Soit $n \in \mathbb{N} \setminus \{0\}$. La congruence modulo n est une relation d'équivalence.

Remarque
La propriété « a est un multiple de b » se traduit également par $a\mathbb{Z} \subset b\mathbb{Z}$ et peut s'écrire $a = 0[b]$.

Exemple

Pour tout entier n impair, $n^2 - 1$ est un multiple de 8.
En effet, en écrivant $n = 2k + 1$ avec $k \in \mathbb{Z}$, on obtient $n^2 - 1 = 4k(k+1)$; or l'un des deux entiers consécutifs k ou $k+1$ est pair donc $n^2 - 1$ est un multiple de $4 \times 2 = 8$.

Rappelons les propriétés opératoires de la congruence.

Proposition 11.4.

Soit $n \in \mathbb{N} \setminus \{0\}$, a, b, c, $d \in \mathbb{Z}$ tels que $a = b[n]$ et $c = d[n]$. Alors, $a + c = b + d[n]$ et $ac = bd[n]$.

Démonstration

▹ Par hypothèse, n divise $a - b$ et $c - d$ donc divise $(a + c) - (b + d)$.

▹ Par hypothèse, n divise $a - b$ et $c - d$ donc divise $(a - b)c + b(c - d) = ac - bd$. ■

On en déduit immédiatement par récurrence sur l'exposant le résultat suivant.

Corollaire 11.5.

Soit $n \in \mathbb{N} \setminus \{0\}$, $k \in \mathbb{N}$, a, $b \in \mathbb{Z}$ tels que $a = b[n]$. Alors, $a^k = b^k[n]$.

Exemple

▹ Pour tout entier naturel n, $2^{3n} - 1$ est un multiple de 7.
En effet, $2^3 = 1 + 7 = 1[7]$ donc $2^{3n} = 1[7]$ et donc $2^{3n} - 1 = 0[7]$.
▹ Plus généralement (le cas précédent correspond aux valeurs $a = 3$, $b = 3n$) : si $a \in \mathbb{N}$ divise $b \in \mathbb{N}$, alors $2^a - 1$ divise $2^b - 1$.

Exemple

Pour tout entier n, $10^{9n+2} + 10^{6n+1} + 1$ est un multiple de 111.
En effet, $10^3 = 1 + 9 \times 111 = 1[111]$ donc

$$\begin{aligned} 10^{9n+2} &= (10^3)^{3n} 10^2 = 100[111], \\ 10^{6n+1} &= (10^3)^{2n} 10 = 10[111], \\ 1 &= 1[111]. \end{aligned}$$

En conclusion, $10^{9n+2} + 10^{6n+1} + 1 = 100 + 10 + 1[111] = 0[111]$.

Rappelons enfin les critères usuels de divisibilité basés sur l'écriture en base 10 des entiers.

Proposition 11.6.

Soit $n = \sum_{k=0}^{r} a_k 10^k$ avec, pour tout $k \in [\![0, r]\!]$, $a_k \in [\![0, 9]\!]$.
▹ n est un multiple de 2 si, et seulement si, a_0 est un multiple de 2.
▹ n est un multiple de 3 si, et seulement si, la somme des chiffres $\sum_{k=0}^{r} a_k$ est un multiple de 3.
▹ n est un multiple de 4 si, et seulement si, $10a_1 + a_0$ est un multiple de 4.
▹ n est un multiple de 5 si, et seulement si, a_0 est un multiple de 5.
▹ n est un multiple de 9 si, et seulement si, la somme des chiffres $\sum_{k=0}^{r} a_k$ est un multiple de 9.
▹ n est un multiple de 11 si, et seulement si, la somme alternée des chiffres $\sum_{k=0}^{r} a_k(-1)^k$ est un multiple de 11.

Démonstration

Il suffit de remarquer que $10 = 0[2]$, $10 = 1[3]$, $100 = 0[4]$, $10 = 0[5]$, $10 = 1[9]$ et $10 = -1[11]$ pour obtenir le reste de 10^k modulo le diviseur concerné. En utilisant les règles opératoires de la congruence, on obtient

▹ $n = a_0[2]$ car $10^k = 0[2]$ pour tout $k \geq 1$,

▹ $n = \sum_{k=0}^{r} a_k[3]$ car $10^k = 1[3]$ pour tout k,

▹ $n = 10a_1 + a_0[4]$ car $10^k = 0[4]$ pour tout $k \geq 2$,
▹ $n = a_0[5]$ car $10^k = 0[4]$ pour tout $k \geq 1$,
▹ $n = \sum_{k=0}^{r} a_k[9]$ car $10^k = 1[9]$ pour tout k,
▹ $n = \sum_{k=0}^{r} a_k(-1)^k[11]$ car $10^k = (-1)^k[11]$ pour tout k. ■

Exemple

L'entier 14146 est divisible par 11 (car $1-4+1-4+6 = 0[11]$) mais pas par 3 (car $1+4+1+4+6 = 1[3]$).

1.2. Division euclidienne

On cherche désormais à écrire la division entre entiers avec quotient et reste.

Proposition 11.7.

Soit $(n, a) \in \mathbb{Z} \times \mathbb{N} \setminus \{0\}$. Il existe un unique couple $(q, r) \in \mathbb{Z} \times \mathbb{N}$ tel que

$$n = qa + r \qquad 0 \leq r \leq a - 1.$$

Les entiers q et r sont respectivement appelés quotient et reste de la division euclidienne de n par a.

Démonstration

- L'unicité est immédiate : si $n = qa + r$ et $n = q'a + r'$ avec $0 \leq r' \leq r < a$, alors $a > r - r' \geq 0$ et $r - r' = (q' - q)a$ est un multiple de a ; par conséquent $r = r'$ et $q = q'$.
- Montrons l'existence pour un entier $n \in \mathbb{N}$ par récurrence forte.
  ▹ Le résultat est immédiat pour $n = 0$ avec $q = r = 0$.
  ▹ Soit $n \in \mathbb{N}$. Si l'existence est établie pour tout entier $k \in [\![1, n]\!]$, montrons-la pour $n+1$. Si $n+1 < a$, alors les entiers $q = 0$ et $r = n+1$ conviennent. Sinon, $n+1-a \in [\![0, n]\!]$ et on conclut avec l'hypothèse de récurrence.
- Pour montrer l'existence pour $n < 0$, on applique le point précédent à $-n$: il existe $(q, r) \in \mathbb{Z} \times \mathbb{N}$ tel que $-n = qa + r$ et $0 \leq r \leq a - 1$. Ainsi, $n = -qa - r$ est la division euclidienne si $r = 0$ et $n = (-q-1)a + (a - r)$ si $r \neq 0$.

■

Remarque

En fait, la preuve de ce résultat fournit une construction algorithmique du quotient et du reste.

Reprenons l'exemple des sous-groupes de $\mathbb{Z}$.

Proposition 11.8.

Soit G un sous-groupe de $\mathbb{Z}$. Alors, il existe $n \in \mathbb{N}$ tel que

$$G = n\mathbb{Z} = \{kn,\ k \in \mathbb{Z}\}.$$

Démonstration

Si $G = \{0\}$, alors $G = 0\mathbb{Z}$. Sinon, G contient un élément non nul x, $0 = x - x$ et $-x = 0 - x$ donc contient au moins un élément strictement positif. Soit $n = \min G \cap \mathbb{N} \setminus \{0\}$. Comme G est stable par différence, $n\mathbb{Z} \subset G$ ($-n = 0 - n \in G$ et si $kn \in G$, alors $(k+1)n = kn - (-n) \in G$).

Réciproquement, si $m \in G$, il existe $(q, r) \in \mathbb{Z}^2$ tel que $m = qn + r$ et $0 \leq r < n$; alors, $r = m - qn \in G$ et par minimalité de n, $r = 0$. Ainsi, $m \in n\mathbb{Z}$. Par double inclusion, $G = n\mathbb{Z}$. ■

1.3. Numération

Appliquons la division euclidienne pour obtenir le résultat de numération dans une base quelconque.

Proposition 11.9.

Soit b un entier supérieur ou égal à 2. Pour tout $n \in \mathbb{N} \setminus \{0\}$, il existe un unique entier $N \in \mathbb{N}$ et un unique $(N+1)$-uplet $(d_0, \dots, d_N) \in [\![0, b-1]\!]^{N+1}$ tels que :

$$n = \sum_{k=0}^{N} d_k b^k, \qquad d_N \neq 0.$$

On note alors $n = \overline{d_N \dots d_0}^b$.

Démonstration

▹ On montre une fois encore l'existence par récurrence forte.

• Le résultat est immédiat pour $n = 1$.

• Soit $n \in \mathbb{N} \setminus \{0\}$. Supposons que la propriété est établie pour tout $k \in [\![1, n]\!]$. Considérons la division euclidienne de $n+1$ par b : il existe $(q, d_0) \in \mathbb{Z}^2$ tel que $n+1 = qb + d_0$ avec $0 \leq d_0 < b$. Si $q = 0$, le résultat est établi. Sinon, $1 \leq q \leq n$ et l'hypothèse de récurrence au rang q entraîne qu'il existe $N' \in \mathbb{N}$ et un unique N'-uplet $(d'_0, \dots, d'_{N'}) \in [\![0, b-1]\!]^{N'+1}$ tels que

$$q = \sum_{k=0}^{N'} d'_k b^k, \qquad d'_{N'} \neq 0.$$

Alors,

$$n = \sum_{k=0}^{N'} d'_k b^{k+1} + d_0.$$

En posant $N = N' + 1$ et $d_{k+1} = d'_k$ pour $k \in [\![0, N']\!]$, on obtient le résultat annoncé.

▹ L'unicité se montre facilement par l'absurde : si

$$\sum_{k=0}^{N} d_k b^k = \sum_{k=0}^{N} d'_k b^k$$

avec $d_N \neq d'_N$ et tous les « chiffres » entre 0 et $b-1$, alors b^N divise

$$\sum_{k=0}^{N-1} (d'_k - d_k) b^k.$$

Or, ce dernier entier est inférieur en valeur absolue à

$$(b-1) \sum_{k=0}^{N-1} b^k = b^N - 1,$$

d'où la contradiction. ■

Remarque

L'étape d'hérédité peut être traitée différemment (ce qui donne une autre construction algorithmique).

Soit $n \in \mathbb{N} \setminus \{0\}$. Supposons que la propriété est établie pour tout $k \in [\![1, n]\!]$. Définissons N comme le plus grand élément de l'ensemble non vide $\{k \in \mathbb{N},\ b^k \leq n+1\}$ puis d_N comme le plus grand élément de

l'ensemble non vide $\{l \in \mathbb{N},\ l.b^N \leq n+1\}$; ce dernier ensemble contient 1 mais pas b donc $d_N \in [\![1, b-1]\!]$. Appliquons l'hypothèse de récurrence à $n+1-d_N b^N \leq n$: il existe $(d_0, \ldots, d_{N-1}) \in [\![0, b-1]\!]^N$ tels que :

$$n+1-d_N b^N = \sum_{k=0}^{N-1} d_k b^k,$$

donc

$$n+1 = \sum_{k=0}^{N} d_k b^k.$$

Exemple

La numération en base 2 ou écriture binaire est essentielle (entre autres pour l'informatique). L'écriture binaire de 108 est $\overline{1101100}^2$.

▹ En utilisant l'idée de la démonstration, on calcule les divisions successives par 2 et l'écriture binaire est composée de la suite des restes (en la lisant de bas en haut sur notre illustration).

$$\begin{array}{ccccc} 108 & = & 2\times 54 & + & 0 \\ 54 & = & 2\times 27 & + & 0 \\ 27 & = & 2\times 13 & + & 1 \\ 13 & = & 2\times 6 & + & 1 \\ 6 & = & 2\times 3 & + & 0 \\ 3 & = & 2\times 1 & + & 1 \\ 1 & = & 2\times 0 & + & 1 \end{array}$$

▹ En utilisant l'idée de la remarque, on retranche à chaque étape la plus grande puissance possible.

$$\begin{aligned} 108 &= 2^6+44 = 2^6+2^5+12 = 2^6+2^5+2^3+2^2 \\ &= 1\times 2^6+1\times 2^5+0\times 2^4+1\times 2^3+1\times 2^2+0\times 2^1+0\times 2^0. \end{aligned}$$

2. PGCD, PPCM

2.1. Définitions

Définition 11.10.

Soit $a, b \in \mathbb{Z}$.
▹ Le PGCD de a et b est l'unique entier naturel $a \wedge b$ tel que $a\mathbb{Z}+b\mathbb{Z} = (a \wedge b)\mathbb{Z}$.
▹ Le PPCM de a et b est l'unique entier naturel $a \vee b$ tel que $a\mathbb{Z} \cap b\mathbb{Z} = (a \vee b)\mathbb{Z}$.

La définition a bien un sens puisque les parties $a\mathbb{Z}+b\mathbb{Z}$ et $a\mathbb{Z}\cap b\mathbb{Z}$ sont non vides et stables par différence donc de la forme $n\mathbb{Z}$ (Proposition 11.8). Le lien avec la notion de divisibilité apparaît clairement avec la proposition suivante, équivalente à la définition.

Proposition 11.11.

Soit $a, b \in \mathbb{Z}$ non tous les deux nuls. Alors,
▹ le PGCD de a et b est le plus grand diviseur commun à a et b ;
▹ le PPCM de a et b est le plus petit multiple commun à a et b.

Démonstration

▹ Montrons le résultat pour le PGCD.

- $a\mathbb{Z} \subset (a \wedge b)\mathbb{Z}$ donc a est un multiple de $a \wedge b$. De même pour b.
- Soit c un diviseur de a et de b. Alors, $a\mathbb{Z} \subset c\mathbb{Z}$ et $b\mathbb{Z} \subset c\mathbb{Z}$ donc
$$(a \wedge b)\mathbb{Z} = a\mathbb{Z} + b\mathbb{Z} \subset c\mathbb{Z},$$
donc c est un diviseur de $a \wedge b$.

▹ Montrons le résultat pour le PPCM.

- $(a \vee b)\mathbb{Z} \subset a\mathbb{Z}$ donc $a \vee b$ est un multiple de a (et de même pour b).
- Soit c un multiple de a et de b. Alors, $c\mathbb{Z} \subset a\mathbb{Z}$ et $c\mathbb{Z} \subset b\mathbb{Z}$ donc
$$c\mathbb{Z} \subset a\mathbb{Z} \cap b\mathbb{Z} = (a \vee b)\mathbb{Z},$$
donc c est un multiple de $a \vee b$.

■

Remarque

Cette définition permet d'aller plus loin et de définir le PGCD d'une famille finie d'entiers : le PGCD des entiers $a_1, \ldots, a_n$ est l'unique entier naturel $a_1 \wedge \ldots \wedge a_n$ tel que
$$a_1\mathbb{Z} + \ldots + a_n\mathbb{Z} = (a_1 \wedge \ldots \wedge a_n)\mathbb{Z}.$$
On remarque en particulier « l'associativité du PGCD » ; pour tous $a, b, c \in \mathbb{Z}$,
$$a \wedge (b \wedge c) = (a \wedge b) \wedge c = a \wedge b \wedge c.$$
Les mêmes remarques restent valables pour le PPCM.

Remarque

La définition de PGCD donne directement la propriété de Bézout : pour tous $a, b \in \mathbb{Z}$, il existe u, $v \in \mathbb{Z}$ tels que $ua + vb = a \wedge b$.
On remarque qu'il n'y a pas unicité de ces coefficients entiers u et v. Par exemple,
$$1 \times 3 + (-1) \times 2 = (-1) \times 3 + 2 \times 2 = 2 \wedge 3.$$

2.2. Algorithme d'Euclide

Précisons tout d'abord quelques propriétés élémentaires du PGCD.

Proposition 11.12.

Soit a, $b \in \mathbb{Z}$. Alors,

- $|a| \wedge |b| = a \wedge b$
- $a \wedge 0 = |a|$
- $\forall k \in \mathbb{Z}, \quad a \wedge (b - ka) = a \wedge b$
- $\forall c \in \mathbb{Z}, \quad (ca) \wedge (cb) = |c|(a \wedge b)$

Démonstration

Il suffit de remarquer que $a\mathbb{Z} = |a|\mathbb{Z}$ et que $a\mathbb{Z} + (b - ka)\mathbb{Z} = a\mathbb{Z} + b\mathbb{Z}$. ■

Ces remarques permettent d'obtenir l'algorithme d'Euclide de calcul du PGCD dont voici une description.

1. On se ramène à des entiers naturels.
2. On retranche au plus grand des deux entiers le plus petit des deux et on recommence jusqu'à ce que l'un des deux entiers soit nul.
3. L'entier restant (celui qui n'est pas nul) est le PGCD.

La deuxième étape peut être réécrite légèrement différemment

2. On remplace le plus grand des deux entiers par le reste de la division euclidienne de celui-ci par le plus petit des deux et on recommence jusqu'à ce que l'un des deux entiers soit nul.

Pour calculer le PGCD de plus de deux entiers, on utilise l'associativité déjà remarquée.

Exemple

Calculons le PGCD de 1789 et 1515 avec l'algorithme d'Euclide.
La première étape est la division de 1789 par 1515, $1789 = 1.1515 + 274$; on effectue ensuite la division de 1515 par le premier reste 274, $1515 = 5.274 + 145$ et on recommence : $274 = 1.145 + 129$, $145 = 1.129 + 16$, $129 = 8.16 + 1$, $16 = 16.1 + 0$: le PGCD est donc le dernier reste non nul : 1.

Exemple

Soit $(F_n)_n$ une suite de Fibonacci définie par ses deux premiers termes et par la relation de récurrence

$$\forall n \in \mathbb{N}, \qquad F_{n+2} = F_{n+1} + F_n.$$

Les restes successifs de l'algorithme d'Euclide appliqué avec les arguments F_{n+1} et F_n sont les n termes de la suite de Fibonacci d'indices strictement inférieurs à n (car la division de F_{n+1} par F_n donne le quotient 1 et le reste F_{n-1}). En particulier, l'entier $F_{n+1} \wedge F_n$ ne dépend pas de n et se trouve donc être égal à $F_1 \wedge F_0$.

Remarque

Ce dernier calcul peut fournir davantage d'informations. Une récurrence sur n permet de montrer que si l'algorithme d'Euclide pour calculer le PGCD des entiers naturels a et $b > a$ requiert n étapes, alors les minorations $b \geq F_{n+2}$ et $a \geq F_{n+1}$ sont vérifiées.

Un corollaire de ces inégalités est que l'on peut majorer le nombre d'étapes pour calculer le PGCD d'entiers a et b en se ramenant à la suite de Fibonacci connue explicitement. Plus précisément, le théorème de Lamé indique le nombre d'étapes dans l'algorithme d'Euclide pour deux entiers naturels est majoré par cinq fois le nombre de chiffres dans l'écriture décimale du plus petit des deux entiers.

2.3. Nombres premiers entre eux

Définition 11.13.

Deux entiers sont premiers entre eux si leur PGCD est égal à 1.

Cette notion apparaît souvent, ne serait-ce que dans la définition du représentant irréductible d'un rationnel.

Remarque

On dit aussi bien que deux entiers a et b sont premiers entre eux ou que a est premier avec b.

Remarquons que la notion « d'entiers premiers entre eux » est définie sans avoir recours aux nombres premiers.

Théorème 11.14. Bézout

Deux entiers a et b sont premiers entre eux si, et seulement si, il existe u, $v \in \mathbb{Z}$ tels que

$$ua + vb = 1.$$

Démonstration

Le sens direct est une simple conséquence de la définition du PGCD. Pour le sens retour, remarquons que le PGCD divise à la fois a et b donc $ua + vb = 1$; par conséquent, il est égal à 1. ■

Conseils méthodologiques

On trouve des entiers u et v satisfaisant cette relation en « remontant » les calculs dans l'algorithme d'Euclide.

Remarque

Soit a et b deux entiers premiers entre eux et u, $v \in \mathbb{Z}$ tels que $ua + vb = 1$. Alors, pour tout $k \in \mathbb{Z}$, $(u - kb)a + (v + ka)b = 1$. On peut donc retrouver les autres relations de Bézout et choisir k de sorte à obtenir certaines propriétés (par exemple la positivité) pour l'un des coefficients.

Déduisons du théorème de Bézout quelques propriétés commodes.

Corollaire 11.15.

▹ Si $a \wedge b = 1$ et $a \wedge c = 1$, alors $a \wedge (bc) = 1$.
▹ Si $a \wedge b = 1$, alors $a^p \wedge b^q = 1$ pour tous p, $q \in \mathbb{N}$.

Démonstration

Le second point se déduit par récurrence du premier point. D'après le théorème de Bézout, il existe u, v, u', $v' \in \mathbb{Z}$ tels que

$$ua + vb = 1, \qquad u'a + v'c = 1.$$

D'où $ua + vb(u'a + v'c) = 1$, soit $(u + vbu')a + (vv')bc = 1$.
D'après le théorème de Bézout, $a \wedge (bc) = 1$. ■

Corollaire 11.16.

Soit a, b, c trois entiers tels que a et b divisent c et $a \wedge b = 1$. Alors, ab divise c.

Démonstration

D'après le théorème de Bézout, il existe $u, v \in \mathbb{Z}$ tel que $ua + vb = 1$ donc $uac + vbc = c$. L'entier b divise c donc ab divise ac ; de même, ab divise bc. Ainsi, ab divise $uac + vbc = c$. ■

Corollaire 11.17. Lemme de Gauss

Soit a, b, c trois entiers tels que a divise bc et $a \wedge b = 1$. Alors a divise c.

Démonstration

D'après le théorème de Bézout, il existe $u, v \in \mathbb{Z}$ tels que $ua + vb = 1$ donc $uac + vbc = c$. a divise les deux termes du premier membre donc divise c. ■

Voici une application utile du lemme de Gauss.

Exemple

Considérons un polynôme à coefficients entiers de degré n

$$P(X)=\sum_{k=0}^{n}\alpha_k X^k.$$

Montrons que si la fraction irréductible $\frac{p}{q}$ est une racine de P, alors p divise α_0 et q divise α_n.
En effet,

$$\sum_{k=0}^{n}\alpha_k p^k q^{n-k}=0,$$

soit

$$\alpha_0 q^n=-\sum_{k=1}^{n}\alpha_k p^k q^{n-k}.$$

On applique alors le lemme de Gauss : p divise tous les termes $\alpha_k p^k q^{n-k}$ pour $k\geq 1$ donc divise $\alpha_0 q^n$; comme p est premier avec q donc q^n, p divise α_0.
On procède de même avec q et α_n.
Ce critère permet une recherche facile des racines rationnelles d'un polynôme à coefficients entiers en donnant une condition nécessaire qui limite à un nombre fini de cas : il suffit ensuite de tester ces quelques valeurs pour vérifier si elles sont racines.
Par exemple, d'après le critère précédent, les racines rationnelles de $X^4-9X^3+13X^2+X+2$ sont parmi $\{-2,-1,1,2\}$; on vérifie rapidement que 2 est la seule racine rationnelle.

3. Nombres premiers

3.1. Définitions

Définition 11.18.

Un entier $p\in\mathbb{N}\setminus\{0,1\}$ est premier si les seuls diviseurs positifs de p sont 1 et p (soit quatre diviseurs si on n'impose pas la positivité).

Remarque

▹ Il n'y a qu'un seul nombre premier pair : 2.
▹ L'entier 1 n'est pas un nombre premier.

Proposition 11.19.

Soit p un nombre premier. Alors, pour tout entier n, $n\wedge p\in\{1,p\}$.

Démonstration

Les seuls diviseurs positifs de p sont 1 et p donc $n\wedge p$ est égal à l'un de ces entiers. ■

Remarque

En fait, cette proposition indique que, pour tout nombre premier p, les entiers sont soit premiers avec p, soit des multiples de p.

Corollaire 11.20.

Pour tout nombre premier p et tout entier $k \in [\![1, p-1]\!]$, p divise $\binom{p}{k}$.

Démonstration
En effet, $p\binom{p-1}{k-1} = k\binom{p}{k}$ donc p divise $k\binom{p}{k}$. Par ailleurs, p est premier avec k (car k n'est pas un multiple de p) donc, d'après le théorème de Gauss, p divise $\binom{p}{k}$. ■

Proposition 11.21. Fermat

Soit p un nombre premier et $n \in \mathbb{Z}$. Alors, $n^p = n[p]$.

Démonstration
Soit p premier fixé.
▹ Pour $n = 1$, le résultat est évident.
▹ Soit $n \in \mathbb{N}$ tel que $n^p = n[p]$. Alors,

$$\begin{aligned}(n+1)^p &= \sum_{j=0}^{p} \binom{p}{j} n^j = n^p + \sum_{j=1}^{p-1} \binom{p}{j} n^j + 1[p] \\ &= n + 0 + 1[p] \\ &= n + 1[p],\end{aligned}$$

d'après l'hypothèse de récurrence.
On a ainsi établi le résultat pour $n \in \mathbb{N}$; en séparant les cas

- $p = 2$ pour lequel $(-n)^2 = n^2[2] = n[2] = -n[2]$ pour tout $n \in \mathbb{N}$,
- p impair pour lequel $(-n)^p = -n^p = -n[p]$ pour tout $n \in \mathbb{N}$,

on obtient le résultat pour tout $n \in \mathbb{Z}$. ■

Corollaire 11.22. Fermat

Soit p un nombre premier et $n \notin p\mathbb{Z}$. Alors, $n^{p-1} = 1[p]$.

Démonstration
On a déjà que p divise $n^p - n = n(n^{p-1} - 1)$. Or, p et n sont premiers entre eux, donc d'après le théorème de Gauss, p divise $n^{p-1} - 1$, soit $n^{p-1} = 1[p]$. ■

Passons à une autre application des nombres premiers.

Proposition 11.23. Euclide

Soit p un nombre premier et des entiers $a_1, \ldots, a_n$. Si p divise $\prod_{k=1}^{n} a_k$, alors il existe $k \in [\![1, n]\!]$ tel que p divise a_k.

Démonstration
Raisonnons par l'absurde. Si p ne divise aucun de ces entiers, alors p est premier avec chacun d'entre eux donc avec le produit : contradiction. ■

Voici une proposition qui nous sera utile pour le théorème fondamental de l'arithmétique.

Proposition 11.24.

Tout entier de valeur absolue supérieure ou égale à 2 admet un diviseur premier.

Démonstration
On procède par récurrence forte.

- 2 admet le nombre premier 2 comme diviseur.
- Soit $n \geq 2$. Supposons que le résultat est vrai pour tout entier $k \in [\![2, n]\!]$. Si $n+1$ est premier, alors $n+1$ est un diviseur premier de $n+1$. Sinon, $n+1 = a.b$ avec $a, b \in [\![2, n]\!]$ et on applique l'hypothèse de récurrence à a ou b.

■

3.2. Théorème fondamental de l'arithmétique

Théorème 11.25.

Pour tout entier $n \geq 2$, il existe un entier $m \geq 1$, des nombres premiers $p_1 > \ldots > p_m$ et des entiers non nuls $r_1, \ldots, r_m$ tels que :

$$n = \prod_{k=1}^{m} p_k^{r_k}.$$

Démonstration
On procède par récurrence forte en s'appuyant sur la proposition 11.24. ■

Remarque
Cette décomposition est en fait unique (à l'ordre près des facteurs). Si $n = \prod_{k=1}^{m} p_k^{r_k} = \prod_{k=1}^{m} p_k^{s_k}$ avec $r_1 < s_1$, alors $p_1^{s_1 - r_1}$ divise $\prod_{k=2}^{m} p_k^{r_k}$ et est premier avec ce nombre : contradiction.

Il convient de ne pas oublier que cette décomposition est algorithmiquement difficile à obtenir.

Définition 11.26.

Soit p un nombre premier. La valuation p-adique d'un entier n non nul est l'entier $v_p(n)$ défini comme l'exposant de p dans la décomposition de n en facteurs premiers avec, par convention, la valeur 0 si le nombre premier p n'apparaît pas.

Exemple

Comme $120 = 2^3 \times 3 \times 5$, $v_2(120) = 3$, $v_3(120) = 1$, $v_5(120) = 1$ et $v_7(120) = 0$.

Voici quelques règles de calcul immédiates en écrivant la décomposition en facteurs premiers.

Proposition 11.27.

Soit p un nombre premier.
▹ Pour tous entiers m et n non nuls, $v_p(mn) = v_p(m) + v_p(n)$.
▹ Pour tout entier n non nul et tout $k \in \mathbb{N}$, $v_p(n^k) = k\, v_p(n)$.

Exemple

Soit $n \in \mathbb{N}$. Montrons que $\sqrt{n} \in \mathbb{Q}$ si, et seulement si, n est le carré d'un entier. Le sens retour est évident. Pour le sens direct, supposons que $\sqrt{n}$ soit égale à la fraction irréductible $\frac{a}{b}$. Alors, $a^2 = nb^2$; en écrivant les décompositions en facteurs premiers de a, b et n, on obtient que, pour tout facteur premier p, $v_p(n) = 2v_p(a) - 2v_p(b)$ est pair : par conséquent, n est un carré.

Proposition 11.28.

Soit a, n deux entiers. Si a est un diviseur de l'entier n de décomposition en facteurs premiers $\prod_{k=1}^{m} p_k^{r_k}$, alors $a = \prod_{k=1}^{m} p_k^{\alpha_k}$ avec pour tout $k \in [\![1, m]\!]$, $\alpha_k \leq r_k$.

Corollaire 11.29.

Soit a, n deux entiers. Alors, a divise n si, et seulement si, pour tout nombre premier p, $v_p(a) \leq v_p(n)$.

On obtient ainsi le dénombrement des diviseurs d'un entier (et la somme d'après les sommes géométriques).

Corollaire 11.30.

L'entier naturel de décomposition en facteurs premiers $\prod_{k=1}^{m} p_k^{r_k}$ admet exactement $\prod_{k=1}^{m}(r_k+1)$ diviseurs positifs.
La somme de tous ses diviseurs est

$$\prod_{k=1}^{m} \frac{p_k^{r_k+1} - 1}{p_k - 1}.$$

Exemple

Les entiers qui admettent un nombre impair de diviseurs positifs sont les carrés d'un entier ; en effet, il faut que tous les facteurs $r_k + 1$ soient impairs et donc que les r_k soient pairs.

Corollaire 11.31.

Considérons les entiers naturels a, b de décompositions en facteurs premiers

$$a = \prod_{k=1}^{m} p_k^{\alpha_k}, \qquad b = \prod_{k=1}^{m} p_k^{\beta_k}.$$

Alors,

$$a \wedge b = \prod_{k=1}^{m} p_k^{\min(\alpha_k, \beta_k)}, \qquad a \vee b = \prod_{k=1}^{m} p_k^{\max(\alpha_k, \beta_k)}.$$

Autrement dit, pour tout premier p,

$$v_p(a \wedge b) = \min(v_p(a), v_p(b)), \qquad v_p(a \vee b) = \max(v_p(a), v_p(b)).$$

Remarque

Cette proposition ne fournit pas un calcul aussi efficace que l'algorithme d'Euclide pour déterminer le PGCD de deux entiers.

Remarque
On vérifie grâce à cette propriété que $(a \wedge b).(a \vee b) = |a.b|$.

Les résultats qui suivent relèvent de la culture mathématique.

Proposition 11.32.

L'ensemble des nombres premiers est infini.

Démonstration
Raisonnons par l'absurde. Supposons que l'ensemble des nombres premiers est fini $\{p_1, \ldots, p_N\}$. L'entier $\prod_{k=1}^{N} p_k + 1$ n'admet aucun diviseur premier : contradiction. ■

Remarque
On connaît des résultats plus fins sur cet ensemble; par exemple, $\pi(n)$, le cardinal de l'ensemble des nombres premiers inférieurs à n vérifie :

$$\pi(n) \underset{\infty}{\sim} \frac{n}{\ln n}.$$

Ce théorème appelé théorème des nombres premiers est dû à de La Vallée-Poussin et Hadamard.

Proposition 11.33.

L'ensemble des nombres premiers de la forme $4n-1$ est infini.

Démonstration
Raisonnons par l'absurde. Supposons que cet ensemble est fini $\{p_1, \ldots, p_N\}$. L'entier $4\prod_{k=1}^{N} p_k - 1$ admet un diviseur premier de la forme $4n-1$ (ses diviseurs premiers sont parmi 2 et les nombres premiers de la forme $4n+1$ car il est congru à $-1[4]$) qui n'est pas dans la liste précédente : contradiction. ■

4. Systèmes congruentiels

Dans cette section, on cherche à résoudre quelques équations diophantiennes (c'est-à-dire équation à inconnues parmi les entiers) congruentielles d'inconnues x et y. Les autres lettres désignent des entiers paramètres.

Exemple

Par définition, l'ensemble des entiers x tels que $x = a[n]$ est $a + n\mathbb{Z}$.

Exemple

Notons S l'ensemble des entiers x tels que $ax = b[n]$.
- Si $a \wedge n = 1$, alors il existe $u, v \in \mathbb{Z}$ tels que $ua + vn = 1$ et $S = bu + n\mathbb{Z}$.
- Si $a \wedge n \neq 1$, alors
  - si $a \wedge n | b$, alors, on divise par $a \wedge n$ et on se ramène au cas précédent.
  - sinon, $S = \varnothing$.

4.1. Cas des modulos premiers entre eux

Notons S l'ensemble des entiers x solutions de $\left\{\begin{array}{lcl} x & = & a[p] \\ x & = & b[q] \end{array}\right.$ avec $p \wedge q = 1$.

D'après le théorème de Bézout, il existe $u, v \in \mathbb{Z}$ tels que $up+vq = 1$. L'entier $avq+bup$ est alors solution du système et on en déduit que l'ensemble des solutions est : $S = avq + bup + pq\mathbb{Z}$.

On remarque, en particulier, qu'il n'y a qu'une solution dans $[\![0, pq-1]\!]$. Pour résoudre un système congruentiel de plus de deux équations, on regroupe les équations deux par deux.

4.2. Cas général

Pour résoudre le système diophantien $\left\{\begin{array}{lcl} x & = & a[m] \\ x & = & b[n] \end{array}\right.$ avec des entiers $m = \prod_i p_i^{m_i}$ et $n = \prod_i p_i^{n_i}$ non nécessairement premiers entre eux donnés par leurs décompositions en facteurs premiers, on utilise le protocole suivant :

1. on se ramène à des « modulos » deux à deux premiers entre eux en écrivant chacune des équations sous la forme du système correspondant :
$$\left\{\begin{array}{lcl} x & = & a[p_1^{m_1}] \\ x & = & a[p_2^{m_2}] \\ & \vdots & \end{array}\right., \qquad \left\{\begin{array}{lcl} x & = & b[p_1^{n_1}] \\ x & = & b[p_2^{n_2}] \\ & \vdots & \end{array}\right.;$$
2. on étudie, pour tout i, le système composé des deux équations correspondant au même facteur premier
$$\left\{\begin{array}{lcl} x & = & a[p_i^{m_i}] \\ x & = & b[p_i^{n_i}]. \end{array}\right.$$
On obtient l'aternative suivante : soit le système n'admet pas de solution, soit les solutions forment une classe de congruence modulo $p_i^{\max(m_i,n_i)}$;
3. on conclut comme précédemment : le résultat obtenu est alors une classe modulo $m \vee n$.

Remarque

Dans la deuxième étape, on a utilisé que si $x = b[p^n]$, alors $x = b[p^m]$ pour tout $m \leq n$ ce qui est élémentaire en réécrivant la définition de congruence. Pour le dire en d'autres termes, une classe modulo p^m pour $m \leq n$ est une réunion de classes modulo p^n.

Exemple

Le système $\left\{\begin{array}{lcl} x & = & 17[28] \\ x & = & 3[98] \end{array}\right.$ est équivalent à

$$\left\{\begin{array}{lclcl} x & = & 17[2^2] & = & 1[2^2] \\ x & = & 17[7] & = & 3[7] \\ x & = & 3[2] & = & 1[2] \\ x & = & 3[7^2] & = & 3[7^2] \end{array}\right.$$

qui équivaut, à son tour, à

$$\left\{\begin{array}{lcl} x & = & 1[2^2] \\ x & = & 3[7^2] \end{array}\right.$$

Comme $-12.2^2 + 1.7^2 = 1$, les solutions sont d'après le point précédent, $101 + 196\mathbb{Z}$.

Exemple

Le système $\left\{\begin{array}{lcl} x & = & 7[12] \\ x & = & 5[15] \end{array}\right.$ est équivalent à

$$\left\{\begin{array}{lclcl} x & = & 7[4] & = & 3[4] \\ x & = & 7[3] & = & 1[3] \\ x & = & 5[3] & = & 2[3] \\ x & = & 5[5] & = & 0[5] \end{array}\right.$$

Il n'y a pas de solutions car les deuxième et troisième équations ne sont pas compatibles.

FICHE

SYNTHÈSE

La principale propriété de l'ensemble $\mathbb{Z}$ exploitée dans ce chapitre est l'existence de la division euclidienne.

Division euclidienne

Soit $(n,a) \in \mathbb{Z} \times \mathbb{N} \setminus \{0\}$. Il existe un unique couple $(q,r) \in \mathbb{Z} \times \mathbb{N}$ tel que

$$n = qa + r \qquad 0 \leq r \leq a-1.$$

Celle-ci permet de montrer l'existence et l'unicité de l'écriture d'un entier dans une base $b > 1$.
Pour en tirer le meilleur profit dans les calculs, on introduit la relation de congruence.

Congruence

▹ Soit $n \in \mathbb{N} \setminus \{0\}$. Pour tous a, $b \in \mathbb{Z}$, $a = b[n]$ si $a - b \in n\mathbb{Z}$.
▹ Pour tous a, b, c, $d \in \mathbb{Z}$ tels que $a = b[n]$ et $c = d[n]$, on a $a + c = b + d[n]$ et $ac = bd[n]$.

Théorème de Fermat

Soit p un nombre premier et $n \in \mathbb{Z}$. Alors, $n^p = n[p]$ et, si $n \notin p\mathbb{Z}$, $n^{p-1} = 1[p]$.

On exploite la divisibilité avec les notions de PGCD et PPCM.

PGCD, PPCM

▹ Le PGCD de deux entiers a et b est le plus grand diviseur commun à a et b. De manière équivalente, c'est l'unique entier naturel $a \wedge b$ tel que

$$a\mathbb{Z} + b\mathbb{Z} = (a \wedge b)\mathbb{Z}.$$

Ainsi, il existe des entiers u et $v \in \mathbb{Z}$ tels que $ua + vb = a \wedge b$.
▹ Le PPCM de deux entiers a et b est le plus petit multiple commun à a et b. De manière équivalente, c'est l'unique entier naturel $a \vee b$ tel que

$$a\mathbb{Z} \cap b\mathbb{Z} = (a \vee b)\mathbb{Z}.$$

Le PGCD se calcule rapidement avec l'algorithme d'Euclide par divisions successives.

Entiers premiers entre eux

▹ Deux entiers a et b sont premiers entre eux si leur PGCD est égal à 1. Ceci équivaut à l'existence d'entiers u, $v \in \mathbb{Z}$ tels que $ua + vb = 1$ (théorème de Bézout).
▹ Soit a, b, c trois entiers tels que a divise bc et $a \wedge b = 1$. Alors a divise c (théorème de Gauss).

Le théorème de Gauss admet de nombreux corollaires et applications qu'il faut savoir identifier.
On introduit alors la notion de nombres premiers.

Nombres premiers

▹ Un entier $p \in \mathbb{N} \setminus \{0, 1\}$ est premier si les seuls diviseurs positifs de p sont 1 et p.
▹ Tout entier s'écrit de manière unique comme produit de nombres premiers.

Les exposants qui apparaissent dans la décomposition en facteurs premiers sont les valuations p-adiques.
Cette décomposition permet de revisiter la notion de divisibilité, de PGCD, de PPCM mais elle est difficile à obtenir donc ne permet pas de remplacer efficacement les algorithmes basés sur la division euclidienne.

EXERCICES

Vrai ou faux ?

		Vrai	Faux
a)	0 est un diviseur de tout entier n.	☐	☐
b)	1326543987321 est divisible par 11.	☐	☐
c)	Si $n = 3[4]$, alors n n'est pas somme de deux carrés d'entiers.	☐	☐
d)	Le nombre de diviseurs positifs de 60 est 12.	☐	☐
e)	S'il existe $u, v \in \mathbb{Z}$ tels que $au + bv = 2$, alors $a \wedge b = 2$.	☐	☐
f)	Soit $k, n \in \mathbb{Z}$. Si k ne divise pas n, alors $k \wedge n = 1$.	☐	☐
g)	Pour tout $n \in \mathbb{N}$, n et $n+1$ sont premiers entre eux.	☐	☐
h)	Pour tout $n \in \mathbb{N}$ impair, n et $n+2$ sont premiers entre eux.	☐	☐
i)	L'équation $3x = 2[25]$ admet une infinité de solutions.	☐	☐
j)	L'équation $3x = 2[24]$ admet une infinité de solutions.	☐	☐
k)	$x = 1[72]$ si, et seulement si, $x = 1[9]$ et $x = 1[8]$.	☐	☐
l)	$x = 13[35]$ si, et seulement si, $x = 6[7]$ et $x = 8[5]$.	☐	☐
m)	Pour tout $a \notin 6\mathbb{Z}$, $a^5 = 1[6]$.	☐	☐

Exercices d'application du chapitre 11

●○○ **Exercice 1**

Déterminer les chiffres a tel que le nombre d'écriture décimale $123a4$ soit divisible par 12.

●○○ **Exercice 2**

Soit x un nombre à (au plus) deux chiffres. Montrer que le nombre à (au plus) six chiffres obtenu en juxtaposant trois fois x est divisible par 37.

●○○ **Exercice 3**

Déterminer un critère simple de parité pour un entier donné par son écriture en base 3.

●○○ **Exercice 4**

Montrer que $n = 10a + b$ avec $b \in [\![0, 9]\!]$ est divisible par 7 si, et seulement si, $a - 2b$ est divisible par 7.

●●○ **Exercice 5**

Montrer que pour tout $n \in \mathbb{N}$, 7 divise $2^{4^n} - 2$.

●○○ **Exercice 6**

Déterminer les entiers naturels n tels que $n-3$ divise n^3-3.

●○○ **Exercice 7**

En réduisant modulo un entier astucieusement choisi, montrer que les équations suivantes n'admettent pas de solution $(x, y) \in \mathbb{Z}^2$:

1. $x^2 - 4y^2 = 1791$,
2. $x^2 + y^2 = 1791$,
3. $x^2 + 6y^2 = 1112$.

●○○ **Exercice 8**

Soit a, $b \in \mathbb{Z}$ et $n \in \mathbb{N} \setminus \{0\}$ tels que $a = b[n]$. Montrer que $a^n = b^n[n^2]$.

●○○ **Exercice 9**

Pour tout entier n, posons $u_n = 2^n + 1$.

1. Montrer que si n est impair et différent de 1, alors u_n n'est pas un nombre premier.
2. Supposons que n est de la forme $2^q(2k+1)$ avec $k \geq 1$. Montrer que u_n n'est pas un nombre premier.
3. En déduire que si u_n est nombre premier, alors n est une puissance de 2.

●●○ **Exercice 10**

Trouver les entiers a, $b \in \mathbb{Z}$ tels que $3^a 7^b = 1[10]$.

●●● **Exercice 11**

Soit p un nombre premier.

1. Montrer que, pour tout $n \in \mathbb{N} \setminus \{0\}$, $\nu_p(n!) = \sum_{k=1}^{\infty} \left\lfloor \frac{n}{p^k} \right\rfloor$.
2. Déterminer le nombre de chiffres 0 à la fin de l'écriture décimale de l'entier 100!.

○○○ **Exercice 12**

Montrer que, pour tout $n \in \mathbb{N}$, les fractions suivantes $\frac{15n+4}{10n+3}$, $\frac{21n+4}{14n+3}$ sont irréductibles.

●○○ **Exercice 13**

Soit a, b deux entiers premiers entre eux. Montrer que ab et $a+b$ sont premiers entre eux.

●○○ **Exercice 14**

Chercher les couples d'entiers $(a, b) \in \mathbb{N}^2$ tels que $a \wedge b = 42$ et $a \vee b = 1680$.

●○○ **Exercice 15**

Calculer, pour tout $(a, b) \in \mathbb{Z}^2$, $(a+b) \wedge (a \vee b)$.

●●● **Exercice 16**

Déterminer les solutions entières $(x, y) \in \mathbb{N}^2$ de l'équation $x \wedge y + x \vee y = y + 4$.

●●○ **Exercice 17**

1. Déterminer les solutions entières (x, y) de l'équation $323x - 391y = 612$.
2. Trouver tous les entiers n tels que $n = 270[323]$ et $n = 100[391]$.

●●● **Exercice 18**

Considérons 1789 entiers tels que leur somme soit nulle. Montrer que la somme de leurs puissances 37ièmes est divisible par 399.

●●● **Exercice 19**

Notons, pour tout entier $n \in \mathbb{N}^*$, $d(n)$ le nombre de diviseurs positifs de n. Fixons $\varepsilon > 0$.

1. En notant $n = p_1^{\alpha_1} \dots p_k^{\alpha_k}$ la décomposition en facteurs premiers de n, montrer l'égalité suivante
$$d(n) = \prod_{i=1}^{k} (\alpha_i + 1).$$
2. En conservant ces notations, montrer que si $p_i \geq 2^{\frac{1}{\varepsilon}}$, alors $\frac{\alpha_i+1}{p_i^{\varepsilon\alpha_i}} \leq 1$.
3. Déterminer le maximum de la fonction $f : x \mapsto \frac{x+1}{2^{\varepsilon x}}$ sur $\mathbb{R}_+$.
4. En déduire qu'il existe C_ε ne dépendant que de ε tel que $d(n) \leq C_\varepsilon n^\varepsilon$.

Exercices d'approfondissement du chapitre 11

EXERCICES

●○○ **Exercice A**

Montrer qu'il n'existe pas d'entiers naturels m et n tels que $3^m + 3^n + 1$ soit le carré d'un entier.
Indication : on pourra par exemple réduire cette quantité modulo un entier judicieusement choisi.

●●○ **Exercice B**

Soit n un entier pair. Montrer que $n+1$ divise $\sum_{k=1}^{n} k^{n+1}$.

●●○ **Exercice C**

Considérons la suite $(u_n)_{n\geq 1}$ définie par $u_1 = 2$ et pour tout entier $n \geq 1$
$$u_{n+1} = \begin{cases} 2.10^n + u_n, & \text{si } 2^{n+1} \text{ divise } u_n\,; \\ 10^n + u_n, & \text{sinon.} \end{cases}.$$

1. Montrer que, pout tout $n \geq 1$, 2^n divise u_n.
2. En déduire que pour tout $n \geq 1$, 2^n admet un multiple entier dont l'écriture décimale ne comporte que des 1 et des 2.

●●○ **Exercice D**

Notons pour tout $k \in \mathbb{N} \setminus \{0\}$, u_k le plus grand diviseur impair de k. Montrer que, pour tout $n \in \mathbb{N} \setminus \{0\}$,
$$\sum_{k=n+1}^{2n} u_k = n^2.$$

●●○ **Exercice E Problème de Frobenius**

Soit a, $b \in \mathbb{N} \setminus \{0,1\}$ deux entiers premiers entre eux et $E = \{ua + vb,\ u, v \in \mathbb{N}\}$ l'ensemble des entiers pouvant se représenter à partir de a et b avec des coefficients positifs.

1. Montrer que $ab - a - b \notin E$.
2. Montrer que, pour tout entier n, il existe un unique couple d'entiers (u_0, v_0) tel que $u_0 a + v_0 b = n$ et $0 \leq v_0 < a$.
3. Montrer qu'un entier n appartient à E si, et seulement si, $u_0 \geq 0$.
4. Montrer que, pour tout entier $n > ab - a - b$, $n \in E$.

●●● **Exercice F**

1. Montrer qu'un diviseur premier impair d'un entier de la forme a^2+1 avec $a \in \mathbb{Z}$ est congru à 1 modulo 4.
2. En déduire que l'ensemble des nombres premiers congrus à 1 modulo 4 est infini.

Problèmes du chapitre 11

●○○ **Problème 1 Théorèmes de Wilson et de Dickson**

On cherche à établir le résultat suivant.

Proposition 11.34. Dickson

Un entier $n > 1$ est premier si, et seulement si, il existe $0 < m < n$ tel que

$$(m-1)!(n-m)! = (-1)^m [n].$$

1. On se propose d'établir le sens direct pour la valeur $m = 1$.
   a. Montrer que pour tout $a \in [\![1, n-1]\!]$, il existe un unique $b \in [\![1, n-1]\!]$ tel que $ab = 1[n]$
   b. En déduire que si n est premier, alors $(n-1)! = -1[n]$.
2. Supposons qu'il existe $0 < m < n$ tel que $(m-1)!(n-m)! = (-1)^m[n]$.
   a. Montrer que n est premier avec tous les entiers $k \in [\![1, n-m]\!]$.
   b. Soit $k \in [\![n-m+1, n-1]\!]$. Montrer que n est premier avec $n-k$, puis avec k.
   c. Conclure.

●●○ **Problème 2 Théorème de Erdős-Pálfy**

Notons, pour tout $n \in \mathbb{N} \setminus \{0\}$ et tout premier p, n_p le reste de la division de n par p.
L'objectif de ce court problème est de montrer le théorème suivant du à Erdös et Pálfy (1987).

Proposition 11.35. Erdős-Pálfy

Soit a et $b \in \mathbb{N} \setminus \{0\}$ tels que, pour tout premier p, $a_p \leq b_p$. Alors, $a = b$.

On raisonne par l'absurde et on considère dans tout ce sujet deux entiers naturels distincts a et b vérifiant l'hypothèse de l'énoncé.

Partie A – Préliminaires

1. Justifier que $a \leq b$.
2. Montrer que les entiers $b-a$ et b vérifient également les hypothèses de l'énoncé.
3. En déduire que l'on peut supposer sans perte de généralité $a \leq b/2$.

Supposons dorénavant $a \leq b/2$.

Partie B – Démonstration de Szegedy

Notons $E_1 = [\![1, a]\!]$ et $E_2 = [\![b-a+1, b]\!]$, puis, pour $i \in \{1,2\}$ et pour p premier,

$$r_i(p^k) = \left|\{n \in E_i,\ p^k \mid n\}\right|$$

et

$$\pi_i = \prod_{n \in E_i} n.$$

Par exemple, avec $a = 57$, $r_1(3^2) = 6$.

1. Montrer que $\frac{\pi_2}{\pi_1}$ est un entier.
2. Soit p un premier.
   a. Rappeler le nombre de multiples de p dans un intervalle de la forme $[\![1, n]\!]$ avec n un entier naturel.
   b. En déduire le nombre de multiples de p dans un intervalle de la forme $[\![m, n]\!]$ avec $m \leq n$ deux entiers naturels.
   c. Montrer que $r_1(p) = r_2(p)$.
3. Montrer que, pour tout $k \in \mathbb{N} \setminus \{0\}$,
$$r_1(p^k) \leq r_2(p^k) \leq r_1(p^k) + 1.$$
   On pourra encadrer $\lfloor x \rfloor + \lfloor y \rfloor$ à partir de $\lfloor x + y \rfloor$.
4. Soit p un premier strictement plus grand que a.
   a. Montrer que $r_2(p) = 0$.
   b. En déduire que $r_1(p^k) = r_2(p^k) = 0$ pour tout $k \in \mathbb{N} \setminus \{0\}$.
5. Soit p un premier. Notons
$$\kappa(p) = \max\{k \in \mathbb{N},\ r_2(p^k) > 0\}.$$
   a. Montrer que
$$v_p(\pi_2/\pi_1) = \sum_{k=2}^{\infty} (r_2(p^k) - r_1(p^k)).$$
   b. En déduire que $v_p(\pi_2/\pi_1) \leq \kappa(p) - 1$.
6. Montrer que l'entier
$$\frac{\pi_2}{\prod\limits_{\substack{p \leq a \\ p \in \mathscr{P}}} p^{\kappa(p)}}$$
   divise
$$\frac{\pi_1}{\prod\limits_{\substack{p \leq a \\ p \in \mathscr{P}}} p}.$$
7. En déduire que l'hypothèse $a \leq b/2$ mène à une contradiction. Conclure.

●●○ **Problème 3 Théorème de Sophie Germain**

1. Soit q un nombre premier strictement supérieur à 5.
   a. Considérons un entier X premier avec q. Montrer que $X^{q-1} = 1[q]$.
   b. Déterminer tous les entiers $x \in [\![1, q-1]\!]$ tels que $x^2 = 1[q]$.
   c. En déduire que si le triplet $(X, Y, Z) \in \mathbb{Z}^3$ satisfait
$$X^{\frac{q-1}{2}} + Y^{\frac{q-1}{2}} + Z^{\frac{q-1}{2}} = 0[q],$$
   alors q divise XYZ.
2. On considère dorénavant un nombre premier p tel que $q = 2p + 1$ est aussi un nombre premier et un triplet d'entiers $(x, y, z) \in \mathbb{Z}^3$ tel que $x^p + y^p + z^p = 0$ tel que p ne divise pas xyz.
   a. Justifier qu'il suffit de considérer le cas où le PGCD de x, y, z est 1 (ce que l'on supposera dans la suite de l'exercice).
   b. Montrer que $y + z$ et $\sum\limits_{k=0}^{p-1} (-y)^k z^{p-1-k}$ sont premiers entre eux.
   c. En déduire qu'il existe $(A, T) \in \mathbb{Z}^2$ tels que $A \wedge T = 1$ et
$$y + z = A^p \qquad \sum_{k=0}^{p-1} (-1)^k y^k z^{p-1-k} = T^p.$$
   On pose de même $x + y = B^p$ et $x + z = C^p$.

PROBLÈMES

d. Montrer que q divise au moins l'un des trois entiers x, y ou z.
On suppose que q divise x.
e. Montrer que q divise ABC puis que q divise A.
f. En considérant T^p et B^p modulo q, exhiber une contradiction.
g. Conclure.

Problème 4 Suite de Lucas

Considérons les suites $(L_n)_{n\in\mathbb{N}}$, $(F_n)_{n\in\mathbb{N}}$ et $(S_n)_{n\in\mathbb{N}}$ définie par $L_0 = 2$, $L_1 = 1$, $F_0 = 0$, $F_1 = 1$ et, pour tout $n \in \mathbb{N}$:

$$L_{n+2} = L_{n+1} + L_n \qquad F_{n+2} = F_{n+1} + F_n \qquad S_n = L_{3.2^n}$$

Posons $\Phi = \frac{1+\sqrt{5}}{2}$, $\overline{\Phi} = \frac{1-\sqrt{5}}{2}$, $\omega = 2+\sqrt{5}$ et $\overline{\omega} = 2-\sqrt{5}$.

Partie A – Premières propriétés

1. a. Calculer, pour tout $n \in \mathbb{N}$, F_n et L_n en fonction de Φ et $\overline{\Phi}$.
   b. Vérifier que, pour tout $k, r \in \mathbb{N}$, $2L_{2k+r} - L_r L_{2k} = 5F_r F_k L_k$.
2. Montrer que pour tout $n \in \mathbb{N}$, $L_{2n} - L_n^2 = 2.(-1)^{n+1}$.
   En déduire une relation de récurrence simple vérifiée par la suite $(S_n)_n$.
3. Vérifier que $S_n = \omega^{2^n} + \overline{\omega}^{2^n}$ pour tout $n \in \mathbb{N}$.

Partie B – Relations de congruence

Dans cette partie on pensera à utiliser la relation obtenue à la question b. de la partie A.

1. Montrer que, pour tout $k, r \in \mathbb{N}$, $2L_{2k+r} = 2(-1)^{k+1} L_r [L_k]$.
2. Montrer, par récurrence sur $\alpha \in \mathbb{N}$, que L_k divise $2^\alpha L_{3^\alpha.k}$.
   Soit $n \geq 5$ un entier impair. Notons r son reste et $3^\alpha.q$, avec q non divisible par 3, son quotient dans la division euclidienne par 4.
3. Montrer que $L_n = -1[L_{2q}]$ ou que $L_n = -4[L_{2q}]$.

Partie C – Théorème de Cohn

1. Rappeler le petit théorème de Fermat.
2. a. Soit p un nombre premier congru à 3 modulo 4. Montrer qu'il n'existe pas d'entier $x \in \mathbb{N}$ tel que $x^2 = -1[p]$.
   b. Soit n un entier congru à 3 modulo 4. Montrer qu'il n'existe pas d'entier $x \in \mathbb{N}$ tel que $x^2 = -1[n]$.
3. Montrer que L_{2n} n'est pas le carré d'un entier.
   Indication : on pourra raisonner par l'absurde en utilisant la question 2. de la partie A.
4. Montrer que pour tout q non divisible par 3, $L_{2q} = 3[4]$.
   Indication : on pourra commencer par montrer que $L_{n+6} = L_n[4]$ pour tout entier n.
5. Déduire des résultats précédents que si $n \notin \{1,3\}$, L_n n'est pas le carré d'un entier.

Partie D – Test de Lucas-Lehmer pour les nombres de Mersenne

Le nombre de Mersenne d'indice $n \in \mathbb{N}$ est l'entier $M_n = 2^n - 1$.

1. Montrer que si M_n est premier, alors n est premier.
   Considérons p un nombre premier tel que M_p divise S_{p-2} et supposons que M_p est composé. Notons q un diviseur premier de M_p strictement inférieur à $\sqrt{M_p}$ et introduisons l'anneau $A = \{x + \sqrt{5}y,\ x, y \in \mathbb{Z}/q\mathbb{Z}\}$ avec $\sqrt{5}$ un élément de carré $5 \in \mathbb{Z}/q\mathbb{Z}$.
2. Montrer que A^* est de cardinal au plus $q^2 - 1$.
3. Montrer que le groupe engendré par $\omega = 2+\sqrt{5}$ contient 2^p éléments.
4. En déduire que M_p est alors premier.

CORRIGÉS

Corrigés des Vrai/Faux

a) Faux, le seul multiple de 0 est 0.

b) Faux, le calcul de la somme alternée des chiffres donne $1326543987321 = 3[11]$.

c) Vrai, les carrés sont congrus à 0 ou 1 modulo 4 donc la somme de deux carrés ne peut être égale à 3 modulo 4.

d) Vrai car $60 = 2^2.3^1.5^1$ donc le nombre de diviseurs positifs est $(2+1)(1+1)(1+1) = 12$.

e) Faux, par exemple, avec $a = u = v = 2$ et $b = -1$, $a \wedge b = 1$.

f) Faux, par exemple, $k = 4$ et $n = 6$.

g) Vrai avec la relation de Bezout $1.(n+1) - 1.n = 1$.

h) Vrai car un diviseur commun à n et $n+2$ divise $(n+2) - n = 2$ et 2 ne divise pas n.

i) Vrai, comme $3 \times (-8) = 1[25]$, les solutions sont $x = -16[25]$.

j) Faux, s'il y avait une solution x, 3 serait un diviseur de 2 car 3 divise $3x$ et 24.

k) Vrai.

l) Vrai.

m) Faux, 6 n'est pas premier et $2^5 = 32 = 2[6]$.

Corrigés des exercices d'application du chapitre 11

Exercice 1

On a évidemment $10 = -2[12]$ donc $10^2 = 4[12]$, $10^3 = 4[12]$, $10^4 = 4[12]$. Ainsi,

$$123a4 = 4.1 + 4.2 + 4.3 - 2a + 4[12]$$

Par conséquent, $123a4$ est divisible par 12 si, et seulement si, $4 - 2a = 0[12]$, c'est-à-dire $a \in \{2, 8\}$.

Exercice 2

▹ Si x est strictement inférieur à 10 (c'est-à-dire n'a qu'un seul chiffre), alors le nombre considéré $x10^2 + x10 + x = 111x = 37 \times 3x$ est divisible par 37.

▹ Si $x \in [\![10, 99]\!]$, écrivons $x = 10a + b$ avec a, $b \in [\![0, 9]\!]$. Le nombre considéré est

$$x10^4 + x10^2 + x = a(10^5 + 10^3 + 10) + b(10^4 + 10^2 + 1).$$

Or, $10^3 = 1 + 9 \times 111 = 1 + 27 \times 37 = 1[37]$, d'où

$$x10^4 + x10^2 + x = a(10^2 + 1 + 10) + b(10^1 + 10^2 + 1)[37] = (a+b)111[37] = 0[37]$$

est divisible par 37.

Exercice 3

Comme, pour tout $n \in \mathbb{N}$, $3^n = 1[2]$, un nombre est donc pair si le nombre de chiffres 1 qui apparaissent dans son écriture en base 3 est pair.

Exercice 4

Comme 2 et 7 sont premiers entre eux, le lemme de Gauss donne

$$\begin{aligned} n = 0[7] &\Leftrightarrow 2n = 0[7] \\ &\Leftrightarrow 2(10a+b) = 0[7] \\ &\Leftrightarrow -a+2b = 0[7] \\ &\Leftrightarrow a-2b = 0[7]. \end{aligned}$$

Par exemple, 2387 est divisible par 7 si, et seulement si, $238-2.7 = 224$ l'est si, et seulement si, $22-2.4 = 14$ l'est : ainsi, 2387 est divisible par 7.

Exercice 5

Commençons par étudier les puissances de 2 modulo 7 en remarquant que $2^3 = 1[7]$:

$k[3]$	0	1	2
$2^k[7]$	1	2	4

Il reste donc à étudier le reste de 4^n modulo 3. Comme $4^n = 1^n[3] = 1[3]$, $2^{4^n} - 2 = 2^1 - 2[7] = 0[7]$.

Exercice 6

Commençons par remarquer que, pour tout $n \in \mathbb{N}$,

$$\begin{aligned} n^3 - 3 &= (n-3+3)^3 - 3 \\ &= 3^3 - 3[n-3] \\ &= 24[n-3]. \end{aligned}$$

Ainsi, $n-3$ divise n^3-3 si, et seulement si, $n-3$ divise 24, soit encore $n-3 \in \{-3,-2,-1,1,2,3,4,6,8,12,24\}$ (car $n \geq 0$). En conclusion, les entiers recherchés sont 0, 1, 2, 4, 5, 6, 7, 9, 11, 15 et 27.

Exercice 7

1. En réduisant modulo 4 la relation $x^2 - 4y^2 = 1791$, on obtient $x^2 = 3[4]$ ce qui n'est jamais réalisé car un carré est congru à 0 ou 1 modulo 4 (selon sa parité).
2. En réduisant modulo 4 la relation $x^2 + y^2 = 1791$, on obtient $x^2 + y^2 = 3[4]$ ce qui n'est jamais réalisé pour les mêmes raisons.
3. En réduisant modulo 3 la relation $x^2 + 6y^2 = 1112$, on obtient $x^2 = 2[3]$ ce qui n'est jamais réalisé puisque $x^2 = 0[3]$ si $x = 0[3]$ et $x^2 = 1[3]$ sinon.

Exercice 8

La formule de Bernoulli donne

$$a^n - b^n = (a-b)\sum_{k=0}^{n-1} a^k b^{n-1-k}.$$

Mais $a = b[n]$ par hypothèse et

$$\sum_{k=0}^{n-1} a^k b^{n-1-k} = \sum_{k=0}^{n-1} a^k a^{n-1-k}[n] = na^{n-1}[n] = 0[n].$$

En conclusion, $a^n = b^n[n^2]$.

Exercice 9

1. Supposons que $n > 1$ est impair. La formule de Bernoulli donne $u_n = 2^n - (-1)^n = 3.\sum_{k=0}^{n-1} 2^k(-1)^{n-1-k}$: u_n n'est pas premier car divisible par $3 \neq u_n$.
2. La congruence $2^{2^q} = -1[u_{2^q}]$ entraîne $2^{2^q(2k+1)} = (-1)^{2k+1}[u_{2^q}] = -1[u_{2^q}]$ et donc u_{2^q} divise u_n : u_n n'est pas premier.
3. Il s'agit simplement de la contraposition de la question précédente.

Les nombres de Fermat sont les entiers $F_n = 2^{2^n} + 1$. Pierre de Fermat avait calculé $F_0 = 3$, $F_1 = 5$, $F_2 = 17$, $F_3 = 257$ et $F_4 = 65537$. Constatant qu'il s'agissait à chaque fois de nombres premiers, il avait conjecturé que c'était le cas pour tous les F_n. Leonhard Euler a infirmé cette intuition en montrant que 641 divisait F_5 (plus précisément, $F_5 = 641 \times 6700417$). Certains arguments récents laissent à penser qu'il n'y a pas d'autres nombres de Fermat premiers.

Exercice 10

On trouve rapidement les puissances suivantes (en notant que $3^4 = 1[10]$ et $7^4 = 1[10]$)

$n[4]$	0	1	2	3
$3^n[10]$	1	3	9	7
$7^n[10]$	1	7	9	3

On en déduit les valeurs de $3^a 7^b[10]$ reportées dans le tableau suivant où $a[4]$ est l'indice de ligne et $b[4]$ l'indice de colonne.

	0	1	2	3
0	1	7	9	3
1	3	1	7	9
2	9	3	1	7
3	7	9	3	1

Le seul moyen d'obtenir $1[10]$ correspond à $a = b[4]$.

Exercice 11

1. D'après les propriétés précédentes, on a, en regroupant les termes de mêmes valeurs

$$\nu_p(n!) = \sum_{l=1}^{n} \nu_p(l) = \sum_{k=1}^{\infty} k.\mathrm{card}\{l \leq n,\ \nu_p(l) = k\}.$$

Or, il y a $\left\lfloor \frac{n}{p^k} \right\rfloor$ multiples de p^k inférieurs à n donc $\left\lfloor \frac{n}{p^k} \right\rfloor - \left\lfloor \frac{n}{p^{k+1}} \right\rfloor$ entiers inférieurs à n de valuation p-adique égale à k. Ainsi

$$\begin{aligned}
\nu_p(n!) &= \sum_{k=1}^{\infty} k\left(\left\lfloor \frac{n}{p^k} \right\rfloor - \left\lfloor \frac{n}{p^{k+1}} \right\rfloor\right) \\
&= \sum_{k=1}^{\infty} k\left\lfloor \frac{n}{p^k} \right\rfloor - \sum_{k=2}^{\infty}(k-1)\left\lfloor \frac{n}{p^k} \right\rfloor \\
&= \sum_{k=1}^{\infty}\left\lfloor \frac{n}{p^k} \right\rfloor.
\end{aligned}$$

2. Le nombre cherché est égal à $\min(\nu_2(100!), \nu_5(100!)) = \nu_5(100!)$. D'après la proposition précédente,

$$\nu_5(100!) = \left\lfloor \frac{100}{5} \right\rfloor + \left\lfloor \frac{100}{5^2} \right\rfloor = 24.$$

Il y a donc 24 zéros à la fin de l'écriture décimale de 100!.

Exercice 12

Il suffit de montrer que, pour chaque fraction, le numérateur et le dénominateur sont premiers entre eux. Pour cela, utilisons le théorème de Bézout et les relations suivantes :

$$-2(15n+4)+3(10n+3)=1$$
$$-2(21n+4)+3(14n+3)=1.$$

Remarque

Ces exemples se généralisent immédiatement au cas d'une fraction $\frac{3qn+4}{2qn+3}$ avec $q \in \mathbb{N}$.

Exercice 13

Commençons par remarquer que a et $a+b$ sont premiers entre eux. En effet, d'après l'algorithme d'Euclide,

$$(a+b)\wedge a=(a+b-a)\wedge a=b\wedge a=1.$$

De même, b et $a+b$ sont premiers entre eux.

Comme $a+b$ est premier avec a et b, il l'est avec ab.

Remarque

Une démonstration alternative serait de partir d'une relation de Bézout $ua+vb=1$ et d'en déduire une relation de Bézout entre $a+b$ et ab, par exemple

$$(au^2+bv^2)(a+b)-(u-v)^2ab=1.$$

Exercice 14

Écrivons $a=42a'$ et $b=42b'$ avec $a'\wedge b'=1$. La seconde condition donne $a'\vee b'=40$ soit encore $a'b'=2^3.5$. Comme a' et b' sont premiers entre eux, ils n'ont pas de facteur premier en commun : les possibilités pour (a',b') sont $(1,40)$, $(8,5)$, $(5,8)$ et $(40,1)$.

En revenant à (a,b), on obtient les solutions

$$(42,1680),\ (336,210),\ (210,336),\ (1680,42).$$

Exercice 15

Notons $d=a\wedge b$ et $a=da'$, $b=db'$. Par construction, $a'\wedge b'=1$ et donc

$$(a+b)\wedge(a\vee b)=d\times(a'+b')\wedge(a'\vee b')=d\times(a'+b')\wedge(a'b').$$

Reste à montrer que $(a'+b')\wedge(a'b')=1$. Soit p un diviseur premier de $a'b'$. Comme a' et b' sont premiers entre eux, soit p divise a' mais pas b', soit p divise b' mais pas a' : dans les deux cas, p ne divise pas $a'+b'$. Par conséquent, $(a'+b')\wedge(a'b')=1$ et le PGCD recherché est $a\wedge b$.

Exercice 16

L'entier $x \wedge y$ divise y donc divise 4 : il est donc égal à 1, 2 ou 4.

▹ Si $x \wedge y = 1$, alors l'équation est $1 + xy = y + 4$, soit $y(x-1) = 3$. Les solutions (x, y) sont $(2,3)$ et $(4,1)$.

▹ Si $x \wedge y = 2$, alors on pose $x = 2x'$ et $y = 2y'$ avec $x' \wedge y' = 1$. L'équation se réécrit alors $2+2x'y' = 2y'+4$ soit $y'(x'-1) = 1$. La seule solution (x', y') est $(2,1)$ et donc la seule solution (x, y) est $(4,2)$.

▹ Si $x \wedge y = 4$, alors on pose $x = 4x'$ et $y = 4y'$ avec $x' \wedge y' = 1$. L'équation se réécrit alors $4+4x'y' = 4y'+4$ soit $y'(x'-1) = 0$. Les solutions (x', y') sont alors $(1, k)$ pour $k \in \mathbb{N}$ et $(1,0)$ (la valeur de x' est ici imposée par la condition $x' \wedge y' = 1$). Les solutions (x, y) sont alors les $(4, 4k)$ pour $k \in \mathbb{N}$ et $(4,0)$;

En conclusion, les solutions sont

$$(2,3),\ (4,1),\ (4,2) \text{ et } (4,4k) \text{ pour } k \in \mathbb{N}.$$

Exercice 17

1. Remarquons que le PGCD de 323 et 391 est 17 et qu'il divise 612. En divisant par 17, l'équation devient $19x - 23y = 36$. Une solution de cette équation est $(-9,-9)$. Par conséquent, (x, y) est solution si, et seulement si, $19(x+9) = 23(y+9)$. Comme 19 est premier avec 23, le lemme de Gauss entraîne l'existence d'un entier k tel que $y = -9 + 19k$ et donc, en reportant, $x = -9 + 23k$. En conclusion, l'ensemble des solutions est

$$\{(-9+23k, -9+19k),\ k \in \mathbb{Z}\}.$$

2. On cherche les entiers n tels qu'il existe des entiers x et y vérifiant $n = 270 + 323x = 100 + 391y$. Ceci entraîne $323x - 391y = -170$ et donc $323x - 391(y-2) = 612$. D'après la première question, il existe un entier k tel que $x = -9 + 23k$ et donc, en reportant, $n = -2637 + 7429k$.
On vérifie laborieusement que ces entiers conviennent.

Exercice 18

Remarquons que $399 = 3 \times 7 \times 19$. Considérons $a_1, \dots a_{1789} \in \mathbb{Z}$ tels que $\sum_{k=1}^{1789} a_k = 0$. Pour montrer que la somme $S = \sum_{k=1}^{1789} a_k^{37}$ est divisible par 399, il suffit de montrer qu'elle est divisible par 3, 7 et 19.

▹ Utilisons le théorème de Fermat avec $p = 3$:

$$\begin{aligned} S &= \sum_{\substack{k=1 \\ a_k \neq 0[3]}}^{1789} a_k^{37}[3] = \sum_{\substack{k=1 \\ a_k \neq 0[3]}}^{1789} a_k^{2\times 18+1}[3] = \sum_{\substack{k=1 \\ a_k \neq 0[3]}}^{1789} a_k^{1}[3] \\ &= \sum_{k=1}^{1789} a_k[3] = 0[3]. \end{aligned}$$

Ainsi, 3 divise S.

▹ On procède de même pour 7 et 19 en remarquant que $37 = 6 \times 6 + 1$ et $37 = 2 \times 18 + 1$.

Exercice 19

1. Un entier naturel divise n si, et seulement si, il est de la forme $p_1^{\beta_1} \dots p_k^{\beta_k}$ avec, pour tout $i \leq k$, $0 \leq \beta_i \leq \alpha_i$. Ainsi, $d(n)$ est le cardinal de

$$\{0, \dots, \alpha_1\} \times \dots \times \{0, \dots, \alpha_k\}$$

donc $d(n) = \prod_{i=1}^{k} (\alpha_i + 1)$.

CORRIGÉS

2. L'hypothèse se réécrit $p_i^\varepsilon \geq 2$ (car $x \mapsto x^\varepsilon$ croissante sur $\mathbb{R}_+$) et donc $p_i^{\varepsilon\alpha_i} \geq 2^{\alpha_i}$. Il suffit alors de vérifier par récurrence sur n que $2^n \geq n+1$. C'est vrai pour $n=0$ et si $2^n \geq n+1$, alors

$$2^{n+1} = 2 \times 2^n \geq 2(n+1) = n+2+n \geq n+2.$$

En conclusion, $p_i^{\varepsilon\alpha_i} \geq \alpha_i + 1$.
3. Cette fonction est dérivable comme produit et composition de fonctions dérivables. De plus, pour tout $x > 0$, $f'(x) = (1-\varepsilon\ln(2)(x+1))2^{-\varepsilon x}$.
Ainsi, f est croissante jusqu'en $x_0 = \frac{1-\varepsilon\ln 2}{\varepsilon\ln 2}$ (on notera y_0 sa valeur en ce point), puis décroissante.
4. Considérons $I_\varepsilon = \{i \leq k,\ p_i \geq 2^{1/\varepsilon}\}$ et $J_\varepsilon = \{i \leq k,\ p_i < 2^{1/\varepsilon}\}$.

$$\begin{aligned}\frac{d(n)}{n^\varepsilon} &= \left(\prod_{i\in I_\varepsilon} \frac{\alpha_i+1}{p_i^{\varepsilon\alpha_i}}\right)\left(\prod_{i\in J_\varepsilon} \frac{\alpha_i+1}{p_i^{\varepsilon\alpha_i}}\right)\\ &\leq \left(\prod_{i\in J_\varepsilon} \frac{\alpha_i+1}{p_i^{\varepsilon\alpha_i}}\right) \leq \left(\prod_{i\in J_\varepsilon} \frac{\alpha_i+1}{2^{\varepsilon\alpha_i}}\right) \leq y_0^{|J_\varepsilon|}\end{aligned}$$

en utilisant tour à tour la première question pour exprimer $d(n)$, la deuxième pour gérer les termes d'indices dans I_ε et le maximum de f. Comme le nombre de nombres premiers inférieurs à $2^{1/\varepsilon}$ est majoré par exemple par $\lfloor 2^{1/\varepsilon}\rfloor$, on obtient un majorant $C = y_0^{\lfloor 2^{1/\varepsilon}\rfloor}$, indépendant de n, de la quantité $\frac{d(n)}{n^\varepsilon}$.

Corrigés des exercices d'approfondissement du chapitre 11

Exercice A

Raisonnons par l'absurde et supposons qu'il existe des entiers naturels m, n et a tels que $3^m+3^n+1 = a^2$.
Remarquons dans un premier temps que 3^m+3^n+1 est impair donc a^2 et a sont impairs. Notons $a = 2k+1$ avec $k \in \mathbb{N}$. Alors, $a^2 = 4k^2+4k+1$ et l'égalité de départ se réécrit $3^m+3^n = 4k(k+1)$.en particulier, 3^m+3^n est un multiple de 8 (puisque k ou $k+1$ est pair).
Or, les puissances de 3 valent 1 modulo 8 si l'exposant est pair, 3 modulo 8 sinon. En aucun cas, la somme de deux puissances de 3 ne peut-être un multiple de 8 : contradiction.

Exercice B

Effectuons le changement d'indice $k = n+1-j$

$$\begin{aligned}\sum_{k=1}^n k^{n+1} &= \sum_{j=1}^n (n+1-j)^{n+1}\\ &= \sum_{j=1}^n (-j)^{n+1}[n+1]\\ &= -\sum_{j=1}^n j^{n+1}[n+1],\end{aligned}$$

car $n+1$ est impair. Par conséquent, $n+1$ divise $2\sum_{k=1}^n k^{n+1}$ et $(n+1)\wedge 2 = 1$ (car $n+1$ est impair) : d'après la lemme de Gauss, $n+1$ divise $\sum_{k=1}^n k^{n+1}$.

Exercice C

1. Procédons par récurrence sur n.
▹ Le résultat est évident pour $n=1$ car 2 divise $u_1=2$.
▹ Soit $n\in\mathbb{N}$ tel que 2^n divise u_n. Si 2^{n+1} divise u_n, alors 2^{n+1} divise $2.10^n+u_n=u_{n+1}$; sinon, $u_n=2^n(2k+1)$ avec $k\in\mathbb{N}$ et donc $u_{n+1}=10^n+u_n=2^n(5^n+1+2k)$ est divisible par 2^{n+1} (car 5^n+1+2k est pair).
Ainsi, pour tout $n\geq 1$, 2^n divise u_n.
2. Une simple récurrence permet de montrer que u_n admet une écriture décimale avec n chiffres égaux à 1 ou 2.
▹ Le résultat est évident pour $n=1$ car $u_1=2$.
▹ Soit $n\in\mathbb{N}$ tel que 2^n divise u_n. Si 2^{n+1} divise u_n, l'écriture décimale de u_{n+1} s'obtient en ajoutant un 2 à gauche de celle de u_n ; sinon, en ajoutant un 1 à gauche. D'où le résultat au rang $n+1$.
D'après la première question, pour tout $n\geq 1$, u_n est un multiple de 2^n dont l'écriture décimale ne comporte que des 1 et des 2.

Exercice D

Raisonnons par récurrence.
▹ Le résultat est évident pour $n=1$.
▹ Soit $n\in\mathbb{N}\setminus\{0\}$ tel que $\displaystyle\sum_{k=n+1}^{2n} u_k=n^2$. Alors,

$$\begin{aligned}\sum_{k=n+2}^{2n+2} u_k &= \sum_{k=n+1}^{2n} u_k + u_{2n+2}+u_{2n+1}-u_{n+1}\\ &= n^2+2n+1=(n+1)^2,\end{aligned}$$

car $u_{2n+2}=u_{n+1}$ et $u_{2n+1}=2n+1$.
Le résultat est établi pour tout $n\in\mathbb{N}$.

Exercice E

1. Raisonnons par l'absurde et supposons qu'il existe u, $v\in\mathbb{N}$ tels que $ab-a-b=u_0a+v_0b$, c'est-à-dire $ab=(u+1)a+(v+1)b$.
On remarque que a divise $(v+1)b$. Comme a et b sont premiers entre eux, le lemme de Gauss indique que a divise $v+1$ donc $v+1\geq a$. Par conséquent,

$$ab=(u+1)a+(v+1)b\geq a+ab$$

contradiction !
2. ▹ Partons d'une relation de Bézout entre a et b et multiplions la par n : il existe u et $v\in\mathbb{Z}$ tels que $ua+vb=n$. Divisons alors v par a : il existe des entiers q et r tels que $v=qa+r$ et $0\leq r<a$. La relation de Bézout de départ se réécrit alors $(u+qb)a+rb=n$: elle est bien de la forme demandée.
▹ Supposons qu'il existe u_0, v_0, u_1, v_1 des entiers tels que $u_0a+v_0b=u_1a+v_1b$, $0\leq v_0<a$ et $0\leq v_1<a$. Alors, $(u_0-u_1)a=(v_1-v_0)b$ et donc a divise v_0-v_1 (lemme de Gauss avec a, b premiers entre eux). Comme $v_0-v_1\in[\![-(a-1),(a-1)]\!]$, $v_0-v_1=0$ et par suite $u_0=u_1$. Cela est bien l'unicité demandée.
3. Le sens retour est évident car dans ce cas $u_0\geq 0$ et $v_0\geq 0$.
Pour le sens direct, procédons comme précédemment en partant de u et $v\in\mathbb{N}$ tels que $ua+vb=n$: on obtient (avec les mêmes notations) la décomposition $(u+qb)a+rb=n$ d'où $v_0=r$ et $u_0=u+qb$. L'entier u_0 est alors positif comme somme de termes positifs (q est positif comme quotient dans la division de deux entiers positifs).
4. Raisonnons par contraposition, si $n\notin E$, alors $u_0\leq -1$ d'après la question précédente. Ainsi, $n=u_0a+v_0b\leq -a+(a-1)b=ab-a-b$.

Exercice F

1. Soit p un diviseur premier impair de a^2+1 : p n'est pas pair ; si p est de la forme $4k+3$ alors

$$(a)^{p-1} = (a)^{4k+2} = (-1)^{2k+1}[p] = -1[p],$$

ce qui contredit le théorème de Fermat. Par conséquent, tous les diviseurs premiers d'un entier de la forme a^2+1 sont congrus à 1 modulo 4.
2. Raisonnons par l'absurde. Supposons que cet ensemble est fini $\{p_1, \dots, p_N\}$. D'après le lemme, l'entier $4\left(\prod_{k=1}^{N} p_k\right)^2 + 1$ admet un diviseur premier de la forme $4n+1$ qui n'est pas dans la liste précédente : contradiction.

Corrigés des problèmes du chapitre 11

Problème 1

1. a. ▹ Soit $a \in [\![1, n-1]\!]$. Comme $a \wedge n = 1$ (car n est premier), il existe des entiers u, $v \in \mathbb{Z}$ tels que $ua + vb = 1$ donc $ua = 1[n]$; quitte à remplacer u par son reste b dans la division par n, on obtient $ba = 1[N]$ et $b \in [\![0, n-1]\!]$. Enfin, il est clair que $b \neq 0$ dans cette égalité car $0 \neq 1[n]$.
▹ Supposons qu'il existe b, $b' \in [\![1, n-1]\!]$ tels que $ab = 1[n]$, et $ab' = 1[n]$. Alors, n divise $a(b-b')$. Comme $a \wedge n = 1$, le lemme de Gauss entraîne que n divise $b-b'$. Or, $b-b' \in [\![-(n-1), n-1]\!]$ donc $b-b' = 0$ (le seul multiple de n dans cet intervalle) : d'où l'unicité attendue.
b. ▹ Commençons par chercher les entiers a tels que $a^2 = 1[n]$ (c'est-à-dire, avec les notations de la question précédente, ceux pour lesquels $a = b$). La condition équivaut à n divise $(a+1)(a-1)$, soit, comme n est premier, n divise $a+1$ ou $a-1$ c'est-à-dire $a = n-1$ ou $a = 1$.
▹ Calculons $(n-1)!$ en regroupant chaque terme a avec son b qui lui est associé par la question précédente, il ne reste d'après le premier point que deux éléments « isolés » 1 et $n-1$ donc $(n-1)! = 1 \times (n-1)[n]$ soit enfin $(n-1)! = -1[n]$.
2. a. D'après le théorème de Bézout, n est premier avec $(n-m)!$ donc avec tous les entiers $k \in [\![1, n-m]\!]$.
b. Si $k \in [\![n-m+1, n-1]\!]$, alors $n-k \in [\![1, m-1]\!]$. Or d'après la propriété de Bézout $n \wedge (m-1)! = 1$ donc $m \wedge (n-k) = 1$.
Comme $n \wedge (n-k) = n \wedge k$, n est premier avec k.
c. n est premier avec tous les entiers de $[\![1, n-1]\!]$ donc est premier.

Problème 2

Partie A – Préliminaires

1. Considérons un nombre premier p strictement supérieur à $\max(a, b)$ (ce qui est possible car l'ensemble des nombres premiers est infini). Alors $a_p = a$ et $b_p = b$, donc l'hypothèse entraîne $a \leq b$.
2. ▹ Comme $a \leq b$ et $b \neq a$ (hypothèse du raisonnement par l'absurde) : $b - a \in \mathbb{N} \setminus \{0\}$.
▹ Soit p un nombre premier. En additionnant les congruences, $b - a = b_p - a_p[p]$; or, $b_p - a_p \in [\![0, b_p]\!] \subset [\![0, p-1]\!]$ donc $(b-a)_p = b_p - a_p \leq b_p$.
En conclusion, les entiers $b-a$ et b vérifient les hypothèses de l'énoncé.
3. Si $a > b/2$, alors $b - a \leq b/2$ et on s'est ramené à un couple $(b-a, b)$ vérifiant les mêmes hypothèses avec la condition additionnelle que la première coordonnée est inférieure ou égale à la moitié de la seconde.

Partie B – Démonstration de Szegedy

1. On trouve $\pi_1 = a!$ et $\pi_2 = \frac{b!}{(b-a)!}$. Le quotient est le coefficient binomial $\binom{b}{a}$ donc est entier.
2. a. Le nombre de multiples de p dans l'intervalle $[\![1, n]\!]$ est $\left\lfloor \frac{n}{p} \right\rfloor$.
b. Le nombre de multiples de p dans l'intervalle $[\![m, n]\!]$ se déduit du nombre de multiples de p dans $[\![1, n]\!]$ et dans $[\![1, m-1]\!]$; ce nombre est donc

$$\left\lfloor \frac{n}{p} \right\rfloor - \left\lfloor \frac{m-1}{p} \right\rfloor.$$

c. Notons $a = \alpha p + a_p$ et $b = \beta p + b_p$. Alors

$$\begin{aligned} r_1(p) &= \left\lfloor \frac{a}{p} \right\rfloor = \alpha \\ r_2(p) &= \left\lfloor \frac{b}{p} \right\rfloor - \left\lfloor \frac{b-a}{p} \right\rfloor \\ &= \beta - \left(\beta - \alpha - \left\lfloor \frac{b_p - a_p}{p} \right\rfloor \right) = \alpha. \end{aligned}$$

Pour la dernière étape, on a exploité l'inégalité

$$0 \leq \frac{b_p - a_p}{p} < 1.$$

Par conséquent, $r_1(p) = r_2(p)$.
3. ▹ Commençons par remarquer que pour tous $x, y \in \mathbb{R}$,

$$\lfloor x \rfloor + \lfloor y \rfloor \leq x + y < \lfloor x \rfloor + \lfloor y \rfloor + 2$$

donc

$$\lfloor x \rfloor + \lfloor y \rfloor \leq \lfloor x + y \rfloor \leq \lfloor x \rfloor + \lfloor y \rfloor + 1.$$

En recentrant sur le terme qui nous intéresse

$$\lfloor x + y \rfloor - 1 \leq \lfloor x \rfloor + \lfloor y \rfloor \leq \lfloor x + y \rfloor.$$

▹ D'après un dénombrement analogue à celui de la question 2,

$$\begin{aligned} r_2(p^k) - r_1(p^k) &= \left\lfloor \frac{b}{p^k} \right\rfloor - \left\lfloor \frac{b-a}{p^k} \right\rfloor - \left\lfloor \frac{a}{p^k} \right\rfloor \\ &\in \{0, 1\} \end{aligned}$$

d'après l'encadrement précédent.
4. a. Comme p est strictement plus grand que a, il n'y a pas de multiple de p dans E_1 donc $r_1(p) = 0$. D'après la question 2, $r_2(p) = r_1(p) = 0$.
b. On vient d'établir qu'il n'y avait pas de multiple de p dans les ensembles E_1 et E_2. Il n'y a donc a fortiori pas de multiple de p^k. D'où $r_1(p^k) = r_2(p^k) = 0$.
5. a. Pour $i \in \{1, 2\}$,

$$v_p(\pi_i) = \sum_{n \in E_i} v_p(n) = \sum_{n \in E_i} \sum_{k=1}^{\infty} \mathbb{1}_{p^k | n} = \sum_{k=1}^{\infty} \sum_{n \in E_i} \mathbb{1}_{p^k | n} = \sum_{k=1}^{\infty} r_i(p^k).$$

Par conséquent,

$$v_p(\pi_2/\pi_1) = v_p(\pi_2) - v_p(\pi_1) = \sum_{k=1}^{\infty} (r_2(p^k) - r_1(p^k)).$$

On conclut en remarquant que, pour $k = 1$, $r_2(p^k) = r_1(p^k)$ (question 2).

b. En majorant chaque terme de la somme précédente par 1, on obtient $\nu_p(\pi_2/\pi_1) \leq \kappa(p)-1$.
6. D'après la question 4, les seuls facteurs premiers de π_1 et π_2 sont inférieurs ou égaux à a. De plus, la valuation p-adique du quotient est majorée par $\kappa(p)-1$. Par conséquent, $\frac{\pi_2}{\pi_1}$ divise $\prod_{\substack{p\leq a\\ p\in\mathscr{P}}} p^{\kappa_p-1}$ de qui en remettant dans l'ordre idoine donne

$$\frac{\pi_2}{\prod_{\substack{p\leq a\\ p\in\mathscr{P}}} p^{\kappa(p)}} \mid \frac{\pi_1}{\prod_{\substack{p\leq a\\ p\in\mathscr{P}}} p}.$$

7. Notons, conformément à l'usage, $\pi(a)$ le nombre de premiers inférieurs ou égaux à a. Le terme de droite est, en majorant chaque terme par a,

$$\frac{\pi_1}{\prod_{\substack{p\leq a\\ p\in\mathscr{P}}} p} = \prod_{\substack{n\leq a\\ n\notin\mathscr{P}}} n \leq a^{a-\pi(a)}.$$

Dans le terme de gauche, il y a par définition de $\kappa(p)$ un facteur divisible par $\kappa(p)$. Quitte à simplifier ces (au plus) $\pi(a)$ termes, on obtient la minoration

$$\frac{\pi_2}{\prod_{\substack{p\leq a\\ p\in\mathscr{P}}} p^{\kappa(p)}} \geq (b-a+1)^{a-\pi(a)}.$$

Or, comme $b \geq 2a$,

$$\frac{\pi_2}{\prod_{\substack{p\leq a\\ p\in\mathscr{P}}} p^{\kappa(p)}} \geq (a+1)^{a-\pi(a)} > a^{a-\pi(a)} \geq \frac{\pi_1}{\prod_{\substack{p\leq a\\ p\in\mathscr{P}}} p}$$

ce qui contredit la divisibilité précédente et termine le raisonnement par l'absurde. Le théorème de Erdös et Pálfy est établi.

Problème 3

1. a. C'est le petit théorème de Fermat.
b. L'équation $x^2 = 1 \mod q$ est équivalente à $(x-1)(x+1) = 0[q]$. Comme q est premier, les seules solutions sont $x = 1$ ou $x = q-1$.
c. Si q ne divise pas XYZ, il est premier avec X Y et Z ; chacun des termes de la somme $X^{\frac{q-1}{2}} + Y^{\frac{q-1}{2}} + Z^{\frac{q-1}{2}}$ vaut ± 1 donc la somme ne peut être égale à 0 modulo q (car $q \geq 5$).
2. a. Si $x \wedge y \wedge z \neq 1$, on peut diviser par cette valeur et on retrouve des nouveaux entiers x', y', z' solution de l'équation de Fermat et avec p ne divisant pas $x'y'z'$.
b. Si $c \neq p$ est un diviseur premier de $y+z$, alors $z = -y \mod c$ et par conséquent,

$$\sum_{k=0}^{p-1} (-y)^k z^{p-1-k} = pz^{p-1}[c] \neq 0[c]$$

(sinon c divise z et donc y or $y+z$ divise x^p donc c divise x : on obtient que c divise $x \wedge y \wedge z = 1$ contradiction). Les diviseurs de $y+z$ ne sont pas diviseurs de $\sum_{k=0}^{p-1}(-y)^k z^{p-1-k}$ d'où le résultat.
c. On remarque que

$$(y+z)\Big(\sum_{k=0}^{p-1}(-y)^k z^{p-1-k}\Big) = -x^p.$$

En utilisant la décomposition en facteurs premiers, on trouve que chacun des termes est une puissance $p^{\text{ième}}$.
d. On utilise l'équation de Fermat et la question 1(c).

e. On remarque que $B^p + C^p - A^p = 2x = 0[q]$ donc d'après la question 1(c), q divise ABC.
Si q divise B alors q divise $B^p - x = y$ et donc q divise z (d'après l'équation de Fermat) : contradiction car $x \wedge y \wedge z = 1$. Donc q ne divise pas BC et donc (théorème de Gauss) q divise A.
f. $T^p = py^{p-1} \mod q$ et $B^p = y[q]$. En utilisant la question 1(b), on en déduit que $p = \pm 1[q]$: contradiction avec $q = 2p+1$.
g. Soit p un nombre premier tel que $2p+1$ soit premier. Si l'équation $x^p + y^p + z^p = 0$ admet des solutions entières, alors p divise xyz.

Problème 4

Partie A – Premières propriétés

1. a. L'équation caractéristique des suites récurrentes linéaires d'ordre 2 $(F_n)_n$ et $(L_n)_n$ est $x^2 - x - 1 = 0$. Ces suites sont donc de la forme $(A\Phi^n + B\overline{\Phi}^n)_n$. En calculant les premiers termes, on trouve $A = B = 1$ pour $(L_n)_n$ et $A = -B = \frac{1}{\sqrt{5}}$ pour $(F_n)_n$.
b. Pour tous k, $r \in \mathbb{N}$,

$$\begin{aligned} 5F_rF_kL_k &= (\Phi^{2k} - \overline{\Phi}^{2k})(\Phi^r - \overline{\Phi}^r) \\ 2L_{2k+r} - L_rL_{2k} &= 2(\Phi^{2k+r} + \overline{\Phi}^{2k+r}) - (\Phi^r + \overline{\Phi}^r)(\Phi^{2k} + \overline{\Phi}^{2k}) \\ &= \Phi^{2k+r} + \overline{\Phi}^{2k+r} - \Phi^r\overline{\Phi}^{2k} - \overline{\Phi}^r\Phi^{2k} \end{aligned}$$

D'où l'égalité recherchée.
2. Pour tout $n \in \mathbb{N}$,

$$L_{2n} - L_n^2 = \Phi^{2n} + \overline{\Phi}^{2n} - (\Phi^n + \overline{\Phi}^n)^2 = -(\Phi\overline{\Phi})^n = 2.(-1)^{n+1}$$

Par conséquent, $S_{n+1} = L_{2.3.2^n} = (L_{3.2^n})^2 + 2(-1)^{3.2^n+1} = S_n^2 - 2$ pour $n > 0$ et $S_1 = S_0 + 2$.
3. Procédons par récurrence. $S_0 = L_3 = 4$ et $\omega + \overline{\omega} = 4$, d'où l'initialisation. Soit $n \in \mathbb{N}$ et supposons la formule établie jusqu'au rang n. Alors,

$$\begin{aligned} S_{n+1} &= (\omega^{2^n} + \overline{\omega}^{2^n})^2 - 2 \\ &= \omega^{2^{n+1}} + \overline{\omega}^{2^{n+1}} + 2(\omega\overline{\omega})^{2^n} - 2. \end{aligned}$$

On conclut en remarquant que $\omega\overline{\omega} = -1$.

Partie B – Relations de congruence

1. Pour tout $k, r \in \mathbb{N}$, $2L_{2k+r} = L_rL_{2k} + 5F_rF_kL_k = L_rL_{2k}[L_k]$; or, $L_{2k} = L_k^2 + 2(-1)^{k+1} = 2(-1)^{k+1}[L_k]$. En conclusion, $2L_{2k+r} = 2(-1)^{k+1}L_r[L_k]$.
2. Le résultat est évident pour $\alpha = 0$. Supposons que la propriété soit vraie pour $\alpha \in \mathbb{N}$. Alors $2^{\alpha+1}L_{3^{\alpha+1}k} = 2^\alpha(2L_{2.3^\alpha k+3^\alpha k}) = 2.(-1)^{3^\alpha k+1}2^\alpha L_{3^\alpha k}[L_{3^\alpha k}] = 0[L_k]$; d'où l'hérédité.
3. Commençons par calculer $2L_n$:

$$\begin{aligned} 2L_n &= 2L_{4.3^\alpha q+r} = 2L_{2.2.3^\alpha q+r} \\ &= 2(-1)^{2.3^\alpha q+1}L_r[L_{2.3^\alpha q}] \\ &= -2L_r[L_{2.3^\alpha q}] \end{aligned}$$

Par ailleurs, nous avons montré que L_{2q} divise $2^\alpha L_{2.3^\alpha q}$ et donc $L_{2.3^\alpha q}$ car L_{2q} est impair (puisque $q \neq 0[3]$). À ce stade nous avons montré $2L_n = -2L_r[L_{2q}]$ et donc $L_n = -L_r[L_{2q}]$ (toujours car L_{2q} est impair). Enfin, notons que $r \in \{1,3\}$ (c'est le reste de la division par 4 de l'entier impair n) et que $L_1 = 1$ et $L_3 = 4$. D'où le résultat annoncé.

Partie C – Théorème de Cohn

1. Voir le cours.
2. a. Si $x^2 = -1[p]$, alors $x^4 = 1[p]$ et donc 4 divise $p-1$ (théorème de Lagrange) ce qui est impossible.
b. $n = 3[4]$ implique que n admet un diviseur premier p congru à 3 modulo 4. Si $x^2 = -1[n]$, alors $x^2 = -1[p]$ d'où la contradiction.
3. Si $L_{2n} = x^2$, alors $x^2 - L_n^2 = (x - L_n)(x + L_n) = 2(-1)^{n+1}$. Or, $x - L_n$ et $x + L_n$ ont la même parité et le reste modulo 4 de leur produit n'est pas 2. D'où la contradiction.
4. La condition $L_{n+6} = L_n[4]$ est vérifiée pour $n \in \{0, 1\}$ donc par construction pour tout $n \in \mathbb{N}$.
Un calcul des premiers termes donne

n	0	1	2	3	4	5
$L_n[4]$	2	1	3	0	3	3

D'où le résultat.
5. Si n est pair, le résultat a déjà été établi. Supposons dorénavant n impair supérieur ou égal à 5. Posons alors, comme à la partie B, $n = r + 4.3^\alpha q$.
- Si $L_n = -1[L_{2q}]$, alors L_n n'est pas un carré car $L_{2q} = 3[4]$.
- Si $L_n = -4[L_{2q}]$, alors, en posant $L_{2q} = 2m+1$, on a $2m = -1[L_{2q}]$ et donc $4m^2 = 1[L_2 q]$. En combinant les deux égalités modulo L_{2q}, on obtient que $m^2 L_n = -1[L_{2q}]$ et que $m^2 L_n$ (et donc L_n) n'est pas un carré.

Partie D – Test de Lucas-Lehmer pour les nombres de Mersenne

1. Par contraposition, si n n'est pas premier, alors il existe p, $q > 1$ tels que $n = pq$ et donc, d'après la formule de Bernoulli,

$$2^n - 1 = (2^p - 1)\Big(\sum_{k=0}^{q-1} 2^k\Big)$$

n'est pas premier.
2. Comme l'anneau A est de cardinal au plus q^2, A^* est de cardinal au plus $q^2 - 1$ (puisque 0 n'est pas inversible).
3. La congruence $S_{p-2} = 0[q]$ devient, en utilisant l'expression explicite de la suite établie à la partie A, $\omega^{2^{p-2}} + \overline{\omega}^{2^{p-2}} = 0[q]$. En multipliant par $\omega^{2^{p-2}}$, on obtient $\omega^{2^{p-1}} + 1 = 0[q]$ et donc $\omega^{2^p} = 1[q]$. Par conséquent, le groupe engendré par ω contient exactement 2^p éléments.
4. D'après les questions précédentes, $2^p \leq q^2 - 1 < M_p - 1 = 2^p - 2$. Cette contradiction entraîne que M_p est premier.

Chapitre 12

Polynômes

1. Définitions et premières propriétés — **2.** Racines de polynômes — **3.** Arithmétique des polynômes

Objectifs et compétences du programme

- Effectuer la division entre deux polynômes.
- Calculer le PGCD de polynômes.
- Comprendre le lien entre PGCD et racines.
- Exploiter les relations coefficients-racines.
- Trouver la multiplicité d'une racine donnée.
- Décomposer un polynôme simple en facteurs irréductibles dans $\mathbb{R}[X]$ ou $\mathbb{C}[X]$.

Carl Friedrich Gauss, mathématicien allemand (1777-1855) est de ces génies des mathématiques que l'on retrouve dans tous les chapitres d'un cours.
Son premier livre,*Disquisitiones Arithmeticae* publié alors qu'il n'a que 24 ans, demeure aujourd'hui une grande référence de l'arithmétique : c'est dans ces pages que l'on trouve pour la première fois la notion de congruence (et la présentation est moderne). Outre cette réussite mathématique, il est connu pour avoir bouleversé l'astronomie depuis son poste à l'observatoire de Göttingen, l'électromagnétisme en collaboration avec Wilhelm Weber, les statistiques avec la méthode des moindres carrés ; il a également travaillé sur le terrain la topographie et la cartographie et en a profité pour fournir les outils pour l'étude des systèmes linéaires...
Vouloir établir la liste des théorèmes ou notions qui portent son nom serait vain.

COURS

Les polynômes sont des objets algébriques relativement simples en apparence mais à l'étonnante complexité : pour les maîtriser complètement, il faut faire appel à des raisonnements arithmétiques (tout ce qui concerne la divisibilité et l'irréductibilité), linéaires (la structure d'espace vectoriel sous-jacente qui sera abordée dans un chapitre ultérieur) et analytiques (notamment lorsque l'on considère la fonction polynomiale associée à un polynôme pour gérer les racines ou les « valeurs » d'un polynôme). Il faut savoir quelles connaissances mobiliser pour résoudre chaque question posée.

1. Définitions et premières propriétés

Dans la première section, on définit les polynômes comme les suites nulles à partir d'un certain rang et on vérifie un certain nombre de propriétés intuitives à partir de cette définition. Ceci n'est pas exigible, ni nécessaire pour la suite du cours.

L'idée de cette construction est de définir un polynôme comme la suite de ces coefficients. Par exemple, le polynôme habituellement écrit X^4+3X^2-2X est visualisé comme

$$0+(-2)X+3X^2+0X^3+1X^4+0X^5+0X^6+\dots$$

(les termes non écrits sont tous nuls) c'est-à-dire défini par la suite

$$(0,-2,3,0,1,0,0,0,\dots).$$

1.1. Contructions

Définition 12.1.

▹ Un polynôme à coefficients dans un corps $\mathbb{K}\subset\mathbb{C}$ est une suite presque nulle d'éléments de $\mathbb{K}$, c'est-à-dire une suite nulle à partir d'un certain rang. Les termes de cette suite sont appelés coefficients du polynôme.
▹ Le polynôme nul est la suite nulle.
▹ Le dernier coefficient non nul d'un polynôme non nul est appelé coefficient dominant de ce polynôme. Si ce coefficient est égal à 1, le polynôme est unitaire.
▹ L'ensemble des polynômes à coefficients dans un corps $\mathbb{K}$ est noté $\mathbb{K}[X]$.

Définition 12.2.

Soit $P=(p_n)_n$, $Q=(q_n)_n$ deux polynômes à coefficients dans un corps $\mathbb{K}$ et $\lambda\in\mathbb{K}$.
▹ La somme de P et Q est le polynôme noté $P+Q$ défini par $(p_n+q_n)_n$.
▹ Le produit $\lambda.P$ est le polynôme défini par $(\lambda p_n)_n$.
▹ Le produit de P et Q est le polynôme noté $P.Q$ défini par $\left(\sum_{k=0}^{n} p_k q_{n-k}\right)_n$.

On voit ci-dessous pourquoi le produit de deux suites presque nulles est bien encore une suite presque nulle.

Définition 12.3.

Le degré d'un polynôme non nul P est l'indice, noté $\deg P$, du dernier coefficient non nul. Par convention, le degré du polynôme nul est $-\infty$.
L'ensemble des polynômes à coefficients dans un corps $\mathbb{K}$ de degré au plus n est noté $\mathbb{K}_n[X]$.

Proposition 12.4.

Soit $P=(p_n)_n$, $Q=(q_n)_n$ deux polynômes à coefficients dans un corps $\mathbb{K}$ et $\lambda\in\mathbb{K}\setminus\{0\}$.
▹ $\deg(P+Q)\leq\max(\deg P,\deg Q)$.
▹ $\deg(\lambda.P)=\deg P$.
▹ $\deg(P.Q)=\deg P+\deg Q$.

Démonstration
▹ Pour tout $n>\max(\deg P,\deg Q)$, $p_n=0$ et $q_n=0$ donc $p_n+q_n=0$.

$$\deg(P+Q)\leq\max(\deg P,\deg Q).$$

▹ Pour tout $n\in\mathbb{N}$, $a_n=0$ si, et seulement si, $\lambda.a_n=0$.
▹ Pour tout $n>\deg P+\deg Q$,

$$\begin{aligned}\sum_{k=0}^{n}p_kq_{n-k}&=\sum_{k=0}^{\deg P}p_kq_{n-k}+\sum_{k=\deg P+1}^{n}p_kq_{n-k}\\&=\sum_{k=0}^{\deg P}p_k.0+\sum_{k=\deg P+1}^{n}0.q_{n-k}=0,\end{aligned}$$

car pour tout $k\leq\deg P$, $n-k>\deg Q$. Par conséquent, $\deg(P.Q)\leq\deg P+\deg Q$.
De plus, si $n=\deg P+\deg Q$,

$$\begin{aligned}\sum_{k=0}^{n}p_kq_{n-k}&=\sum_{k=0}^{\deg P-1}p_k\underbrace{q_{n-k}}_{=0}+p_{\deg P}q_{\deg Q}+\sum_{k=\deg P+1}^{n}\underbrace{p_k}_{=0}q_{n-k}\\&=p_{\deg P}q_{\deg Q}\neq0.\end{aligned}$$

Donc $\deg(P.Q)=\deg P+\deg Q$. ■

Remarque
On peut définir parallèlement au degré la valuation d'un polynôme non nul P comme l'indice, noté $\operatorname{val}P$, du premier coefficient non nul. Par convention, la valuation du polynôme nul est $+\infty$.
Soit $P=(p_n)_n$, $Q=(q_n)_n$ deux polynômes à coefficients dans un corps $\mathbb{K}$ et $\lambda\in\mathbb{K}\setminus\{0\}$.
▹ $\operatorname{val}(P+Q)\geq\min(\operatorname{val}P,\operatorname{val}Q)$.
▹ $\operatorname{val}(\lambda.P)=\operatorname{val}P$.
▹ $\operatorname{val}(P.Q)=\operatorname{val}P+\operatorname{val}Q$.

1.2. Structure d'anneau

On vérifie rapidement la propriété suivante.

Proposition 12.5.

L'ensemble $\mathbb{K}[X]$ muni des lois définies à la section précédente est un anneau.

Pour simplifier les notations, on utilise le symbole de Kronecker défini par

$$\delta_{a,b}=\begin{cases}1 & \text{si } a=b,\\ 0 & \text{sinon.}\end{cases}$$

En fait, on peut aussi montrer que $\mathbb{K}$ se plonge dans $\mathbb{K}[X]$ avec l'injection $\lambda \mapsto (\lambda\delta_{0,n})_n$; les polynômes ainsi obtenus sont appelés polynômes constants. On identifie $\lambda \in \mathbb{K}$ et le polynôme $(\lambda\delta_{0,n})_n$ et on pose $X = (\delta_{1,n})_n$.

Proposition 12.6.

Pour tout $k \in \mathbb{N}$, $X^k = (\delta_{k,n})_n$.

Démonstration

Démontrons cette propriété par récurrence sur k.

▹ Par définition, $X^0 = 1$ et $X^1 = X$ donc la propriété est vraie aux rangs 0 et 1.

▹ Soit $k \in \mathbb{N}$; supposons la propriété vraie au rang k et calculons le $n^{\text{ième}}$ coefficient de $X^{k+1} = X^k.X$

$$\sum_{j=0}^{n} \delta_{1,j}\delta_{k,n-j} = \delta_{k,n-1} = \delta_{k+1,n}.$$

■

Remarque

Désormais, on adopte alors la notation usuelle pour le polynôme $P = (p_n)_{n\in\mathbb{N}}$ de degré p, à savoir $P(X) = \sum_{n=0}^{p} p_n X^n$ ou $P(X) = \sum_{n=0}^{+\infty} p_n X^n$ avec la convention $p_n = 0$ pour $n > p$.

1.3. Divisibilité

Définition 12.7.

Soit P et Q deux polynômes.

▹ Le polynôme P est un multiple de Q s'il existe $A \in \mathbb{K}[X]$ tel que $P = AQ$.

▹ Le polynôme P est un diviseur de Q si Q est un multiple de P.

▹ Les polynômes P et Q sont associés si P et Q se divisent l'un l'autre.

▹ Deux polynômes P et Q sont congrus modulo un autre polynôme non nul A si A est un diviseur de $P - Q$. On note encore $P = Q[A]$.

Remarque

Comme dans l'anneau $\mathbb{Z}$, la relation de congruence modulo un polynôme non nul est compatible avec l'addition et le produit.

Exemple

Soit $n \in \mathbb{N}$. Le polynôme $X - 1$ est un diviseur de $X^n - 1$ puisque la formule de Bernoulli dans l'anneau $\mathbb{K}[X]$ donne

$$X^n - 1 = (X-1)\sum_{k=0}^{n-1} X^k.$$

Exemple

Soit $n \in \mathbb{N}$. Le polynôme $X+1$ est un diviseur de $X^{2n+1}+1$ puisque la formule de Bernoulli avec X et -1 dans l'anneau $\mathbb{K}[X]$ donne

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

En revanche, $X+1$ ne divise pas $X^{2n}+1$. En effet, $X=-1[X+1]$ donc $X^{2n}=1[X+1]$ et par conséquent, $X^{2n}+1=2[X+1]$.

Exemple

Soit $P \in \mathbb{K}[X]$. Le polynôme $P(X)-X$ divise $P(P(X))-X$.
▹ En effet, en notant $P(X)=\sum_{k=0}^{n} a_k X^k$,

$$\begin{aligned} P(P(X))-P(X) &= \sum_{k=0}^{n} a_k (P(X)^k - X^k) \\ &= (P(X)-X)\sum_{k=1}^{n} a_k \sum_{j=0}^{k-1} P(X)^j X^{k-1-j}. \end{aligned}$$

Ainsi, $P(X)-X$ divise $P(P(X))-P(X)$ donc $P(P(X))-X=(P(P(X))-P(X))+(P(X)-X)$.
▹ Réécrivons ce résultat à l'aide de la relation de congruence. On a $P(X)=X[P(X)-X]$ donc

$$\begin{aligned} P(P(X)) &= \sum_{k=0}^{n} a_k P(X)^k = \sum_{k=0}^{n} a_k X^k [P(X)-X] \\ &= P(X)[P(X)-X] = X[P(X)-X]. \end{aligned}$$

Ainsi, $P(X)-X$ divise $P(P(X))-X$.

Proposition 12.8.

Les polynômes inversibles de $\mathbb{K}[X]$ sont les polynômes constants non nuls.

Démonstration
Soit P un polynôme inversible et Q son inverse. En passant au degré, la relation $PQ=1$ entraîne que $\deg P+\deg Q=\deg 1=0$ donc $\deg P=\deg Q=0$.
La réciproque est évidente. ■

1.4. Dérivation

Définition 12.9.

La dérivée (formelle) de $P=\sum_{k=0}^{p} p_k X^k$ est le polynôme $P'=\sum_{k=1}^{p} p_k k X^{k-1}=\sum_{k=0}^{p-1} p_{k+1}(k+1)X^k$.

Par construction, le degré de la dérivée d'un polynôme non constant P est $\deg P-1$.
Regardons quelques règles de manipulation.

Proposition 12.10.

Soit P, $Q \in \mathbb{K}[X]$, $\lambda \in \mathbb{K}$. Alors,

$$(P+\lambda Q)' = P' + \lambda Q', \qquad (PQ)' = P'Q + PQ'.$$

Une démonstration est nécessaire ici car il ne s'agit pas de la dérivation des fonctions déjà étudiée (d'ailleurs, la définition de la dérivation formelle ne dépend pas du corps auquel appartiennent les coefficients). Montrons le résultat pour le produit.

Démonstration

Par linéarité, il suffit de vérifier le résultat pour $P = X^k$ et $Q = X^j$. Alors,

$$(PQ)' = (k+j)X^{k+j-1},$$
$$P'Q + PQ' = kX^{k-1}X^j + X^k jX^{j-1} = (k+j)X^{k+j-1}.$$

■

La règle sur le produit permet avec une récurrence d'obtenir la formule de Leibniz pour la dérivée $n^{\text{ième}}$ formelle.

Proposition 12.11. Formule de Leibniz

Pour tous P, $Q \in \mathbb{K}[X]$,

$$(PQ)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} P^{(k)} Q^{(n-k)}.$$

2. Racines de polynômes

Définition 12.12.

La fonction polynomiale associée au polynôme $\sum_{k=0}^{n} a_k X^k$ est la fonction

$$\left\{ \begin{array}{ccc} \mathbb{K} & \longrightarrow & \mathbb{K} \\ x & \longmapsto & \sum_{k=0}^{n} a_k x^k \end{array} \right.$$

Exemple

La fonction réelle associée au polynôme $X^2 + X$ est la fonction $x \mapsto x^2 + x$. On a utilisé des lettres différentes X et x pour distinguer l'objet algébrique (le polynôme) de l'objet analytique (la fonction).
On remarque que dans le cas réel la dérivation des polynômes correspond à la dérivation des fonctions : la fonction polynomiale associée à la dérivée d'un polynôme P est la dérivée de la fonction polynomiale associée à P.

Remarque
Calculer la valeur de la fonction polynomiale P en $x_0 \in \mathbb{K}$ peut se faire de plusieurs manières :

- de manière naïve, on calcule les puissances de x_0 puis on réalise la combinaison linéaire de ces puissances ;
- de manière plus subtile, on utilise l'écriture
$$P(x_0) = a_0 + x_0(a_1 + x_0(a_2 + \ldots + x_0(a_{n-1} + x_0 a_n)\ldots))$$
et on calcule depuis le terme le plus intérieur aux parenthèses.
Cette méthode efficace (en nombre de produits effectués) s'appelle algorithme de Horner.

2.1. Racines

Définition 12.13.

Une racine d'un polynôme est un zéro de la fonction polynomiale associée.

Cette définition nous ramène dans les complexes (en fait dans les éléments de $\mathbb{K}$) et permet d'obtenir quelques propriétés élémentaires.

Proposition 12.14.

Un complexe z est racine de $P \in \mathbb{R}[X]$ si, et seulement si, $\overline{z}$ est racine de P.

Démonstration
Notons $P(X) = \sum_{k=0}^{n} a_k X^k$. Alors, z est racine de P si, et seulement si,
$$\sum_{k=0}^{n} a_k z^k = 0$$
soit en conjugant et en utilisant que tous les coefficients sont réels,
$$\sum_{k=0}^{n} a_k \overline{z}^k = 0$$
ce qui est bien la définition de $\overline{z}$ est racine de P. ■

On obtient alors la caractérisation arithmétique (c'est-à-dire *via* la divisibilité et sans recours à la fonction polynomiale associée) des racines.

Proposition 12.15.

Un complexe a est une racine de $P \in \mathbb{K}[X]$ si, et seulement si, $X - a$ divise P.

Démonstration
En effet, si $P(X) = \sum_{k=0}^{n} a_k X^k$ et si on note $P(a) = \sum_{k=0}^{n} a_k a^k \in \mathbb{K}$ la valeur de la fonction polynomiale associée, alors, la formule de Bernoulli donne
$$\begin{aligned} P(X) - P(a) &= \sum_{k=1}^{n} a_k (X^k - a^k) \\ &= (X-a) \sum_{k=1}^{n} a_k \sum_{j=0}^{k-1} a^{k-1-j} X^j. \end{aligned}$$

Ainsi, $X-a$ divise P si, et seulement si, $X-a$ divise la constante $P(a)$ (puisque $P=P(a)[X-a]$) et donc si, et seulement si, la valeur $P(a)$ est nulle. ■

Proposition 12.16.

Le degré d'une polynôme majore le nombre de ses racines distinctes.

Démonstration

Soit $a_1,\dots,a_m$ les racines deux à deux distinctes d'un polynôme P. Une récurrence élémentaire permet d'établir que $\prod_{k=1}^{m}(X-a_k)$ divise P ; par conséquent, $\deg P$ est supérieur ou égal à $\deg\prod_{k=1}^{m}(X-a_k)=m$. ■

Remarque

Cette proposition est essentielle : régulièrement on exhibe $n+1$ racines pour un polynôme de degré au plus n pour montrer que ce polynôme est nul.

En particulier, lorsque les coefficients appartiennent au corps $\mathbb{C}$, cela permet de démontrer qu'il existe un unique polynôme associé à une fonction polynomiale donnée.

2.2. Interpolation de Lagrange

Un problème d'interpolation consiste à chercher un polynôme tel que la fonction polynomiale associée prenne certaines valeurs fixées en des arguments fixés.

Proposition 12.17. Lagrange

Soit $a_0,\dots,a_n$ des complexes deux à deux distincts et $b_0,\dots,b_n$ des complexes. Il existe un unique polynôme $P\in\mathbb{C}_n[X]$ tel que $P(a_k)=b_k$ pour tout $k\in[\![0,n]\!]$.

Démonstration

L'unicité procède de la section précédente : s'il existe deux polynômes P et Q de degré au plus n vérifiant ces conditions, alors $P-Q$ admet $n+1$ racines (à savoir $a_0,\dots,a_n$) et est de degré au plus n donc est nul.

▹ Pour l'existence, commençons par un cas simple où tous les b_j sont nuls sauf l'un d'entre eux qui vaut 1.

Le $k^{\text{ième}}$ polynôme de Lagrange associé aux points $a_0,\dots,a_n$ est le polynôme

$$L_k(X)=\prod_{\substack{j=0\\ j\neq k}}^{n}\frac{X-a_j}{a_k-a_j}.$$

On vérifie simplement que $L_k\in\mathbb{C}_n[X]$ puis que $L_k(a_j)=\delta_{j,k}$ pour tout $j\in[\![0,n]\!]$.

▹ Notons $L_0,\dots,L_n$ les polynômes de Lagrange aux points $a_0,\dots,a_n$. Le polynôme $P=\sum_{k=0}^{n}b_kL_k$ vérifie les conditions de l'énoncé ; en effet, pour tout $j\in[\![0,n]\!]$,

$$P(a_j)=\sum_{k=0}^{n}b_kL_k(a_j)=b_jL_j(a_j)=b_j.$$

■

Exemple

Soit a, b et c trois complexes distincts, α, β et $\gamma \in \mathbb{C}$. Le seul polynôme P de degré au plus 2 tel que $P(a)=\alpha$, $P(b)=\beta$ et $P(c)=\gamma$ est donné par

$$\alpha\frac{(X-b)(X-c)}{(a-b)(a-c)}+\beta\frac{(X-a)(X-c)}{(b-a)(b-c)}+\gamma\frac{(X-a)(X-b)}{(c-a)(c-b)}.$$

Les autres polynômes vérifiant ces conditions diffèrent d'un multiple de $(X-a)(X-b)(X-c)$.

Exemple

Soit $P \in \mathbb{C}[X]$ tel que $P(\mathbb{Q}) \subset \mathbb{Q}$. Montrons que P est à coefficients rationnels. En effet, en notant n le degré de P et $L_0, \dots, L_n$ les polynômes de Lagrange aux points 0, 1, …, n, on obtient que $L_k \in \mathbb{Q}[X]$ (d'après l'expression explicite de K_k) donc

$$P=\sum_{k=0}^{n} P(k)L_k \in \mathbb{Q}[X].$$

2.3. Multiplicité d'une racine

Introduisons un concept de multiplicité associé à une racine.

Définition 12.18.

Soit $P \in \mathbb{K}[X]$ un polynôme et a une racine de P.
La multiplicité de a dans P est l'entier $r \in \mathbb{N} \setminus \{0\}$ tel qu'il existe $Q \in \mathbb{K}[X]$ vérifiant $P(X)=(X-a)^r Q(X)$ et a n'est pas racine de Q.
Si la multiplicité d'une racine est 1 (respectivement 2), la racine est simple (respectivement double).

Afin de caractériser les racines multiples, précisons la formule de Taylor pour les polynômes.

Proposition 12.19. Formule de Taylor

Soit $P \in \mathbb{K}_n[X]$ et $a \in \mathbb{K}$. Alors, $P(X)=\sum_{k=0}^{n} \frac{P^{(k)}(a)}{k!}(X-a)^k$.

Démonstration

Soit $P \in \mathbb{K}_n[X]$.

▹ Écrivons $P(X)=\sum_{k=0}^{n} a_k X^k$. En dérivant à k reprises et en spécifiant en 0, on obtient $P^{(k)}(0)=a_k k!$ donc la formule est établie pour $a=0$.

▹ Pour obtenir la formule avec $a \in \mathbb{K}$, il suffit d'appliquer le résultat précédent à $Q(X)=P(X+a)$. Comme

$$Q(X)=\sum_{k=0}^{n} \frac{Q^{(k)}(0)}{k!}X^k,$$

et $Q^{(k)}(0)=P^{(k)}(a)$ pour tout k, on obtient

$$P(X)=\sum_{k=0}^{n} \frac{P^{(k)}(a)}{k!}(X-a)^k.$$

■

En utilisant cette formule, on obtient la caractérisation « analytique » (c'est-à-dire en terme de valeurs de fonctions polynomiales) des racines de multiplicité r.

Proposition 12.20.

Soit $P \in \mathbb{K}[X]$ un polynôme. Une racine a de P est de multiplicité r si, et seulement si, $P^{(k)}(a) = 0$ pour tout $k \leq r-1$ et $P^{(r)}(a) \neq 0$.

Démonstration

▹ D'après la formule de Taylor, a est de multiplicité au moins r dans P si, et seulement si, $(X-a)^r$ divise $\sum_{k=0}^{r-1} \frac{P^{(k)}(a)}{k!}(X-a)^k$. Pour des considérations de degré, cela équivaut à la nullité de ce dernier polynôme donc à $P^{(k)}(a) = 0$ pour tout $k \leq r-1$.

▹ Le complexe a est de multiplicité exactement r dans P s'il est de multiplicité au moins r et pas de multiplicité au moins $r+1$. D'après le point précédent, cela équivaut à $P^{(k)}(a) = 0$ pour tout $k \leq r-1$ et $P^{(r)}(a) \neq 0$. ■

Par convention, une racine de multiplicité 0 est donc un complexe qui n'est pas racine.

Corollaire 12.21.

Soit $P \in \mathbb{R}[X]$ et $a \in \mathbb{C}$. Alors, a est racine de P de multiplicité r si, et seulement si, $\overline{a}$ est racine de P de multiplicité r.
En d'autres termes, les racines conjuguées d'un polynôme réel ont la même multiplicité.

Démonstration

Il suffit de remarquer que, pour tout $k \in \mathbb{N}$, $\overline{P^{(k)}(a)} = P^{(k)}(\overline{a})$ (car les coefficients de $P^{(k)}$ sont réels). Ainsi, pour tout $k \in \mathbb{N}$, $P^{(k)}(a)$ est nul si, et seulement si, $P^{(k)}(\overline{a})$ l'est. On conclut avec la caractérisation précédente de la multiplicité. ■

Une conséquence de la définition de multiplicité est le raffinement de la proposition 12.16.

Proposition 12.22.

Le degré majore le nombre de racines comptées avec leur multiplicité.

2.4. Théorème de D'Alembert-Gauss

Définition 12.23.

Un polynôme est scindé s'il est égal à un produit de polynômes de degré 1.

Remarque

Un polynôme de degré $n \geq 1$ est scindé si, et seulement s'il admet n racines comptées avec leur multiplicité.

Énonçons maintenant le théorème (admis) de D'Alembert-Gauss (aussi quelquefois appelé théorème fondamental de l'algèbre).

Théorème 12.24. D'Alembert-Gauss

Tout polynôme de $\mathbb{C}[X]$ non constant admet une racine. De manière équivalente, tout polynôme complexe est scindé.

Remarque

Les racines complexes d'un polynôme réel sont deux à deux conjuguées puisque, si $P \in \mathbb{R}[X]$, alors, pour tout $z \in \mathbb{C}$, $\overline{P(z)} = P(\overline{z})$; ainsi $P(z) = 0$ si, et seulement si, $P(\overline{z}) = 0$. De plus, des racines conjuguées d'un polynôme réel ont la même multiplicité.

Corollaire 12.25.

Un polynôme $P \in \mathbb{C}[X]$ non constant est surjectif de $\mathbb{C}$ dans $\mathbb{C}$.

Démonstration

Soit $z \in \mathbb{C}$. Le polynôme non constant $P - z$ admet une racine $a \in \mathbb{C}$ d'après le théorème de d'Alembert-Gauss. Ainsi, $P(a) = z$ et z admet un antécédent par P. ■

2.5. Relations coefficients-racines

Définition 12.26.

Soit $x_1, \dots, x_n \in \mathbb{K}$. Les fonctions symétriques élémentaires de $x_1, \dots, x_n$ sont les scalaires définis par

$$\forall k \in [\![1, n]\!], \qquad \sigma_k = \sum_{A \in \mathscr{P}_k([\![1,n]\!])} \prod_{j \in A} x_j$$

où $\mathscr{P}_k([\![1, n]\!])$ désigne l'ensemble des parties de $[\![1, n]\!]$ à k éléments.
En particulier,

$$\sigma_1 = \sum_{i=1}^{n} x_i, \qquad \sigma_2 = \sum_{i<j} x_i x_j, \qquad \sigma_n = \prod_{i=1}^{n} x_i.$$

Exprimons maintenant les relations coefficients-racines.

Proposition 12.27. Relations coefficients-racines

Soit $P = \sum_{k=0}^{n} a_k X^k$ un polynôme scindé dont les racines sont $x_1, \dots, x_n$ (avec leur multiplicité). Alors, pour tout $k \in [\![1, n]\!]$, la $k^{\text{ième}}$ fonction symétrique de $x_1, \dots, x_n$ est $\sigma_k = (-1)^k \frac{a_{n-k}}{a_n}$.

Ce résultat provient du développement de la formule du polynôme comme produit.

On retrouve, avec ce théorème, les conditions sur la somme et le produit des racines d'un polynôme de degré 2.

Exemple

Les racines du polynôme $X^2 - 2\cos(\theta)X + 1$ sont de somme $2\cos(\theta)$ et de produit 1. Rappelons que ces racines sont $e^{i\theta}$ et $e^{-i\theta}$.

Exemple

Le polynôme $\sqrt{2}X^2-(2+\sqrt{2})X+2$ admet 1 comme racine évidente. Le produit des racines est $\frac{2}{\sqrt{2}}$ donc l'autre racine est $\sqrt{2}$.

Remarque

En fait, toute expression polynomiale symétrique de n scalaires peut s'exprimer polynomialement à l'aide des fonctions symétriques élémentaires. Par exemple, avec trois scalaires x, y, z et σ_1, σ_2, σ_3 les fonctions symétriques associées,

$$\begin{aligned} x^2+y^2+z^2 &= \sigma_1^2-2\sigma_2, \\ x^3+y^3+z^3 &= \sigma_1^3+3\sigma_3-3\sigma_1\sigma_2. \end{aligned}$$

En effet,

$$\begin{aligned} \sigma_1^3 &= (x+y+z)^3 = x^3+y^3+z^3+3x^2(y+z)+3y^2(x+z)+3z^2(x+y)+6xyz \\ &= -2(x^3+y^3+z^3)+3x^2(x+y+z)+3y^2(x+y+z)+3z^2(x+y+z)+6xyz \\ &= -2(x^3+y^3+z^3)+3(x^2+y^2+z^2)\sigma_1+6\sigma_3 \\ &= -2(x^3+y^3+z^3)+3(\sigma_1^2-2\sigma_2)\sigma_1+6\sigma_3. \end{aligned}$$

3. Arithmétique des polynômes

3.1. Division euclidienne

Proposition 12.28.

Soit A et B deux polynômes tels que $B \neq 0$. Il existe un unique couple de polynômes (Q,R) tel que $A = QB+R$ et $\deg R < \deg B$.

Démonstration

Séparons l'unicité et l'existence.

▹ Soit deux couples (Q,R) et (Q',R') qui vérifient les conditions de l'énoncé. On a alors : $B(Q-Q') = R'-R$ donc

$$\deg B > \deg(R-R') = \deg(B(Q-Q')).$$

Par conséquent, $Q = Q'$ et $R = R'$.

▹ Montrons le résultat d'existence par récurrence sur le degré de A.

- Si $\deg A < \deg B$, alors $A = 0.B+A$.
- Soit $n \in \mathbb{N}$. Supposons que le résultat est établi pour tout polynôme de degré au plus n et considérons A de degré $n+1$. Notons αX^{n+1} et βX^p les termes dominants de A et B. Si $n+1 < p$, on conclut comme dans la première étape ; sinon, on applique l'hypothèse de récurrence à $A-\frac{\alpha}{\beta}X^{n+1-p}B$ et il existe (Q',R') tels que :
$$A-\frac{\alpha}{\beta}X^{n+1-p}B = Q'B+R',$$
et $\deg R' < p$. On conclut à l'existence souhaitée avec $Q = Q'+\frac{\alpha}{\beta}X^{n+1-p}$ et $R = R'$.

■

Exemple

On obtient avec l'argument d'hérédité une méthode pour effectuer la division euclidienne de deux polynômes. Écrivons cette division sur l'exemple de la division de X^4+7X^3+1 par X^2+1 :

$$\begin{array}{rrrrr|l} X^4 & +7X^3 & & & +1 & X^2+1 \\ \cline{6-6} & 7X^3 & -X^2 & & +1 & X^2+7X-1 \\ & & -X^2 & -7X & +1 & \\ & & & -7X & +2 & \end{array}$$

On remarque la disposition « en potence » comme pour les calculs sur les entiers.

Exemple

Dans de nombreux exercices (notamment en algèbre linéaire), il n'est pas nécessaire d'effectuer la division euclidienne complète. Lorsqu'il suffit de connaître le reste, on peut l'obtenir en utilisant les racines de B et de ses dérivées en cas de racines multiples. Détaillons quelques exemples.

- Calculons le reste R de la division de P par $(X-a)(X-b)$ avec $a \neq b$. En évaluant en a et b, on obtient $R(a)=P(a)$ et $R(b)=P(b)$; or R est de degré au plus 1 donc
$$R(X)=P(a)\frac{X-b}{a-b}+P(b)\frac{X-a}{b-a}.$$
Par exemple, pour calculer le reste de la division de X^n par $X^2-2\cos(\theta)X+1$ avec $\theta \neq 0[\pi]$, on utilise les deux racines $e^{i\theta}$ et $e^{-i\theta}$ et on obtient
$$\frac{\sin(n\theta)}{\sin(\theta)}X-\frac{\sin((n-1)\theta)}{\sin(\theta)}.$$
Ce type de calculs se généralise facilement avec le procédé d'interpolation de Lagrange.
- Calculons le reste R de la division de P par $(X-a)^2$. En évaluant P et la dérivée P' en a, on obtient $P(a)=R(a)$ et $P'(a)=R'(a)$; or R est de degré au plus 1 donc
$$R(X)=P'(a)(X-a)+P(a).$$

3.2. PGCD, PPCM

Définition 12.29.

Soit A et B deux polynômes.
Le PGCD de A et B est l'unique polynôme unitaire $A \wedge B$ tel que
- $A \wedge B$ divise A et B ;
- tout diviseur de A et B divise $A \wedge B$.

Le PPCM de A et B est l'unique polynôme unitaire $A \vee B$ tel que
- A et B divisent $A \vee B$;
- $A \vee B$ divise tout multiple de A et B.

Proposition 12.30. Bézout

Soit A, $B \in \mathbb{K}[X]$. Alors, il existe U, $V \in \mathbb{K}[X]$ tels que $UA+VB=A \wedge B$.

La démonstration repose sur l'algorithme d'Euclide identique à celui pour les entiers.

Plus précisément, on définit $R_0 = A$, $R_1 = B$ et pour tout $k \in \mathbb{N}$, R_{k+2} est le reste de la division euclidienne de R_k par R_{k+1}. $A \wedge B$ est le polynôme unitaire associé au dernier terme non nul de la suite $(R_k)_k$. La relation de Bézout s'obtient en remontant les calculs.

Exemple

Pour calculer le PGCD de $A = X^4 + 7X^3 + 1$ et $B = X^2 + 1$, on obtient successivement les restes suivants

- $R_1 = X^4 + 7X^3 + 1$
- $R_2 = X^2 + 1$
- $R_3 = -7X + 2$
- $R_4 = \frac{53}{49}$
- $R_5 = 0$

Donc, $A \wedge B = 1$.

Exemple

Pour tous $m, n \in \mathbb{N}$, $(X^m - 1) \wedge (X^n - 1) = X^{m \wedge n} - 1$.
En effet, si $a = bq + r$, alors $(X^a - 1) = X^r(X^{bq} - 1) + (X^r - 1) = Q(X)(X^b - 1) + (X^r - 1)$ avec $Q \in \mathbb{R}[X]$ donc

$$(X^a - 1) \wedge (X^b - 1) = (X^b - 1) \wedge (X^r - 1).$$

Donc si $(r_n)_n$ est la suite des restes dans l'algorithme d'Euclide pour les entiers a et b, la suite des (polynômes) restes $(X^{r_n} - 1) \wedge (X^{r_{n+1}} - 1)$ est constante ; en calculant ce résultat en $n = 0$ et pour la dernière valeur de la suite, on obtient

$$\begin{aligned}(X^a - 1) \wedge (X^b - 1) &= (X^{a \wedge b} - 1) \wedge 1 \\ &= X^{a \wedge b} - 1.\end{aligned}$$

3.3. Polynômes premiers entre eux

Définition 12.31.

Deux polynômes sont premiers entre eux si leur PGCD est égal à 1.

Remarque

On fera attention en couplant PGCD et racines.

▹ Si deux polynômes ont une racine commune a, alors ils ne sont pas premiers entre eux car le PGCD est divisible par $X - a$. En particulier, ceci permet simplement de savoir si un polynôme admet une racine multiple : les racines multiples de P sont exactement les racines du PGCD de P et de P' (qui se calcule facilement avec l'algorithme d'Euclide).

▹ Toutefois, la réciproque est fausse : les polynômes réels $X^2 + 1$ et $(X^2 + 1)^2$ ne sont pas premiers entre eux pourtant ils n'ont pas de racine réelle commune.

D'après la définition de PGCD, on a directement le théorème de Bézout pour les polynômes :

Proposition 12.32.

Soit A et B deux polynômes. A et B sont premiers entre eux si, et seulement s'il existe des polynômes U et V tels que $UA + VB = 1$.

Démonstration
Le sens direct correspond à la définition de PGCD. Pour le sens retour, l'égalité implique que $A \wedge B$ est un diviseur unitaire de 1 donc est égal à 1. ■

Comme pour les entiers, on en déduit un certain nombre de corollaires.

Corollaire 12.33.

Soit $A, B_1, \ldots, B_n$ des polynômes tels que, pour tout $k \leq n$, B_k est premier avec A. Alors, les polynômes A et $\prod_{k=1}^{n} B_k$ sont premiers entre eux.

Démonstration
Montrons le résultat avec deux polynômes B_1 et B_2 premiers avec A. D'après le théorème de Bézout, il existe $U_1, U_2, V_1, V_2 \in \mathbb{K}[X]$ tels que

$$U_1 A + V_1 B_1 = 1, \qquad U_2 A + V_2 B_2 = 1.$$

Donc

$$(U_1 + V_1 B_1 U_2)A + V_1 V_2 B_1 B_2 = U_1 A + V_1 B_1 (U_2 A + V_2 B_2) = 1.$$

D'après le théorème de Bézout, A et $B_1 B_2$ sont premiers entre eux. ■

Corollaire 12.34.

Soit A et B des polynômes premiers entre eux. Alors, les polynômes obtenus comme puissances de A et de B sont premiers entre eux.

Exemple

Si $\alpha \neq \beta$, alors $(X-\alpha)^n$ et $(X-\beta)^p$ sont premiers entre eux pour tout $n, p \in \mathbb{N}$ d'après le corollaire précédent et le théorème de Bézout appliqué à la relation

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

Corollaire 12.35.

Soit A, B et C des polynômes tels que A et B sont premiers entre eux, A divise C et B divise C. Alors, AB divise C.

Démonstration
D'après le théorème de Bézout, il existe U et V tels que $UA + VB = 1$ donc $UCA + VBC = C$. Par hypothèse, A divise C donc AB divise CB ; de même, AB divise CA. Par conséquent, AB divise C. ■

Une autre conséquence est le lemme de Gauss.

Corollaire 12.36. Lemme de Gauss

Soit A, B et C des polynômes tels que A et B sont premiers entre eux et A divise BC. Alors, A divise C.

Démonstration
D'après le théorème de Bézout, il existe U et V tels que $UA + VB = 1$ donc $UCA + VBC = C$. Par hypothèse, A divise le premier membre donc divise C. ■

Utilisons enfin le théorème de Bézout pour la résolution d'équations « diophantiennes » linéaires sur les polynômes.

Proposition 12.37.

L'ensemble des polynômes P tels que

$$\begin{cases} P &= R[A] \\ P &= Q[B] \end{cases}$$

avec A et B premiers entre eux, est

$$\{QUA + RVB + CAB,\ C \in \mathbb{K}[X]\},$$

avec U, V des polynômes tels que $AU + BV = 1$.

Démonstration

D'après le théorème de Bézout, il existe U et V tels que $UA + VB = 1$. Le polynôme $P_0 = QUA + RVB$ est solution du système car

$$QUA + RVB = RVB[A] = R(1 - UA)[A] = R[A]$$
$$QUA + RVB = QUA[B] = Q(1 - VB)[B] = Q[B].$$

Alors, P est solution de P_0 si, et seulement si, A et B divisent $P - P_0$, c'est-à-dire si $P - P_0$ est un multiple de AB. ■

Exemple

Cherchons les polynômes P tels que le reste de P dans la division par $X^3 + 1$ est $X^2 + 1$ et le reste de P dans la division par $X^2 + 1$ est $X + 1$.
Le PGCD de $X^3 + 1$ et $X^2 + 1$ est 1 donc l'algorithme d'Euclide donne la relation de Bézout

$$\frac{1}{2}(X+1)(X^3+1) + \frac{1}{2}(-X^2 - X + 1)(X^2+1) = 1.$$

Ainsi les solutions sont de la forme

$$(X+1)\frac{1}{2}(X+1)(X^3+1) + (X^2+1)\frac{1}{2}(-X^2 - X + 1)(X^2+1) + C(X)(X^2+1)(X^3+1)$$

avec $C \in \mathbb{K}[X]$.

3.4. Polynômes irréductibles

Passons aux polynômes irréductibles, l'analogue des nombres premiers dans l'arithmétique des entiers.

Définition 12.38.

Un polynôme P non-constant est irréductible si ses seuls diviseurs sont associés à 1 ou à P.

Proposition 12.39.

- Les polynômes de degré 1 sont toujours irréductibles.
- Les polynômes de degré 2 à discriminants négatifs sont irréductibles dans $\mathbb{R}[X]$.

Démonstration

- Si un polynôme P de degré 1 s'écrit AB, alors $\deg A + \deg B = 1$ donc $\deg A = 0$ ou $\deg B = 0$.
- Si un polynôme P de degré 2 s'écrit AB, alors $\deg A + \deg B = 2$ donc $\deg A = 0$ ou $\deg B = 0$ ou

$\deg A = \deg B = 1$. Or, ce dernier cas entraîne que A admet une racine dans $\mathbb{R}$ ce qui est impossible si le discriminant est strictement négatif. ■

Exemple

Il faut faire attention au corps dans lequel sont les coefficients car l'irréductibilité dépend de celui-ci. Par exemple,
▹ le polynôme X^3-2 n'est pas irréductible dans $\mathbb{C}[X]$ ni dans $\mathbb{R}[X]$ (car il admet le diviseur $X-\sqrt[3]{2}$) mais est irréductible dans $\mathbb{Q}[X]$: s'il était réductible, il serait produit d'un polynôme de degré 1 et d'un polynôme de degré 2 donc admettrait une racine rationnelle ce que l'on peut contredire avec les arguments arithmétiques sur les entiers.
▹ le polynôme X^2+1 est irréductible dans $\mathbb{R}[X]$ mais ne l'est pas dans $\mathbb{C}[X]$ puisqu'il s'écrit

$$X^2+1=(X+i)(X-i).$$

Proposition 12.40.

Les polynômes irréductibles de $\mathbb{C}[X]$ sont les polynômes de degré 1.

Démonstration
D'après le théorème de D'Alembert-Gauss, les polynômes de $\mathbb{C}[X]$ sont scindés donc seuls les polynômes de degré 1 peuvent être irréductibles. ■

Proposition 12.41.

Les polynômes irréductibles de $\mathbb{R}[X]$ sont les polynômes de degré 1 et les polynômes de degré 2 à discriminants strictement négatifs.

Démonstration
Soit P un polynôme de $\mathbb{R}[X]$ irréductible. S'il admet une racine réelle, il est nécessairement de degré 1. Sinon, d'après le théorème de D'Alembert-Gauss, P admet une racine $a \in \mathbb{C} \setminus \mathbb{R}$; comme ce polynôme est réel, $\overline{a}$ est aussi racine de P. Ainsi, $(X-a)(X-\overline{a})$ divise P : donc P est de degré 2 à discriminant strictement négatif. ■

Conseils méthodologiques

La preuve dans le cas réel indique comment retrouver les facteurs irréductibles de degré 2 en regroupant les racines complexes conjuguées. Ainsi, pour obtenir la décomposition dans $\mathbb{R}[X]$, il suffit de bien comprendre la décomposition de $\mathbb{C}[X]$.

Exemple

Factorisons X^4+1 dans $\mathbb{R}[X]$. Les racines complexes de ce polynôme sont les racines quatrièmes de -1. Par conséquent,

$$\begin{aligned} X^4+1 &= \prod_{k=0}^{3}(X-e^{\frac{(2k+1)i\pi}{4}}) \\ &= \left[(X-e^{\frac{i\pi}{4}})(X-e^{-\frac{i\pi}{4}})\right]\left[((X-e^{\frac{3i\pi}{4}})(X-e^{-\frac{3i\pi}{4}})\right] \\ &= (X^2+\sqrt{2}X+1)(X^2-\sqrt{2}X+1). \end{aligned}$$

Exemple

Factorisons $X^{2n}-1$ sur $\mathbb{R}$ ou sur $\mathbb{C}$ en produits de polynômes irréductibles.

$$\begin{aligned} X^{2n}-1 &= \prod_{k=0}^{2n-1}(X-e^{i\frac{k\pi}{n}}) \\ &= (X-1)(X+1)\prod_{k=1}^{n-1}\left(X^2-2\cos\left(\frac{k\pi}{n}\right)X+1\right). \end{aligned}$$

Passons alors au problème de la factorisation d'un polynôme comme produit de polynômes irréductibles.

Proposition 12.42.

Soit P un polynôme unitaire irréductible. Pour tout polynôme A, le PGCD $A\wedge P$ est égal à 1 ou à P.

Démonstration

Le PGCD $A\wedge P$ est un diviseur de P donc associé à 1 ou à P ; comme il est unitaire, il est égal à 1 ou à P. ■

Proposition 12.43. Euclide

Soit P un polynôme unitaire irréductible. Si P divise un produit de polynômes, alors P divise l'un des facteurs de ce produit.

Démonstration

Si P ne divise aucun facteur, alors le PGCD avec chaque facteur est égal à 1 (d'après la proposition précédente) donc le PGCD avec le produit est égal à 1, ce qui contredit l'hypothèse. ■

Passons au théorème fondamental de l'arithmétique des polynômes.

Proposition 12.44.

Soit A un polynôme non constant. Il existe un entier $m\in\mathbb{N}\setminus\{0\}$, m polynômes irréductibles unitaires $P_1,\dots,P_m$ deux à deux distincts, m entiers non nuls $\alpha_1,\dots,\alpha_m$ et un scalaire λ tels que

$$A=\lambda\prod_{k=1}^{m}P_k^{\alpha_k}.$$

Cette décomposition est unique à l'ordre des facteurs près.

On démontre ce théorème (autant l'existence que l'unicité) par récurrence sur le degré de A.

Proposition 12.45.

Soit A de décomposition en facteurs irréductibles $A=\lambda\prod_{k=1}^{m}P_k^{\alpha_k}$. Les diviseurs de A sont de la forme $\mu\prod_{k=1}^{m}P_k^{\beta_k}$ avec $\mu\in\mathbb{K}$ et pour tout $k\leq m$, $\beta_k\leq\alpha_k$.

Proposition 12.46.

▹ Le PGCD des polynômes de décomposition en facteurs irréductibles $\lambda \prod_{k=1}^{m} P_k^{\alpha_k}$ et $\mu \prod_{k=1}^{m} P_k^{\beta_k}$ est donné par la décomposition $\prod_{k=1}^{m} P_k^{\min(\alpha_k,\beta_k)}$.
▹ Le PPCM correspondant est obtenu par la formule $\prod_{k=1}^{m} P_k^{\max(\alpha_k,\beta_k)}$.

FICHE

SYNTHÈSE

L'étude des polynômes rassemble des aspects arithmétiques (divisibilité, PGCD, polynômes irréductibles...), des aspects algébriques et des aspects analytiques (l'étude des fonctions polynomiales et des racines).

Arithmétique

▹ Soit A et B deux polynômes tels que $B \neq 0$. Il existe un unique couple de polynômes (Q, R) tel que $A = QB + R$ et $\deg R < \deg B$.
▹ Le PGCD de A et B est l'unique polynôme unitaire $A \wedge B$ tel que
- $A \wedge B$ divise A et B ;
- tout diviseur de A et B divise $A \wedge B$.

Les résultats arithmétiques (lemme de Gauss, d'Euclide...) de $\mathbb{Z}$ se généralisent assez simplement à $\mathbb{K}[X]$.
Les racines sont importantes pour comprendre la notion de polynôme réel ou complexe.

Racines

Soit $P \in \mathbb{K}[X]$ un polynôme et $a \in \mathbb{K}$.
▹ L'élément a est une racine de P si $X - a$ divise P.
Ceci équivaut à dire que a est un zéro de la fonction polynomiale associée, c'est-à-dire si $P(a) = 0$.
▹ La multiplicité de a comme racine de P est l'entier $r \in \mathbb{N} \setminus \{0\}$ tel qu'il existe $Q \in \mathbb{K}[X]$ vérifiant $P(X) = (X - a)^r Q(X)$ et a n'est pas racine de Q.
Ceci équivaut à la condition $P^{(k)}(a) = 0$ pour tout $k \leq r - 1$ et $P^{(r)}(a) \neq 0$.

Relations entre coefficients et racines

Soit $P = \sum_{k=0}^{n} a_k X^k$ un polynôme scindé dont les racines sont $x_1, \ldots, x_n$ (avec leur multiplicité).
Alors, pour tout $k \in [\![1, n]\!]$, la $k^{\text{ième}}$ fonction symétrique de $x_1, \ldots, x_n$ est

$$\sigma_k = (-1)^k \frac{a_{n-k}}{a_n}.$$

Interpolation de Lagrange

Soit $a_0, \ldots, a_n$ des complexes deux à deux distincts.
▹ Le $k^{\text{ième}}$ polynôme de Lagrange associé aux points $a_0, \ldots, a_n$ est le polynôme

$$L_k(X) = \prod_{\substack{j=0 \\ j \neq k}}^{n} \frac{X - a_j}{a_k - a_j}.$$

C'est l'unique polynôme de degré inférieur ou égal à n qui vaut 0 en a_j pour $j \neq k$ et 1 en a_k.
▹ Soit $b_0, \ldots, b_n$ des complexes. Il existe un unique polynôme $P \in \mathbb{C}_n[X]$ tel que $P(a_k) = b_k$ pour tout $k \in [\![0, n]\!]$. Ce polynôme est donné par

$$\sum_{k=0}^{n} b_k L_k.$$

L'équivalent de la notion de nombres premiers dans $\mathbb{Z}$ est la notion de polynômes irréductibles dans $\mathbb{K}[X]$.

Polynômes irréductibles

▹ Un polynôme P non-constant est irréductible si ses seuls diviseurs sont associés à 1 ou à P.
▹ Tout polynôme de $\mathbb{C}[X]$ se décompose de manière unique sous la forme $\lambda \prod_{k=1}^{m} P_k^{\alpha_k}$ avec $P_1, \ldots, P_m$ des polynômes irréductibles unitaires deux à deux distincts, m entiers non nuls $\alpha_1, \ldots, \alpha_m$ et un scalaire λ.
▹ Les polynômes irréductibles de $\mathbb{C}[X]$ sont les polynômes de degré 1.
▹ Les polynômes irréductibles de $\mathbb{R}[X]$ sont les polynômes de degré 1 et les polynômes de degré 2 à discriminant strictement négatif.

EXERCICES

Vrai ou faux ?

		Vrai	Faux
a)	Pour tous $P, Q \in \mathbb{R}[X]$, $\deg(P.Q) = \deg P + \deg Q$.	☐	☐
b)	Les inversibles de l'anneau $\mathbb{R}[X]$ sont exactement les polynômes constants non nuls.	☐	☐
c)	$2X$ est un diviseur de X dans $\mathbb{R}[X]$.	☐	☐
d)	Le reste dans la division entre deux polynômes à coefficients entiers est à coefficients entiers.	☐	☐
e)	Deux polynômes premiers entre eux n'ont pas de racine commune.	☐	☐
f)	Deux polynômes complexes premiers entre eux n'ont pas de racine commune.	☐	☐
g)	Deux polynômes sans racine commune sont premiers entre eux.	☐	☐
h)	Deux polynômes complexes sans racine commune sont premiers entre eux.	☐	☐
i)	Si un polynôme P divise A et un polynôme Q divise B, alors $P \wedge Q$ divise $A \wedge B$.	☐	☐
j)	Si P est premier avec Q et P divise QR, alors P divise R.	☐	☐
k)	$X^3 + X + 1$ est irréductible dans $\mathbb{R}[X]$.	☐	☐
l)	$X^4 + 1$ est irréductible dans $\mathbb{R}[X]$.	☐	☐
m)	Si z est une racine de multiplicité n de P, alors $P^{(n)}(z) = 0$.	☐	☐
n)	Le polynôme $P/(P \wedge P')$ est à racines simples.	☐	☐
o)	La somme des racines complexes (comptées avec leur multiplicité) de $X^{1515} + 18X^{1514} + 42X + 7$ est 18.	☐	☐
p)	Le produit des racines complexes (comptées avec leur multiplicité) de $X^{1515} + 18X^{1514} + 42X + 7$ est 7.	☐	☐

Exercices d'application du chapitre 12

●○○ **Exercice 1**

Déterminer les réels α tels que $x \leq x^2 + \alpha$ pour tout $x \in \mathbb{R}$.

●○○ **Exercice 2**

Soit P et Q deux polynômes réels distincts de degré n. Montrer que $P^3 - Q^3$ est de degré supérieur ou égal à $2n$.

●●○ **Exercice 3**

Trouver le polynôme $P \in \mathbb{R}[X]$ de degré minimal tel que

- le reste de la division euclidienne de P par $X^2 + X + 1$ est $X - 1$;
- le reste de la division euclidienne de P par $(X-1)^2$ est $-X + 2$.

●○○ **Exercice 4**

Soit $n \in \mathbb{N} \setminus \{0\}$. Trouver les complexes a, $b \in \mathbb{C}$ tels que $(X-1)^2$ divise $aX^{n+1} + bX^n + 1$.

●○○ **Exercice 5**

Soit $n \in \mathbb{N}^*$ et $(\theta_1, \dots, \theta_n) \in \mathbb{R}^n$. Posons $P(X) = \prod_{k=1}^{n} (\cos\theta_k + X \sin\theta_k)$. Déterminer le reste de la division euclidienne de P par $X^2 + 1$.

●○○ **Exercice 6**

Montrer que, pour tout $(m, n, p, q) \in \mathbb{N}^4$, $X^3 + X^2 + X + 1$ divise $X^{4m+3} + X^{4n+2} + X^{4p+1} + X^{4q}$.

●○○ **Exercice 7**

Soit $\alpha \in \mathbb{R} \setminus \pi\mathbb{Z}$. Factoriser le polynôme $(X+1)^n - e^{2i\alpha}(X-1)^n$ dans $\mathbb{C}$.

●●○ **Exercice 8**

Soit $n \in \mathbb{N} \setminus \{0\}$.

1. Montrer qu'il existe un unique couple (P, Q) de polynômes de degrés strictement inférieurs à n tels que $(1-X)^n P(X) + X^n Q(X) = 1$.
2. Montrer que $Q(X) = P(1-X)$.
3. Montrer qu'il existe $\lambda \in \mathbb{R}$ tel que $(1-X)P'(X) - nP(X) = \lambda X^{n-1}$.
4. En déduire les coefficients de P.

●●○ **Exercice 9**

1. Montrer que, si A et B sont deux polynômes réels qui sont chacun somme de deux carrés de polynômes, alors il en est de même pour AB.
2. Soit $A \in \mathbb{R}[X]$. Déduire de la question précédente qu'il existe deux polynômes à coefficients réels P et Q tels que $A = P^2 + Q^2$ si, et seulement si, pour tout $x \in \mathbb{R}$, $A(x) \geq 0$.

●○○ **Exercice 10**

Déterminer les polynômes $P \in \mathbb{R}[X]$ tels que $P(k) = k^{1789}$ pour tout $k \in \mathbb{N}$.

●○○ **Exercice 11**

Soit $P \in \mathbb{R}[X]$ tel que, pour tout $1 \leq k \leq n$, $P(k) = \frac{1}{k}$. Montrer que $\deg P \geq n - 1$.

●●○ **Exercice 12**

Trouver les réels a tels que le polynôme $2X^3 - X^2 - 7X + a$ admette deux racines de somme 1.

●○○ **Exercice 13**

Soit $P(X) = X^3 + X + 1$ de racines complexes z_1, z_2 et z_3. Calculer $z_1^4 + z_2^4 + z_3^4$.

EXERCICES

●●○ **Exercice 14**

Montrer que l'ensemble des polynômes unitaires, de degré $n > 0$, à coefficients entiers et à racines complexes dans $\mathbb{U}$ est fini.

●●● **Exercice 15**

Soit $P \in \mathbb{C}[X]$ un polynôme non-constant dont toutes les racines appartiennent au demi-plan d'équation $\mathfrak{Im}(z) > 0$. Soit A (respectivement B) le polynôme obtenu à partir de P en remplaçant les coefficients par leurs parties réelles (respectivement imaginaires). Montrer que les polynômes A et B sont scindés sur $\mathbb{R}$.

●●○ **Exercice 16**

Montrer qu'un polynôme $P \in \mathbb{R}[X]$ non-constant unitaire est scindé sur $\mathbb{R}$ si, et seulement si, pour tout $z \in \mathbb{C}$, $|P(z)| \geq |\mathfrak{Im}(z)|^{\deg P}$.

●○○ **Exercice 17**

Soit $(a_0, a_1, \ldots, a_{n-1}) \in \mathbb{C}^n$ et $P(X) = X^n + \sum_{k=0}^{n-1} a_k X^k$. Montrer que les racines complexes de P ont un module majoré par $\max\left(1, \sum_{k=0}^{n-1} |a_k|\right)$.

●●● **Exercice 18**

Soit $P \in \mathbb{Q}[X]$ de degré 5 qui admet une racine multiple dans $\mathbb{C}$. Montrer que P admet une racine dans $\mathbb{Q}$.

Exercices d'approfondissement du chapitre 12

●●● **Exercice A**

Soit $P \in \mathbb{Z}[X]$ de degré $n > 1$. Montrer que $P(\mathbb{Z})$ contient au plus n entiers consécutifs.

●●● **Exercice B**

Soit $P, Q \in \mathbb{Z}[X]$ premiers entre eux; posons, pour tout $n \in \mathbb{N}$, $u_n = P(n) \wedge Q(n)$. Montrer que la suite $(u_n)_n$ est périodique.

●●● **Exercice C**

Soit $E = \{3, 5, 7, 11, 13, 17\}$ et 50 réels $(a_k)_{k\in[0,49]}$ tels que

$$\forall p \in E,\ \forall r \in [\![1, p-1]\!], \qquad \sum_{k=r[p]} a_k = \sum_{k=0[p]} a_k.$$

Montrer que les réels $(a_k)_{k\in[0,49]}$ sont tous nuls. On pourra chercher des racines du polynôme $\sum_{k=0}^{49} a_k X^k$.

●●○ **Exercice D**

Soit $P \in \mathbb{R}[X]$ scindé à racines simples dans $\mathbb{R}$ et $\alpha \in \mathbb{R}^*$. Montrer que $P^2 + \alpha^2$ est scindé à racines simples dans $\mathbb{C}$.

●●○ **Exercice E**

Un polynôme réel est stable si toutes ses racines complexes sont de partie réelle strictement négative.

1. Montrer que si un polynôme est unitaire et stable alors ses coefficients sont strictement positifs.
2. Soit $P = \prod_{k=1}^{n} (X - \lambda_k) \in \mathbb{R}[X]$ et $Q = \prod_{1\leq i<j\leq n} (X - \lambda_i - \lambda_j)$. Montrer que P est stable si, et seulement si, les coefficients de P et Q sont strictement positifs.

●●● **Exercice F**

Soit z_1, z_2 deux complexes distincts et $P, Q \in \mathbb{C}[X]$ tels que $P^{-1}(\{z_1\}) = Q^{-1}(\{z_1\})$ et $P^{-1}(\{z_2\}) = Q^{-1}(\{z_2\})$. Montrer que $P = Q$.

Problèmes du chapitre 12

●○○ **Problème 1 Problème de Waring**

Pour tout $n \geq 2$, notons $k(n)$ le plus petit entier $k \geq 1$ tels que

$$\forall P \in \mathbb{C}[X],\ \exists (Q_1, \ldots, Q_k) \in \mathbb{C}[X]^k, \qquad P(X) = Q_1(X)^n + \ldots + Q_k(X)^n.$$

1. Justifier que, pour tout $n \geq 2$, $k(n)$ est également le plus petit entier $k \geq 1$ tels que

$$\exists (Q_1, \ldots, Q_k) \in \mathbb{C}[X]^k, \qquad X = Q_1(X)^n + \ldots + Q_k(X)^n.$$

2. Calculer $k(2)$.
3. Soit $n > 2$, $Q_1, Q_2 \in \mathbb{C}[X]$.
   a. Déterminer $A \subset \mathbb{C}$ tel que

$$Q_1(X)^n + Q_2(X)^n = \prod_{\eta \in A} (Q_1(X) - \eta Q_2(X)).$$

   b. Montrer que, pour $n > 2$, $k(n) \geq 3$.
4. a. Soit l'application (linéaire) $\Delta : P \mapsto P(X+1) - P(X)$.
   Montrer que, pour tout $1 \leq k \leq n-1$,

$$\Delta^k(X^n) = \sum_{j=0}^{k} \binom{k}{j} (-1)^{k+j} (X+j)^n.$$

   b. Montrer que $k(n) \leq n$ pour tout n.
   Indication : on pourra utiliser $\Delta^{n-1}(X^n)$.

●○○ **Problème 2 Théorème de Block-Thielmann**

L'objectif de cet exercice est de caractériser les familles de polynômes $(P_n)_{n \in \mathbb{N}^*}$ qui satisfont :

$$\begin{cases} \forall n \in \mathbb{N}^* & \deg(P_n) = n \\ \forall (n,m) \in (\mathbb{N}^*)^2 & P_n \circ P_m = P_m \circ P_n \end{cases} \tag{12.1}$$

où $\circ$ désigne la composition (des polynômes).

Partie A – Préliminaires

1. Montrer que la famille $(X^n)_{n \in \mathbb{N}^*}$ satisfait (12.1).
2. a. Montrer qu'il existe une unique famille $(T_n)_{n \in \mathbb{N}^*}$ telle que

$$\begin{cases} \forall n \in \mathbb{N}^* & \deg(T_n) = n \\ \forall (n,x) \in \mathbb{N}^* \times \mathbb{R} & T_n(\cos(x)) = \cos(nx) \end{cases}$$

   b. Montrer que $(T_n)_{n \in \mathbb{N}^*}$ satisfait les conditions (12.1).
3. Montrer que l'ensemble des polynômes de degré 1 est un groupe pour la loi $\circ$.
   Dans la suite, P^{-1} désigne l'inverse pour la loi $\circ$ d'un polynôme P de degré 1.

PROBLÈMES

Partie B – Caractérisation

Soit $(P_n)_{n\in\mathbb{N}^*}$ une famille de polynômes vérifiant les conditions (12.1).

1. Établir l'existence du polynôme A de degré 1 et d'un réel a tels que
$$A\circ P_2\circ A^{-1} = X^2 + a.$$
2. Montrer que la famille $(Q_n = A\circ P_n\circ A^{-1})_{n\in\mathbb{N}^*}$ satisfait les conditions (12.1) et que, pour tout $n\in\mathbb{N}^*$, Q_n est unitaire.
3. En calculant $Q_2\circ Q_3$, montrer que $a\in\{0,-2\}$.
4. Soit $n\in\mathbb{N}^*$. Montrer, par l'absurde, qu'il n'y a qu'un seul polynôme unitaire de degré n qui commute avec Q_2.
5. a. Supposons $a=0$. Calculer Q_n pour tout entier $n\in\mathbb{N}^*$.
   b. Supposons $a=-2$. Trouver B un polynôme de degré 1 tel que
$$B\circ T_2\circ B^{-1} = Q_2.$$
   En déduire la suite $(Q_n)_{n\in\mathbb{N}^*}$.
6. Conclure.

Problème 3 Polynômes irréductibles à coefficients entiers

Définition 12.47.

Un polynôme P à coefficients entiers est irréductible sur $\mathbb{Z}$ si pour tout couple (A,B) de polynômes à coefficients entiers tels que $P=AB$, soit A soit B est égal à 1 ou -1.

Partie A – Critère de Perron

1. Soit $P = X^n + \sum_{k=0}^{n-1} a_k X^k \in \mathbb{C}[X]$ tel que
$$|a_{n-1}| > 1 + \sum_{k=0}^{n-2} |a_k|.$$
   a. Montrer que P n'admet pas de racine de module 1.
   b. Supposons que P admette une racine z_1 de module strictement supérieur à 1 et notons
$$P = (X-z_1)\sum_{k=0}^{n-1} b_k X^k.$$
      i. Calculer b_{n-1}, b_0 puis montrer que, pour tout $k\in[\![1,n-1]\!]$,
$$a_k = b_{k-1} - z_1 b_k.$$
      ii. En déduire que
$$|b_{n-2} - z_1| \geq 1 + |b_{n-2}| + (|z_1|-1)\sum_{k=0}^{n-2}|b_k|,$$
      puis que
$$1 > \sum_{k=0}^{n-2}|b_k|.$$
      iii. En déduire que $\sum_{k=0}^{n-1} b_k X^k$ n'a pas de racine de module supérieur à 1.

c. Conclure que P admet au plus une racine de module supérieur ou égal à 1 et que celle-ci est simple.

2. Soit $P = X^n + \sum_{k=0}^{n-1} a_k X^k$ à coefficients entiers tel que $a_0 \neq 0$ et

$$|a_{n-1}| > 1 + \sum_{k=0}^{n-2} |a_k|.$$

Supposons que $P = AB$ avec A et B deux polynômes non constants à coefficients entiers.

a. En étudiant les coefficients constants de A et B, montrer que ces polynômes admettent une racine de module supérieur ou égal à 1.

b. Conclure.

Partie B – Critère de Polya

1. Soit P un polynôme à coefficients entiers de degré n et des entiers $\alpha_0 < \alpha_1 < \ldots < \alpha_n$.

a. Montrer que le coefficient dominant de P est

$$\sum_{k=0}^{n} P(\alpha_k) \prod_{\substack{l=0 \\ l \neq k}}^{n} \frac{1}{\alpha_k - \alpha_l}.$$

b. En déduire qu'il existe $k \in [\![0, n]\!]$ tel que $|P(\alpha_k)| \geq n! 2^{-n}$.

2. Soit P un polynôme à coefficients entiers de degré n, $\alpha_1, \alpha_2, \ldots, \alpha_n$ des entiers deux à deux distincts non racines de P tels que

$$\forall k \in [\![1, n]\!], \qquad |P(\alpha_k)| < M! 2^{-M},$$

où $M = \left\lfloor \frac{n+1}{2} \right\rfloor$.

Supposons que $P = AB$ avec A et B non constants à coefficients entiers.

a. Montrer que l'on peut supposer que $\deg A \geq M$.

b. Vérifier que $(\deg A)! 2^{-\deg A} \geq M! 2^{-M}$.

c. Conclure.

Partie C – Exemples classiques

Soit $\alpha_1 < \ldots < \alpha_n$ des entiers.

1. Montrer que le polynôme $\prod_{k=1}^{n} (X - \alpha_k) - 1$ est irréductible sur $\mathbb{Z}$.
2. Montrer que le polynôme $\prod_{k=1}^{n} (X - \alpha_k) + 1$ est irréductible sur $\mathbb{Z}$ pour n impair.
3. Montrer que le polynôme $\prod_{k=1}^{n} (X - \alpha_k)^2 + 1$ est irréductible sur $\mathbb{Z}$.

PROBLÈMES

CORRIGÉS

Corrigés des Vrai/Faux

a) Vrai.

b) Vrai.

c) Vrai, $X = \frac{1}{2}.2X$.

d) Faux, la division de X par $2X+1$ est $X = \frac{1}{2}(2X+1) - \frac{1}{2}$.

e) Vrai.

f) Vrai.

g) Faux, avec par exemple X^2+1 et $(X^2+1)^2$.

h) Vrai car les polynômes complexes sont scindés d'après le théorème de D'Alembert-Gauss.

i) Vrai.

j) Vrai.

k) Faux, les polynômes irréductibles sont de degré 1 ou 2.

l) Faux, les polynômes irréductibles sont de degré 1 ou 2.

m) Faux, la condition est $P(z) = P'(z) = \ldots = P^{(n-1)}(z) = 0$.

n) Vrai.

o) Faux, c'est -18.

p) Faux, c'est -7.

Corrigés des exercices d'application du chapitre 12

Exercice 1

Il suffit de déterminer les $\alpha \in \mathbb{R}$ tel que le polynôme $X^2 - X + \alpha$ soit positif, c'est-à-dire admette au plus une racine (car le coefficient dominant est positif). Il faut donc et il suffit que le discriminant soit négatif : $1 - 4\alpha \leq 0$. Les réels recherchés forment donc l'intervalle $\left[\frac{1}{4}, +\infty\right[$.

Exercice 2

Écrivons $P^3 - Q^3 = (P-Q)(P-jQ)(P-j^2Q)$. En regardant les coefficients de X^n dans les trois facteurs, on trouve que deux au moins sont non nuls. Par conséquent, aucun des termes n'est nul et deux sont de degré n, le produit est donc de degré au moins $2n$.

Exercice 3

Résolvons complètement le système

$$\begin{cases} P(X) & = X-1 & [X^2+X+1] \\ P(X) & = -X+2 & [(X-1)^2] \end{cases}$$

Pour cela, procédons en plusieurs étapes.

▹ Calculons une relation de Bézout entre les polynômes $A(X) = X^2+X+1$ et $B(X) = (X-1)^2$. Les étapes de l'algorithme d'Euclide sont

$$A(X) = 1.B(X) + 3X$$
$$B(X) = \frac{1}{3}(X-2).3X + 1$$

Ainsi, les polynômes $A(X)$ et $B(X)$ sont premiers entre eux et on obtient la relation de Bézout en remontant les calculs :

$$-\frac{1}{3}(X-2)A(X) + \frac{1}{3}(X+1)B(X) = 1.$$

▹ Une solution du système initial est donc

$$\begin{aligned} P_0(X) &= -\frac{1}{3}(X-2)A(X)(-X+2) + \frac{1}{3}(X+1)B(X)(X-1) \\ &= \frac{1}{3}(2X^4 - 5X^3 + X^2 + 2X + 3). \end{aligned}$$

En effet, avec la relation de Bézout,

$$\begin{aligned} P_0(X) &= \frac{1}{3}(X+1)B(X)(X-1)\,[A(X)] \\ &= (1 + \frac{1}{3}(X-2)A(X))(X-1)\,[A(X)] \\ &= (X-1)\,[A(X)]. \end{aligned}$$

L'autre égalité se montre de la même manière.

▹ Résolvons désormais le système. Un polynôme P est une solution du système si, et seulement si,

$$\begin{cases} P(X) & = P_0(X) & [X^2+X+1] \\ P(X) & = P_0(X) & [(X-1)^2] \end{cases}$$

c'est-à-dire si $P - P_0$ est divisible par X^2+X+1 et par $(X-1)^2$ donc par $(X^2+X+1)(X-1)^2$ (puisque ces polynômes sont premiers entre eux). Ainsi, l'ensemble des solutions du système initial est

$$\{P_0(X) + R(X)(X^2+X+1)(X-1)^2,\ R \in \mathbb{R}[X]\}.$$

▹ Revenons enfin à la question de l'énoncé. Le polynôme de degré minimal vérifiant le système correspond à $R(X) = -\frac{2}{3}$ (seule possibilité pour avoir un polynôme de degré strictement inférieur à 4) et on obtient

$$P(X) = -X^3 + \frac{1}{3}X^2 + \frac{4}{3}X + \frac{1}{3}.$$

Exercice 4

On cherche ici $a,\ b \in \mathbb{C}$ tels que 1 soit racine double de $P(X) = aX^{n+1} + bX^n + 1$, c'est-à-dire tels que $P(1) = P'(1) = 0$. On obtient le système

$$\begin{cases} a & + & b & = -1 \\ (n+1)a & + & nb & = 0 \end{cases}$$

dont l'unique solution est donnée par $a = n$ et $b = -1-n$.

Exercice 5

Notons la division euclidienne $P = Q.(X^2+1) + aX + b$ avec $Q \in \mathbb{R}[X]$, a, $b \in \mathbb{R}$. En évaluant en i, on obtient $b + ai = P(i) = \exp\left(i\sum_{k=1}^{n}\theta_k\right)$. Le reste recherché est donc

$$\cos\left(\sum_{k=1}^{n}\theta_k\right)X + \sin\left(\sum_{k=1}^{n}\theta_k\right).$$

Exercice 6

On remarque tout d'abord que $X^4 - 1 = (X-1)(X^3+X^2+X+1)$ donc $X^4 = 1[X^3+X^2+X+1]$. Avec les règles de manipulations des congruences, on obtient, pour $(m,n,p,q) \in \mathbb{N}^4$,

$$\begin{aligned} X^{4m+3} + X^{4n+2} + X^{4p+1} + X^{4q} &= X^3 + X^2 + X + 1[X^3+X^2+X+1] \\ &= 0[X^3+X^2+X+1] \end{aligned}$$

ce qui signifie que X^3+X^2+X+1 divise tous les polynômes de la forme $X^{4m+3} + X^{4n+2} + X^{4p+1} + X^{4q}$.

Exercice 7

Remarquons tout d'abord que 1 n'est pas racine de ce polynôme. Ensuite, $z \in \mathbb{C} \setminus \{1\}$ est racine de $(X+1)^n - e^{2i\alpha}(X-1)^n$ si, et seulement si,

$$\left(\frac{z+1}{z-1}\right)^n = e^{2i\alpha}.$$

Cette condition équivaut à l'existence d'un $k \in [\![0, n-1]\!]$ tel que $\frac{z+1}{z-1} = e^{\frac{2i(\alpha+k\pi)}{n}}$ soit tel que

$$z = \frac{e^{\frac{2i(\alpha+k\pi)}{n}}+1}{e^{\frac{2i(\alpha+k\pi)}{n}}-1} = \frac{-i}{\tan\frac{\alpha+k\pi}{n}}.$$

Comme on a calculé toutes les racines la factorisation est immédiate en calculant le coefficient dominant

$$(X+1)^n - e^{2i\alpha}(X-1)^n = (1-e^{2i\alpha})\prod_{k=0}^{n}(X + \frac{i}{\tan\frac{\alpha+k\pi}{n}}).$$

Exercice 8

1. ▹ Établissons déjà l'unicité. Soit P_1, P_2, Q_1, Q_2 de degrés strictement inférieurs à n tels que $(1-X)^n P_1(X) + X^n Q_1(X) = (1-X)^n P_2(X) + X^n Q_2(X)$. Alors, $(1-X)^n(P_1 - P_2) = X^n(Q_2 - Q_1)$. Or, X^n et $(1-X)^n$ sont premiers entre eux donc le lemme de Gauss entraîne que X^n divise le polynôme $P_1 - P_2$ de degré strictement inférieur à n : ainsi, $P_1 = P_2$ et par suite $Q_1 = Q_2$.
▹ Élevons à la puissance $2n-1$ l'égalité $1 = (1-X) + X$

$$\begin{aligned} 1 &= \sum_{k=0}^{2n-1}\binom{2n-1}{k}(1-X)^k X^{2n-1-k} \\ &= \sum_{k=0}^{n-1}\binom{2n-1}{k}(1-X)^k X^{2n-1-k} + \sum_{k=n}^{2n-1}\binom{2n-1}{k}(1-X)^k X^{2n-1-k} \\ &= X^n\sum_{k=0}^{n-1}\binom{2n-1}{k}(1-X)^k X^{n-1-k} + (1-X)^n\sum_{k=n}^{2n-1}\binom{2n-1}{k}(1-X)^{k-n} X^{2n-1-k}, \end{aligned}$$

ce qui fournit bien une décomposition de la forme souhaitée.
2. En substituant $1-X$ à X dans l'identité précédente, on obtient $X^n P(1-X) + (1-X)^n Q(X) = 1$, $\deg P(1-X) < n$ et $\deg Q(X) < n$: l'unicité établie à la première question donne $Q(X) = P(1-X)$.

3. D'après les questions précédentes, $(1-X)^n P(X) = -X^n P(1-X)+1$ ce qui entraîne en dérivant,

$$(1-X)^{n-1}(-nP(X)+(1-X)P'(X)) = X^{n-1}(-nP(1-X)+XP'(1-X)).$$

En utilisant le lemme de Gauss, on obtient que $-nP(X)+(1-X)P'(X)$ est un multiple de X^{n-1} et que son degré est strictement inférieur à n : il existe $\lambda \in \mathbb{R}$ tel que $(1-X)P'(X)-nP(X) = \lambda X^{n-1}$.
4. Notons $P = \sum_{k=0}^{n-1} a_k X^k$ et reportons dans l'équation ci-dessus de sorte à identifier les coefficients des deux membres. Pour le terme de degré $n-1$, on obtient $(2n-1)a_{n-1} = -\lambda$ et pour les termes de degré $k \in [\![0, n-2]\!]$, $a_{k+1} = a_k \frac{n+k}{k+1}$. On peut alors remarquer qu'en évaluant l'expression de l'énoncé en 0, on obtient que $a_0 = P(0) = 1$ puis en déduire pour tout $k \in [\![1, n-1]\!]$,

$$a_k = \prod_{j=0}^{k-1} \frac{n+j}{j+1} = \frac{(n+k-1)!}{k!(n-1)!} = \binom{n+k-1}{k}.$$

Exercice 9

1. Soit P_1, Q_1, P_2 et $Q_2 \in \mathbb{R}[X]$ tels que $A = P_1^2+Q_1^2$ et $B = P_2^2+Q_2^2$. On vérifie immédiatement en s'inspirant de la formule de multiplicativité du module des nombres complexes que

$$AB = (P_1^2+Q_1^2)(P_2^2+Q_2^2) = (P_1P_2-Q_1Q_2)^2+(P_1Q_2+P_2Q_1)^2.$$

Ainsi, AB est somme de deux carrés de polynômes.
2. Le sens direct est immédiat car le carré d'un réel est toujours positif. Pour le sens retour, considérons les facteurs irréductibles de A dans sa décomposition sur $\mathbb{R}[X]$.
Chaque facteur P irréductible de degré 1 apparaît dans A avec un exposant pair $2k$ (car sinon, A change de signe en la racine de P) et $P^{2k} = (P^k)^2+0^2$ est somme de deux carrés.
Chaque facteur irréductible de degré 2 est à discriminant strictement négatif donc son écriture canonique fournit une écriture sous forme de somme de deux carrés.
En utilisant la première question, on en déduit que le polynôme A est somme de deux carrés.

Exercice 10

Pour un tel polynôme P, le polynôme $P-X^{1789}$ admet une infinité de racines donc est nul : d'où $P = X^{1789}$ est la seule solution.

Exercice 11

Le polynôme $XP(X)-1$ admet n racines donc est de degré au moins n. Par conséquent, $\deg P \geq n-1$.

Exercice 12

Soit α, β et γ les trois racines (complexes *a priori*) du polynôme. D'après les relations entre coefficients et racines, $\alpha+\beta+\gamma = \frac{1}{2}$, $\alpha\beta+\beta\gamma+\gamma\alpha = -\frac{7}{2}$ et $\alpha\beta\gamma = -\frac{a}{2}$. Si on suppose que $\alpha+\beta = 1$, la première relation donne $\gamma = -\frac{1}{2}$ et donc $\alpha\beta = a$. En reportant dans la deuxième condition, on obtient $a-\frac{1}{2} = -\frac{7}{2}$ soit $a = -3$.
Réciproquement, pour $a = -3$, α et β sont les racines de X^2-X-3 (d'après la somme et le produit) donc égales à $\frac{1\pm\sqrt{13}}{2}$.

Exercice 13

Commençons par effectuer la division euclidienne de X^4 par P : $X^4 = XP(X)-X^2-X$. Ainsi, la quantité $s = z_1^4+z_2^4+z_3^4$ vaut

$$\begin{aligned} s &= -(z_1^2+z_2^2+z_3^2)-(z_1+z_2+z_3) \\ &= -(z_1+z_2+z_3)^2+2(z_1z_2+z_1z_3+z_2z_3)-(z_1+z_2+z_3) = 2, \end{aligned}$$

d'après les relations entre coefficients et racines.

Exercice 14

Soit $P(X) = X^n + \sum_{k=0}^{n-1} a_k X^k$ un tel polynôme et $z_1, \ldots, z_n$ ses racines. D'après les relations entre coefficients et racines, pour tout $k \in [\![0, n-1]\!]$,

$$a_k = (-1)^{n-k} \sum_{\substack{A \subset [\![1,n]\!] \\ |A| = n-k}} \prod_{l \in A} z_l,$$

donc par inégalité triangulaire,

$$|a_k| \leq \sum_{\substack{A \subset [\![1,n]\!] \\ |A| = n-k}} 1 = \binom{n}{n-k} \leq 2^n.$$

Ainsi, chacun de coefficients entiers a_k prend au plus un nombre fini de valeurs et par conséquent, il y a un nombre fini de tels polynômes.

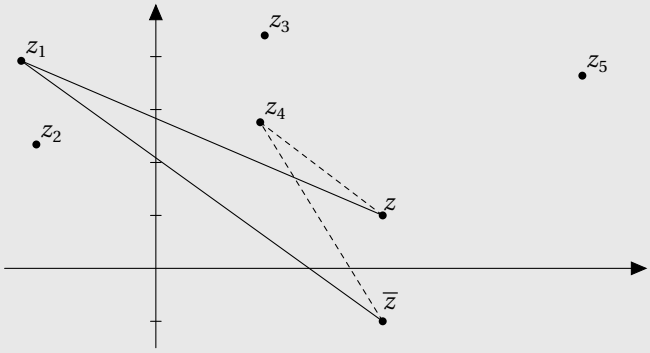
Exercice 15

Il suffit de faire le calcul pour A (quitte à multiplier le polynôme P par i). Soit z une racine de A ; montrons que z est réel.

$$\begin{aligned} |P(z)|^2 &= |A(z) + i\,B(z)|^2 = |B(z)|^2 \\ &= |B(\overline{z})|^2 = |P(\overline{z})|^2. \end{aligned}$$

Notons $P(z) = a \prod_{k=1}^{n} (z - z_k)$.

- Si $\mathfrak{Im}(z) > 0$, alors, pour tout k, $|z - z_k| < |\overline{z} - z_k|$ (car la droite d'équation $\mathfrak{Im}(z) = 0$ est la médiatrice de la droite joignant les points d'affixes z et $\overline{z}$) ;



ainsi $|P(z)| < |P(\overline{z})|$: contradiction.

- Si $\mathfrak{Im}(z) < 0$, alors, pour tout k, $|z - z_k| > |\overline{z} - z_k|$; ainsi $|P(z)| > |P(\overline{z})|$: contradiction.

Ainsi, $\mathfrak{Im}(z) = 0$, c'est-à-dire $z \in \mathbb{R}$.

En conclusion, A est scindé sur $\mathbb{R}$.

Exercice 16

Le retour est élémentaire : pour toute racine $z \in \mathbb{C}$ de P, on a $|P(z)| = 0$ et donc $\mathfrak{Im}(z) = 0$ d'après l'inégalité. Ainsi, toutes les racines de P sont réelles et P est scindé sur $\mathbb{R}$.

Pour le sens direct, notons $P(X) = \prod_{k=1}^{n} (X - z_i)$ avec $z_1, z_2, \ldots, z_n \in \mathbb{R}$. Alors, pour tout $z \in \mathbb{C}$,

$$|P(z)| = \prod_{k=1}^{n} |z - z_i| \geq \prod_{k=1}^{n} |\mathfrak{Im}(z - z_i)| = \prod_{k=1}^{n} |\mathfrak{Im}(z)| = |\mathfrak{Im}(z)|^n,$$

car, pour tout $k \in [\![1, n]\!]$, $\mathfrak{Im}(z_k) = 0$.

Exercice 17

Soit $z \in \mathbb{C}$ une racine de P. Si $|z| \leq 1$, alors $|z| \leq \max\left(1, \sum_{k=0}^{n-1} |a_k|\right)$.
Sinon,

$$|z|^n = \left| -\sum_{k=0}^{n-1} a_k z^k \right| \leq \sum_{k=0}^{n-1} |a_k||z|^k \leq \sum_{k=0}^{n-1} |a_k||z|^{n-1} \text{ car } |z| > 1.$$

Ainsi, en simplifiant par $|z|^{n-1}$,

$$|z| \leq \sum_{k=0}^{n-1} |a_k| \leq \max\left(1, \sum_{k=0}^{n-1} |a_k|\right).$$

Exercice 18

Soit $z \in \mathbb{C}$ une racine multiple de P. Si $z \in \mathbb{Q}$, le résultat est établi ; sinon, on introduit un polynôme $M \in \mathbb{Q}[X]$ admettant z comme racine et de degré minimal parmi les polynômes non nuls à coefficients rationnels et admettant z comme racine. Comme $z \notin \mathbb{Q}$, $\deg M > 1$.

▹ Montrons que si un polynôme $A \in \mathbb{Q}[X]$ admet z comme racine, alors il est un multiple de M. Effectuons la division euclidienne de A par M : $A = QM + R$ avec $\deg R < \deg M$. Comme $R(z) = A(z) - Q(z)M(z) = 0$, la minimalité de M entraîne $R = 0$.

▹ D'après le point précédent, il existe $A \in \mathbb{Q}[X]$ tel que $P = AM$. Si z est racine multiple de M, alors le polynôme $\frac{M}{M \wedge M'}$ admet aussi z comme racine et est de degré strictement inférieur à $\deg M$: contradiction. Ainsi, z est racine simple de M et donc racine de A. D'après le premier point, il existe $B \in \mathbb{Q}[X]$ tel que $A = BM$ et donc $P = BM^2$. Or, $\deg M \geq 2$ et $\deg P = 5$ donc $\deg B = 1$ et B admet une racine rationnelle : on a ainsi exhibé une racine rationnelle pour P.

Corrigés des exercices d'approfondissement du chapitre 12

Exercice A

Soit $x_1, x_2, \ldots, x_m \in \mathbb{Z}$ tels que, pour tout $k \leq m$, $P(x_k) = P(x_1) + k - 1$. Pour tout $k \leq m-1$, l'entier $x_{k+1} - x_k$ divise $P(x_{k+1}) - P(x_k) = 1$ donc est égal à ± 1. Comme les x_k sont deux à deux distincts (car leurs images par P sont distinctes), il existe $\varepsilon \in \{\pm 1\}$ tel que

$$\forall k \leq m, \qquad x_k = x_1 + \varepsilon(k-1).$$

Le polynôme $Q(X) = P(x_1) + \varepsilon(X - x_1)$ et P coïncident en m points ; si $m > n$, alors $Q = P$: contradiction. En conclusion, $m \leq n$.

Exercice B

▹ D'après le théorème de Bézout dans $\mathbb{Q}[X]$, il existe $A, B \in \mathbb{Q}[X]$ tels que $AP + BQ = 1$. Quitte à multiplier par le PPCM des dénominateurs des coefficients de A et de B, on obtient $U, V \in \mathbb{Z}[X]$ et $d \in \mathbb{N} \setminus \{0\}$ tels que $UP + VQ = d$. Ainsi, pour tout $n \in \mathbb{N}$, u_n divise d.

▹ Pour tout $n \in \mathbb{N}$, la formule de Taylor donne

$$\begin{aligned} P(n+d) &= \sum_{k=0}^{\deg P} \frac{P^{(k)}(n)}{k!} d^k \\ &= P(n) + d \sum_{k=1}^{\deg P} \frac{P^{(k)}(n)}{k!} d^{k-1}. \end{aligned}$$

Ainsi, u_n divise $P(n+d)$. De même, u_n divise $Q(n+d)$ donc u_{n+d}.
▹ Pour tout $n \in \mathbb{N}$, la formule de Taylor donne

$$\begin{aligned} P(n) &= P(n+d-d) \\ &= \sum_{k=0}^{\deg P} \frac{P^{(k)}(n+d)}{k!}(-d)^k \\ &= P(n+d) - d\sum_{k=1}^{\deg P} \frac{P^{(k)}(n+d)}{k!}(-d)^{k-1}. \end{aligned}$$

Ainsi, u_{n+d} divise $P(n)$. De même, u_{n+d} divise $Q(n)$ donc u_n.
Comme la suite (u_n) est à valeurs positives, on en déduit que $u_n = u_{n+d}$ pour tout $n \in \mathbb{N}$.

Exercice C

Posons $P = \sum_{k=0}^{49} a_k X^k$. Alors, pour tout $p \in E$ et tout $z \in \mathbb{U}_p \setminus \{1\}$,

$$\begin{aligned} P(z) &= \sum_{k=0}^{49} a_k z^k = \sum_{r=0}^{p-1} \sum_{\substack{k=0 \\ k=r[p]}}^{49} a_k z^k = \sum_{r=0}^{p-1} \sum_{\substack{k=0 \\ k=r[p]}}^{49} a_k z^r \\ &= \sum_{r=0}^{p-1} \sum_{\substack{k=0 \\ k=0[p]}}^{49} a_k z^r \text{ avec l'hypothèse de l'énoncé,} \\ &= \sum_{\substack{k=0 \\ k=0[p]}}^{49} a_k \sum_{r=0}^{p-1} z^r = 0 \text{ car } z \in \mathbb{U}_p \setminus \{1\}. \end{aligned}$$

Ainsi, tous les éléments de $\mathbb{U}_3 \cup \mathbb{U}_5 \cup \mathbb{U}_7 \cup \mathbb{U}_{11} \cup \mathbb{U}_{13} \cup \mathbb{U}_{17} \setminus \{1\}$ sont racines de P. Or, pour p et q premiers distincts, $\mathbb{U}_p \cap \mathbb{U}_q = \{1\}$ donc $\mathbb{U}_p \setminus \{1\}$ et $\mathbb{U}_q \setminus \{1\}$ sont disjoints. Par suite,

$$|\mathbb{U}_3 \cup \mathbb{U}_5 \cup \mathbb{U}_7 \cup \mathbb{U}_{11} \cup \mathbb{U}_{13} \cup \mathbb{U}_{17} \setminus \{1\}| = 2+4+6+10+12+16 = 50.$$

Ainsi, P est un polynôme de degré au plus 49 qui admet 50 racines, il est donc nul : tous les coefficients $(a_k)_{k\in[0,49]}$ sont nuls.

Exercice D

Supposons que $Q = P^2 + \alpha^2$ admette une racine multiple z, donc que $Q(z) = Q'(z) = 0$ soit encore $P(z)^2 + \alpha^2 = 0$ et $2P(z)P'(z) = 0$. Or, les racines de P (par hypothèse) et de P' (déduit de l'hypothèse avec le théorème de Rolle) sont toutes réelles : la condition $2P(z)P'(z) = 0$ entraîne que $z \in \mathbb{R}$ et donc $P(z)^2 \geq 0$. Par conséquent, $P(z)^2 + \alpha^2 > 0$ (car $\alpha \in \mathbb{R}^*$) : contradiction.
Ainsi, le polynôme Q n'admet que des racines complexes simples : il est scindé à racines simples dans $\mathbb{C}$.

Exercice E

1. Un polynôme unitaire stable s'écrit comme produit de termes de la forme $X - \lambda$ où λ est une racine strictement négative et de la forme $(X - \alpha - i\beta)(X - \alpha + i\beta) = X^2 - 2\alpha X + \alpha^2 + \beta^2$ où $\alpha + i\beta$ est une racine non réelle de partie réelle α strictement négative. Par conséquent, les coefficients de P sont strictement positifs.
2. Pour le sens direct, il suffit d'appliquer la première question à P et Q (qui est aussi réel car ses racines non réelles sont deux à deux conjuguées et stable car la partie réelle d'une somme de réels strictement négatifs est strictement négative).
Pour le sens retour, il est évident que P et Q ne peuvent avoir de racines réelles positives (pour tout $x \in \mathbb{R}_+$, $P(x) \geq P(0) > 0$). Or, si $\alpha + i\beta$ est une racine non réelle de P, alors $2\alpha = \alpha + i\beta + \alpha - i\beta$ est une racine réelle de Q donc strictement négative.

Exercice F

Considérons $R = P - Q$; notons $n = \deg P \geq \deg Q$, $p_1 = \operatorname{card} P^{-1}(\{z_1\})$ et $p_2 = \operatorname{card} P^{-1}(\{z_2\})$.

- Les p_1 éléments de $P^{-1}(\{z_1\})$ sont les racines de $P - z_1$ et la somme de leurs multiplicités est n. Elles sont alors aussi racines de $P' = (P - z_1)'$ avec pour somme des multiplicités $n - p_1$.
- De même, les p_2 éléments de $P^{-1}(\{z_2\})$ sont alors aussi racines de $P' = (P - z_2)'$ avec pour somme des multiplicités $n - p_2$.

Au final, nous avons trouvé des racines de P' avec pour multiplicité totale $2n - p_1 - p_2$ donc, en comparant au degré, $2n - p_1 - p_2 \leq n - 1$ soit $p_1 + p_2 \geq n + 1$. Mais les $p_1 + p_2$ éléments de $P^{-1}(\{z_1\}) \cup P^{-1}(\{z_2\})$ sont des racines de R qui est de degré au plus n : ainsi, $R = 0$ donc $P = Q$.

Corrigés des problèmes du chapitre 12

Problème 1

1. Si X s'écrit comme la somme de k puissances $n^{\text{ièmes}}$ de polynômes $X = Q_1^n(X) + \ldots + Q_k^n(X)$, alors tout polynôme $P \in \mathbb{C}[X]$ s'écrit $P(X) = Q_1^n(P(X)) + \ldots + Q_k^n(P(X))$ donc est aussi somme de k puissances $n^{\text{ièmes}}$ de polynômes. Ainsi, $k(n)$ est le nombre minimal de polynômes nécessaires pour écrire X comme somme de puissances $n^{\text{ièmes}}$.

2. Le monôme X n'est pas le carré d'un polynôme puisque son degré est impair donc $k(2) \geq 2$. De plus,

$$X = \left(X + \frac{1}{4}\right)^2 + \left(iX - \frac{i}{4}\right)^2.$$

D'où, $k(2) = 2$.

3. a. Soit A désigne l'ensemble des racines $n^{\text{ième}}$ de -1. Alors

$$Q_1(X)^n + Q_2(X)^n = \prod_{\eta \in A} (Q_1(X) - \eta Q_2(X)).$$

b. Soit $n \geq 3$ tel que $k(n) = 2$; il existe $Q_1, Q_2 \in \mathbb{C}[X]$ tels que

$$X = Q_1(X)^n + Q_2(X)^n = \prod_{\eta \in A} (Q_1(X) - \eta Q_2(X)).$$

Ainsi, il y a $n - 1$ polynômes constants parmi les $Q_1(X) - \eta Q_2(X)$; en effectuant la différence de deux de ces polynômes, on obtient que Q_2 est constant et l'on a une contradiction en considérant le degré $1 = \deg X = n \deg Q_1$.

4. a. Montrons par récurrence, pour tout $k \leq n - 1$, que

$$\Delta^k X^n = \sum_{j=0}^{k} \binom{k}{j} (-1)^{k+j} (X + j)^n.$$

▹ Cette égalité est évidemment vraie pour $k = 0$.

▹ Soit $k < n-1$. Supposons la formule établie au rang k et calculons $\Delta^{k+1}X^n$.

$$\begin{aligned}
\Delta^{k+1}X^n &= \sum_{j=0}^{k}\binom{k}{j}(-1)^{k+j}(X+1+j)^n - \sum_{j=0}^{k}\binom{k}{j}(-1)^{k+j}(X+j)^n \\
&= \sum_{j=1}^{k+1}\binom{k}{j-1}(-1)^{k+j+1}(X+j)^n - \sum_{j=0}^{k}\binom{k}{j}(-1)^{k+j}(X+j)^n \\
&= \sum_{j=1}^{k}\binom{k+1}{j}(-1)^{k+1+j}(X+j)^n - (-1)^k X^n + (X+k+1)^n \\
&= \sum_{j=0}^{k+1}\binom{k+1}{j}(-1)^{k+1+j}(X+j)^n.
\end{aligned}$$

b. D'après le point précédent, le polynôme $\Delta^{n-1}X^n$ s'écrit comme somme de n puissances $n^{\text{ièmes}}$ de polynômes ; en effet, posons pour tout $j \in \{0,\dots,n-1\}$, $\eta_j \in \mathbb{C}$ tel que $\eta_j^n = \binom{n-1}{j}(-1)^{n-1+j}$ et remarquons que

$$\Delta^{n-1}X^n = \sum_{j=0}^{n-1}\binom{n-1}{j}(-1)^{n-1+j}(X+j)^n = \sum_{j=0}^{n-1}(\eta_j(X+j))^n.$$

Remarquons enfin que $\Delta^k X^n$ est un polynôme de degré $n-k$ et de coefficient dominant $\frac{n!}{(n-k)!}$; en effet, $\Delta^0 X^n = X^n$ et si $\Delta^k X^n = \frac{n!}{(n-k)!}X^{n-k} + P(X)$ avec $P \in \mathbb{C}_{n-1-k}[X]$, alors

$$\begin{aligned}
\Delta^{k+1}X^n &= \frac{n!}{(n-k)!}(X+1)^{n-k} + P(X+1) - \frac{n!}{(n-k)!}X^{n-k} + P(X) \\
&= \frac{n!}{(n-k)!}\sum_{j=0}^{n-k-1}\binom{n-k}{j}X^j + P(X+1) - P(X) \\
&= \frac{n!}{(n-k)!}\binom{n-k}{n-k-1}X^{n-k-1} + \frac{n!}{(n-k)!}\sum_{j=0}^{n-k-2}\binom{n-k}{j}X^j + P(X+1) - P(X) \\
&= \frac{n!}{(n-k-1)!}X^{n-k-1} + \frac{n!}{(n-k)!}\sum_{j=0}^{n-k-2}\binom{n-k}{j}X^j + P(X+1) - P(X);
\end{aligned}$$

or, le polynôme $\frac{n!}{(n-k)!}\sum_{j=0}^{n-k-2}\binom{n-k}{j}X^j + P(X+1) - P(X)$ est de degré strictement inférieur à $n-k-1$ ce qui établit le résultat annoncé.
En conclusion, $\Delta^{n-1}X^n$ est un polynôme de degré 1 qui s'écrit comme somme de n puissances $n^{\text{ièmes}}$ de polynômes donc il existe $(a,b) \in \mathbb{R}^* \times \mathbb{C}$ tel que $\Delta^{n-1}X^n = aX + b$ et $Q_1,\dots,Q_n \in \mathbb{C}[X]$ tels que

$$\Delta^{n-1}X^n = Q_1(X)^n + \dots + Q_n(X)^n.$$

Ainsi, $X = Q_1(\frac{X-b}{a})^n + \dots + Q_n(\frac{X-b}{a})^n$ et donc, d'après la première question, $k(n) \leq n$.

Problème 2

Partie A – Préliminaires

1. La condition sur le degré est évidente et $(X^n)^m = X^{mn} = (X^m)^n$ pour tout $m,n \in \mathbb{N}^*$.
2. a. ▹ Supposons qu'il existe deux polynômes T_n et T_n' vérifiant ces conditions ; alors, pour tout $x \in \mathbb{R}$, $(T_n - T_n')(\cos(x)) = 0$. Le polynôme $T_n - T_n'$ admet une infinité de racines (tout l'intervalle $[-1,1]$) donc est nul : $T_n = T_n'$.

▹ En utilisant les formules d'Euler, on obtient, pour tout $x \in \mathbb{R}$:

$$\begin{aligned}\cos(nx) = \mathfrak{Re}(e^{inx}) &= \mathfrak{Re}((\cos x + i\sin x)^n)\\ &= \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} \binom{n}{2k} (-1)^k (1-\cos^2 x)^k \cos^{n-2k} x.\end{aligned}$$

Ainsi, le polynôme $T_n(X) = \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} \binom{n}{2k} (-1)^k (1-X^2)^k X^{n-2k}$ convient (le terme dominant étant $2^{n-1}X^n$).
b. La condition sur le degré est vérifiée d'après le point précédent.
Par ailleurs, pour tout $x \in \mathbb{R}$, $T_m \circ T_n(\cos x) = T_m(\cos(nx)) = \cos(mnx)$. Par unicité de la famille $(T_n)_{n\in\mathbb{N}^*}$, on obtient que $T_m \circ T_n = T_{mn}$ (donc est aussi égal à $T_n \circ T_m$). En conclusion, la famille $(T_n)_{n\in\mathbb{N}^*}$ satisfait les conditions (12.1).
3. Notons G l'ensemble des polynômes de degré 1.
- La composition des polynômes est une loi de composition interne sur G car $\deg P \circ Q = \deg P \deg Q$.
- La loi $\circ$ est associative, d'élément neutre $X \in G$.
- Le symétrique du polynôme $\alpha X + \beta$ avec $\alpha \neq 0$ est $\frac{1}{\alpha}X - \frac{\beta}{\alpha}$.

Par conséquent, G est un groupe pour la loi $\circ$.

Partie B – Caractérisation

1. En notant $A(X) = \alpha X + \beta$ et $P(X) = a_2X^2 + a_1X + a_0$, on obtient :

$$A \circ P_2 \circ A^{-1} = \frac{a_2}{\alpha}X^2 + \left(a_1 - 2\frac{\beta.a_2}{\alpha}\right)X + \frac{\beta^2 a_2}{\alpha} - \beta a_1 + \beta.$$

Il suffit alors de poser $\alpha = a_2$ (qui est non nul car $a_2 \neq 0$) et $\beta = \frac{a_1}{2}$ pour avoir le résultat annoncé.
2. La condition sur le degré est vérifiée puisque $\deg Q_n = \deg A \deg P_n \deg A^{-1} = \deg P_n$. De plus, pour tout $n, m \in \mathbb{N}^*$:

$$\begin{aligned}Q_m \circ Q_n &= A \circ P_m \circ A^{-1} \circ A \circ P_n \circ A^{-1} = A \circ P_m \circ P_n \circ A^{-1}\\ &= A \circ P_n \circ P_m \circ A^{-1} = A \circ P_n \circ A^{-1} \circ A \circ P_m \circ A^{-1}\\ &= Q_n \circ Q_m.\end{aligned}$$

Ainsi, la famille $(Q_n = A \circ P_n \circ A^{-1})_{n\in\mathbb{N}^*}$ satisfait les conditions (12.1). En utilisant que $Q_2 \circ Q_n = Q_n \circ Q_2$, on obtient que a_n, le coefficient dominant de Q_n, vérifie $a_n^2 = a_n$ d'où $a_n = 1$: Q_n est unitaire.
3. Notons $Q_3(X) = X^3 + \alpha X^2 + \beta X + \gamma$. La condition $Q_2 \circ Q_3 = Q_3 \circ Q_2$ entraîne (après calculs) $3a = 2\beta$ et $3a^2 + \beta = \beta^2$ d'où $a \in \{0, -2\}$.
4. S'il existe deux polynômes unitaires de degré n, P_n et Q_n, qui commutent avec Q_2, alors $P_n(X^2+a) = P_n(X)^2 + a$ et $Q_n(X^2+a) = Q_n(X)^2 + a$. Par conséquent,

$$(P_n - Q_n)(X^2+a) = (P_n - Q_n)(X)(P_n + Q_n)(X)$$

et

$$2\deg(P_n - Q_n) = \deg(P_n - Q_n) + \deg(P_n + Q_n)$$

ce qui est contradiction avec $\deg(P_n - Q_n) < n = \deg(P_n + Q_n)$.
5. a. Supposons $a = 0$. Le polynôme X^n est de degré n et commute avec $Q_2 = X^2$: d'après la propriété d'unicité, $Q_n(X) = X^n$.
b. Supposons $a = -2$. On trouve $B \circ T_2 \circ B^{-1} = Q_2$ pour $B = 2X$ (car $T_2(X) = 2X^2 - 1$). Le polynôme $B \circ T_n \circ B^{-1}$ est de degré n et commute avec $Q_2(X) = X^2$: par unicité $Q_n = B \circ T_n \circ B^{-1}$.
6. Les seules familles qui vérifient (12.1) sont conjuguées avec la base canonique ou la famille des polynômes de Tchebychev.

CORRIGÉS

Problème 3

Partie A – Critère de Perron

1. a. Supposons que z soit une racine de P de module 1. Alors

$$a_{n-1}z^{n-1} = -z^n - \sum_{k=0}^{n-2} a_k z^k.$$

En passant au module et en utilisant l'inégalité triangulaire, on obtient

$$|a_{n-1}| \leq 1 + \sum_{k=0}^{n-2} |a_k|,$$

ce qui contredit l'hypothèse sur le polynôme P.
b. i. En développant, on obtient successivement

$$\begin{aligned} P &= (X - z_1)\sum_{k=0}^{n-1} b_k X^k \\ &= \sum_{k=0}^{n-1} b_k X^{k+1} - \sum_{k=0}^{n-1} z_1 b_k X^k \\ &= \sum_{k=1}^{n} b_{k-1} X^k - \sum_{k=0}^{n-1} z_1 b_k X^k \\ &= b_{n-1}X^n + \sum_{k=1}^{n-1} (b_{k-1} - z_1 b_k) X^k + b_0. \end{aligned}$$

En identifiant les coefficients, on obtient $b_{n-1} = 1$, $-b_0 z_1 = a_0$ puis la relation demandée.
ii. Utilisons maintenant l'hypothèse de l'énoncé

$$\begin{aligned} |b_{n-2} - z_1| = |a_{n-1}| &> 1 + \sum_{k=0}^{n-2} |a_k| \\ &> 1 + \sum_{k=1}^{n-2} |b_{k-1} - z_1 b_k| + |b_0 z_1| \\ &> 1 + \sum_{k=1}^{n-2} (|z_1 b_k| - |b_{k-1}|) + |b_0 z_1| \\ &> 1 + \sum_{k=0}^{n-2} |z_1|.|b_k| - \sum_{k=0}^{n-2} |b_k| + |b_{n-2}| \\ &> 1 + |b_{n-2}| + (|z_1| - 1)\sum_{k=0}^{n-2} |b_k|. \end{aligned}$$

Majorons maintenant (toujours par inégalité triangulaire), $|b_{n-2} - z_1| < |b_{n-2}| + |z_1|$,

$$|z_1| - 1 > (|z_1| - 1)\sum_{k=0}^{n-2} |b_k|$$

soit en divisant par le réel strictement positif $|z_1| - 1$,

$$1 > \sum_{k=0}^{n-2} |b_k|.$$

iii. Supposons que $\sum_{k=0}^{n-1} b_k X^k$ admette une racine z de module supérieur à 1. Alors,

$$z^{n-1} = -\sum_{k=0}^{n-2} b_k z^k.$$

En passant au module (avec l'inégalité triangulaire),

$$|z^{n-1}| \leq \sum_{k=0}^{n-2} |b_k z^k| \leq |z^{n-1}| \sum_{k=0}^{n-2} |b_k|.$$

Ceci contredit la question précédente (après simplifications par le réel strictement positif $|z^{n-1}|$). Ainsi, Q n'admet pas de racine de module supérieur à 1.
c. Une fois factorisé par $X - z_1$ où z_1 est de module strictement supérieur à 1, P n'admet plus de racine de module supérieure ou égale à 1.
2. a. Les coefficients constants de A et B sont des entiers non nuls donc supérieurs ou égaux à 1. Comme, d'après les relations entre coefficients et racines, ces termes sont les produits des racines du polynôme correspondant, cela signifie que A et B admettent au moins une racine de module supérieur ou égal à 1. Ainsi, P admet deux racines ou une racine double de module supérieur ou égal à 1, ce qui contredit la question précédente.
b. En conclusion, A ou B est constant et cette constante divise le coefficient dominant de P donc est égal à 1 ou -1 : P est irréductible sur $\mathbb{Z}$.

Partie B – Critère de Polya

a. i. D'après la formule d'interpolation de Lagrange (P est de degré n et on interpole en $n+1$ points),

$$P = \sum_{k=0}^{n} P(\alpha_k) \prod_{\substack{l=0 \\ l \neq k}}^{n} \frac{X - \alpha_l}{\alpha_k - \alpha_l}.$$

Alors le coefficient dominant de P est celui devant X^n, à savoir

$$\sum_{k=0}^{n} P(\alpha_k) \prod_{\substack{l=0 \\ l \neq k}}^{n} \frac{1}{\alpha_k - \alpha_l}.$$

ii. Comme le polynôme P est à coefficients entiers,

$$1 \leq \left| \sum_{k=0}^{n} P(\alpha_k) \prod_{\substack{l=0 \\ l \neq k}}^{n} \frac{1}{\alpha_k - \alpha_l} \right| \leq \max_k |P(\alpha_k)| . \sum_{k=0}^{n} \prod_{\substack{l=0 \\ l \neq k}}^{n} \frac{1}{|\alpha_k - \alpha_l|}$$

donc

$$\begin{aligned} \max_k |P(\alpha_k)| &\geq \frac{1}{\sum_{k=0}^{n} \prod_{\substack{l=0 \\ l \neq k}}^{n} \frac{1}{|\alpha_k - \alpha_l|}} \geq \frac{1}{\sum_{k=0}^{n} \prod_{\substack{l=0 \\ l \neq k}}^{n} \frac{1}{|k-l|}} \\ &\geq \frac{1}{\sum_{k=0}^{n} \frac{1}{k!(n-k)!}} \geq \frac{n!}{\sum_{k=0}^{n} \binom{n}{k}} = n! 2^{-n}. \end{aligned}$$

Ainsi, il existe $k \in [\![0, n]\!]$ tel que $P(\alpha_k) \geq k! 2^{-k}$.
b. i. Si $\deg A < M$, alors $\deg B = n - \deg A \geq M$ et il suffit de permuter les rôles de A et B.
ii. Effectuons le quotient :

$$\frac{(\deg A)! 2^{-\deg A}}{M! 2^{-M}} = \frac{\deg A}{2} \frac{\deg A - 1}{2} \cdots \frac{M+1}{2} \geq 1,$$

d'où le résultat attendu.

iii. D'après la question 1 de cette partie, il existe $k \in [\![1,n]\!]$ tel que $|A(\alpha_k)| \geq (\deg A)!2^{-\deg A}$. Or, $A(\alpha_k)$ est un diviseur de $P(\alpha_k)$ donc

$$|A(\alpha_k)| \leq |P(\alpha_k)| \leq M!2^{-M}.$$

Les deux dernières inégalités sont contradictoires d'après la question b. Par conséquent, l'un des polynômes A ou B est constant et divise le coefficient dominant de P donc est égal à 1 ou -1.

Partie C – Exemples classiques

1. Écrivons $\prod_{k=1}^{n}(X-\alpha_k)-1 = A(X)B(X)$ avec A et B à coefficients entiers.
Pour tout $k \in [\![1,n]\!]$, $A(\alpha_k)B(\alpha_k) = -1$ donc $A(\alpha_k) = -B(\alpha_k)$. Ainsi, $A+B$ admet les α_k pour racines.
▹ Si l'un des polynômes A ou B est non constant, alors $A+B$ est de degré au plus $n-1$ donc nul puisqu'il y a n racines distinctes. Mais alors, $\prod_{k=1}^{n}(X-\alpha_k)-1 = -A(X)^2$ ce qui est impossible (comme on le voit en considérant le coefficient dominant).
▹ Sinon, l'un des polynômes est constant donc égal à ± 1 car cette constante doit diviser le coefficient dominant de $A(X)B(X)$.
En conclusion, $\prod_{k=1}^{n}(X-\alpha_k)-1$ est irréductible sur $\mathbb{Z}$.
2. Écrivons $\prod_{k=1}^{n}(X-\alpha_k)+1 = A(X)B(X)$ avec A et B à coefficients entiers.
Pour tout $k \in [\![1,n]\!]$, $A(\alpha_k)B(\alpha_k) = 1$ donc $A(\alpha_k) = B(\alpha_k)$. Ainsi, $A-B$ admet les α_k pour racines.
▹ Si l'un des polynômes A ou B est non constant, alors $A-B$ est de degré au plus $n-1$ donc nul puisqu'il y a n racines distinctes. Mais alors, $\prod_{k=1}^{n}(X-\alpha_k)-1 = A(X)^2$ ce qui est impossible (pour des raisons de degré) lorsque n est impair.
▹ Sinon, l'un des polynômes est constant donc égal à ± 1 car cette constante doit diviser le coefficient dominant de $A(X)B(X)$.
En conclusion, pour n impair, $\prod_{k=1}^{n}(X-\alpha_k)+1$ est irréductible sur $\mathbb{Z}$.
3. Écrivons $\prod_{k=1}^{n}(X-\alpha_k)^2+1 = A(X)B(X)$ avec A et B à coefficients entiers.
Pour tout $k \in [\![1,n]\!]$, $A(\alpha_k)B(\alpha_k) = 1$ donc $A(\alpha_k) = B(\alpha_k) \in \{\pm 1\}$. Par ailleurs, les polynômes A et B sont de signe constant car la fonction $x \mapsto A(x)B(x)$ est continue et ne s'annule pas. Ainsi, il existe $\varepsilon \in \{\pm 1\}$ tel que, pour tout $k \in [\![1,n]\!]$, $A(\alpha_k) = B(\alpha_k) = \varepsilon$.
▹ Si $\deg A < n$ (ou de manière analogue si $\deg B < n$), alors $A(X)-\varepsilon$ est le polynôme nul.
▹ Sinon $\deg A = \deg B = n$, alors $A-B$ est de degré au plus $n-1$ (car les deux polynômes sont unitaires) et admet n racines donc est nul. Mais alors, $\prod_{k=1}^{n}(X-\alpha_k)^2+1 = A(X)^2$ donc

$$1 = \left(A(X) - \prod_{k=1}^{n}(X-\alpha_k)\right)\left(A(X) + \prod_{k=1}^{n}(X-\alpha_k)\right),$$

ce qui est impossible pour des raisons de degré.
En conclusion, $\prod_{k=1}^{n}(X-\alpha_k)^2+1$ est irréductible sur $\mathbb{Z}$.

Chapitre 13

Fractions rationnelles

1. Définitions — **2.** Décomposition en éléments simples (cas scindé) — **3.** Décomposition en éléments simples (cas général)

Objectifs et compétences du programme

- Trouver l'écriture irréductible d'une fraction rationnelle.
- Généraliser la notion de degré aux fractions rationnelles.
- Calculer la partie entière d'une fraction donnée.
- Connaître la forme de la décomposition en éléments simples.
- Calculer dans des cas simples la décomposition en éléments simples.

Édouard Lucas, mathématicien français (1842-1891)
Nous avons déjà rencontré Lucas pour la suite d'entiers qui porte son nom et qui généralise celle de Fibonacci; il a aussi laissé à l'arithmétique un test de primalité particulièrement adapté aux nombres de Mersenne, désormais appelé test de Lucas-Lehmer. Il a notamment montré que $2^{2^7-1}-1$ est premier.
On le retrouve dans un exercice de ce chapitre pour le théorème dit de Gauss-Lucas (même s'il n'a jamais connu Gauss).
On lui attribue également le problème des tours de Hanoï et la « Pipopilette » ou jeux des petits carrés.

COURS

Comme on est passé des entiers aux rationnels, on peut définir des fractions à partir des polynômes : ce sont les fractions rationnelles. On peut alors les manipuler à partir des propriétés des polynômes (degré, racines, irréductibilité...). L'intérêt majeur des fractions rationnelles est qu'il existe une forme « canonique » permettant de nombreux calculs : la décomposition en éléments simples.

1. Définitions

1.1. Fractions rationnelles

Proposition 13.1.

La relation $\sim$ définie sur l'ensemble $\mathbb{K}[X] \times \mathbb{K}[X] \setminus \{0\}$ par $(P_1, Q_1) \sim (P_2, Q_2)$ si $P_1 Q_2 = P_2 Q_1$ est une relation d'équivalence.

La vérification des propriétés de relation d'équivalence (réflexivité, symétrie, transitivité) ne pose aucun problème.

Définition 13.2.

Une fraction rationnelle sur $\mathbb{K}$ est une classe d'équivalence de $\mathbb{K}[X] \times \mathbb{K}[X] \setminus \{0\}$ pour $\sim$. Chaque couple d'une classe d'équivalence est appelé un représentant de la fraction.
Leur ensemble est noté $\mathbb{K}(X)$.

Notation

On adopte la notation $\frac{P}{Q}$ pour la classe du couple $(P, Q) \in \mathbb{K}[X] \times \mathbb{K}[X] \setminus \{0\}$, P et Q sont alors appelés respectivement numérateur et dénominateur de (de représentant de) la fraction.

On remarque que « simplifier » une fraction correspond à un autre choix de représentant dans la classe d'équivalence.
On vérifie immédiatement que la définition des lois de composition internes ci-dessous ne dépend pas des représentants choisis.

Définition 13.3.

Soit deux fractions rationnelles de représentants $\frac{P_1}{Q_1}$ et $\frac{P_2}{Q_2}$. Alors,

- la somme est la fraction dont un représentant est $\frac{P_1 Q_2 + P_2 Q_1}{Q_1 Q_2}$;
- le produit est la fraction dont un représentant est $\frac{P_1 P_2}{Q_1 Q_2}$.

Proposition 13.4.

L'ensemble $\mathbb{K}(X)$ est un corps.

Définition 13.5.

Un représentant $\frac{P}{Q}$ d'une fraction rationnelle est irréductible si P et Q sont premiers entre eux.

En utilisant le PGCD du numérateur et du dénominateur d'un représentant, on obtient la proposition suivante.

Proposition 13.6.

Toute fraction rationnelle admet un représentant irréductible.
De plus, les numérateurs (respectivement dénominateurs) de deux représentants irréductibles d'une même fraction rationnelle sont associés.

Exemple

Le représentant $\frac{X^6+1}{X^3+1}$ est irréductible ; en effet, $(X^6+1)\wedge(X^3+1)=1$ car

$$X^6+1=(X^3-1)(X^3+1)+2.$$

Définition 13.7.

Les racines (respectivement les pôles) de multiplicité r d'une fraction rationnelle sont les racines de multiplicité r du numérateur (respectivement du dénominateur) d'un représentant irréductible de cette fraction.

Exemple

La fraction $\frac{X^2-X}{(X^2-1)^2}$ n'admet qu'une racine 0 et deux pôles 1 et -1 ce dernier étant de multiplicité 2.
En effet,

$$\frac{X^2-X}{(X^2-1)^2}=\frac{(X-1)X}{(X-1)^2(X+1)^2}=\frac{X}{(X-1)(X+1)^2}.$$

1.2. Degré

On remarque que si $\frac{P_1}{Q_1}=\frac{P_2}{Q_2}$, alors $P_1Q_2=P_2Q_1$ donc

$$\deg P_1+\deg Q_2=\deg P_2+\deg Q_1$$

soit encore

$$\deg P_1-\deg Q_1=\deg P_2-\deg Q_2$$

ce qui permet de justifier de la correction de la définition du degré d'une fraction rationnelle (c'est-à-dire de l'indépendance du représentant).

Définition 13.8.

Le degré d'une fraction rationnelle est la différence du degré du numérateur et du degré du dénominateur d'un représentant (quelconque) de cette fraction.

Remarque
En identifiant le polynôme P avec la fraction de représentant $\frac{P}{1}$, on remarque que le degré d'une fraction étend le concept de degré des polynômes.

Remarque
Les polynômes non nuls sont des fractions rationnelles de degré positif.
La réciproque est fausse : la fraction $\frac{X^2+1}{X}$ n'est pas un polynôme mais est pourtant de degré $1>0$.

On étend les propriétés de degré des polynômes.

Proposition 13.9.

Soit $\frac{P_1}{Q_1}$, $\frac{P_2}{Q_2}$ deux fractions rationnelles à coefficients dans $\mathbb{K}$ et $\lambda \in \mathbb{K} \setminus \{0\}$.
▹ $\deg(\frac{P_1}{Q_1}+\frac{P_2}{Q_2}) \leq \max(\deg \frac{P_1}{Q_1}, \deg \frac{P_2}{Q_2})$.
▹ $\deg(\lambda.\frac{P_1}{Q_1}) = \deg \frac{P_1}{Q_1}$.
▹ $\deg(\frac{P_1}{Q_1}.\frac{P_2}{Q_2}) = \deg \frac{P_1}{Q_1} + \deg \frac{P_2}{Q_2}$.

Démonstration
▹ Par définition de l'addition de fractions rationnelles et du degré :

$$\begin{aligned}
\deg\Big(\frac{P_1}{Q_1}+\frac{P_2}{Q_2}\Big) &= \deg \frac{P_1Q_2+P_2Q_1}{Q_1Q_2} \\
&= \deg(P_1Q_2+P_2Q_1) - \deg(Q_1Q_2) \\
&\leq \max(\deg(P_1Q_2), \deg(P_2Q_1)) - \deg(Q_1Q_2) \\
&\leq \max(\deg P_1 + \deg Q_2, \deg P_2 + \deg Q_1) - \deg Q_1 - \deg Q_2 \\
&\leq \max(\deg P_1 - \deg Q_1, \deg P_2 - \deg Q_2) = \max(\deg \tfrac{P_1}{Q_1}, \deg \tfrac{P_2}{Q_2}).
\end{aligned}$$

▹ Il suffit dans ce cas de remarquer que $\deg \lambda P_1 = \deg P_1$.
▹ Par définition du produit et du degré :

$$\begin{aligned}
\deg\Big(\frac{P_1}{Q_1}.\frac{P_2}{Q_2}\Big) &= \deg \frac{P_1P_2}{Q_1Q_2} \\
&= \deg(P_1P_2) - \deg(Q_1Q_2) \\
&= \deg P_1 + \deg P_2 - \deg Q_1 - \deg Q_2 \\
&= \deg P_1 - \deg Q_1 + \deg P_2 - \deg Q_2 \\
&= \deg \tfrac{P_1}{Q_1} + \deg \tfrac{P_2}{Q_2}.
\end{aligned}$$

■

1.3. Partie entière

Définition 13.10.

La partie entière de la fraction de représentant $\frac{P}{Q}$ est le quotient de division euclidienne de P par Q.

Proposition 13.11.

La partie entière de la fraction de représentant $\frac{P}{Q}$ est l'unique polynôme E tel que $\deg(\frac{P}{Q}-E)<0$, c'est-à-dire s'il existe $R\in\mathbb{K}[X]$ vérifiant

$$\frac{P}{Q}=E+\frac{R}{Q},\qquad \deg R<\deg Q.$$

Démonstration

▹ Il est clair que le quotient E de la division de P par Q vérifie cette condition puisque la condition sur le degré du reste $\deg(P-EQ)<\deg Q$ se réécrit $\deg(\frac{P}{Q}-E)<0$.

▹ Supposons qu'il existe deux polynômes E_1 et E_2 vérifiant ces conditions. Alors,

$$E_1-E_2=\Big(\frac{P}{Q}-E_2\Big)-\Big(\frac{P}{Q}-E_1\Big).$$

Or, le premier membre est un polynôme et le second est de degré strictement négatif : les deux membres sont donc nuls et $E_1=E_2$. ■

Exemple

Calculons la partie entière de $\frac{X^n}{X-1}$ pour $n\geq 2$. Comme

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

on obtient

$$\frac{X^n}{X-1}=\sum_{k=0}^{n-1}X^k+\frac{1}{X-1},$$

donc la partie entière est $\sum_{k=0}^{n-1}X^k$.

Une simple conséquence de cette définition est le résultat suivant.

Proposition 13.12.

La partie entière de la combinaison linéaire de fractions rationnelles est la combinaison linéaire des parties entières de ces fractions rationnelles.

Démonstration

Soit R_1 et R_2 des fractions de parties entières E_1 et E_2 respectivement, $\lambda\in\mathbb{K}$. Alors,

$$\deg((R_1+\lambda R_2)-(E_1+\lambda E_2))\leq\max(\deg(R_1-E_1),\deg(\lambda(R_2-E_2)))<0.$$

Ainsi, la partie entière de la fraction rationnelle $R_1+\lambda R_2$ est $E_1+\lambda E_2$. ■

L'objectif du reste de ce cours est alors d'obtenir une écriture simple et manipulable des fractions rationnelles en général et des fractions rationnelles de degré négatif en particulier.

2. Décomposition en éléments simples (dénominateur scindé)

On suppose dans un premier temps que les fractions rationnelles considérées admettent un dénominateur scindé.

En vertu du théorème de D'Alembert-Gauss, tous les polynômes non constants de $\mathbb{C}[X]$ sont scindés ; par conséquent, les résultats de cette section s'appliquent aux fractions de $\mathbb{C}(X)$ (mais pas seulement).

2.1. Théorème de décomposition

Commençons par le cas particulier où le dénominateur est scindé avec un unique pôle.

Proposition 13.13.

Soit $P \in \mathbb{K}[X]$, $a \in \mathbb{K}$ et $\alpha \in \mathbb{N} \setminus \{0\}$. Il existe un unique $E \in \mathbb{K}[X]$ et une unique famille $(c_j)_{j \leq \alpha_i}$ de scalaires tels que :

$$\frac{P}{(X-a)^\alpha} = E + \sum_{j=1}^{\alpha} \frac{c_j}{(X-a)^j}.$$

Démonstration

On a déjà établi l'existence et l'unicité de la partie entière ; il reste à discuter le cas où le polynôme P est de degré strictement inférieur à α.

Décomposons le polynôme au numérateur avec la formule de Taylor en a : $P = \sum_{k=0}^{\alpha-1} a_k (X-a)^k$; par conséquent

$$\frac{P}{(X-a)^\alpha} = \sum_{k=0}^{\alpha-1} \frac{a_k}{(X-a)^{\alpha-k}}.$$

L'unicité découle de l'unicité des coefficients de $P(X+a)$. ■

Remarque

Rappelons que l'on peut obtenir cette décomposition en éléments simples par divisions successives du numérateur (et des quotients qui apparaissent) par $X-a$.

Revenons au cas général d'un dénominateur scindé.

Proposition 13.14.

Soit $P \in \mathbb{K}[X]$, $a_1, \ldots, a_q \in \mathbb{K}$ et $\alpha_1, \ldots, \alpha_q \in \mathbb{N} \setminus \{0\}$.
Il existe un unique $E \in \mathbb{K}[X]$ et une unique famille $(c_{i,j})_{i \leq q, j \leq \alpha_i}$ de scalaires tels que :

$$\frac{P}{\prod_{i=1}^{q} (X-a_i)^{\alpha_i}} = E + \sum_{i=1}^{q} \sum_{j=1}^{\alpha_i} \frac{c_{i,j}}{(X-a_i)^j}.$$

Démonstration
La démonstration s'appuie sur le théorème de Bézout pour la famille des polynômes premiers entre eux composée des $\prod_{\substack{j=1\\j\neq i}}^{q}(X-a_i)^{\alpha_i}$ pour $i\in[\![1,q]\!]$. En effet, il existe des polynômes $U_1,\dots,U_q$ tels que

$$1=\sum_{i=1}^{q}U_i\prod_{\substack{j=1\\j\neq i}}^{q}(X-a_i)^{\alpha_i}.$$

En multipliant par P puis en divisant par le dénominateur, on obtient

$$\frac{P}{\prod_{i=1}^{q}(X-a_i)^{\alpha_i}}\sum_{i=1}^{q}\frac{PU_i}{(X-a_i)^{\alpha_i}}.$$

L'existence suit immédiatement avec le cas particulier précédemment traité. ■

Donnons un exemple pratique de décomposition quelquefois appelé « dérivée logarithmique ».

Proposition 13.15.

Soit $P=\prod_{i=1}^{n}(X-a_i)^{\alpha_i}$ un polynôme scindé avec les racines $a_1,\dots,a_n$ deux à deux distincts. Alors,

$$\frac{P'}{P}=\sum_{i=1}^{n}\frac{\alpha_i}{X-a_i}.$$

Démonstration
La formule de dérivation donne directement

$$P'=\sum_{i=1}^{n}\alpha_i(X-a_i)^{\alpha_i-1}\prod_{j\neq i}(X-a_j)^{\alpha_j}$$

et le résultat suit immédiatement en divisant par P. ■

Remarque
On peut mémoriser ce résultat en le voyant comme une forme de « dérivée logarithmique » (même si cela n'a pas de sens réel puisque nous serions bien en peine de définir le logarithme d'un polynôme à coefficients dans un corps $\mathbb{K}$).

2.2. Obtention pratique

Nous disposons maintenant de la forme de la décomposition en éléments simples ; nous savons déterminer la partie entière par une division euclidienne et il reste désormais à déterminer les différents autres coefficients. Dressons une petite liste (non-exhaustive) de méthodes pour trouver les coefficients ou des relations entre ces coefficients.

▹ On calcule le coefficient qui correspond au terme de plus bas degré pour un pôle a (le seul coefficient si a est un pôle simple) en multipliant par la plus grande puissance de $(X-a)$ puis en évaluant en a.

Exemple

Décomposons la fraction $\frac{1}{(X-2)(X-1)}$. D'après le résultat général, la décomposition en éléments simples est de la forme

$$\frac{1}{(X-2)(X-1)} = \frac{a}{X-2} + \frac{b}{X-1}.$$

En multipliant par $X-2$, on obtient

$$\frac{1}{X-1} = a + b\frac{X-2}{X-1},$$

et l'évaluation en 2 nous donne $a = 1$. De même, $b = -1$.

Passons au cas général.

Proposition 13.16.

Soit a un pôle de multiplicité r de la fraction irréductible $\frac{P}{Q}$.

▹ Si $Q(X) = (X-a)^r R(X)$, alors le coefficient de $\frac{1}{(X-a)^r}$ dans la décomposition en éléments simples est

$$\frac{P(a)}{R(a)}.$$

▹ Le coefficient de $\frac{1}{(X-a)^r}$ dans la décomposition en éléments simples est

$$r!\frac{P(a)}{Q^{(r)}(a)}.$$

Remarque

Ces deux résultats sont de même nature ; on utilise le premier lorsque le dénominateur est déjà factorisé, le second sinon (les dérivées sont plus faciles à calculer que la factorisation).

Déduisons la seconde formule de la première.

Démonstration

En effet, si on écrit $Q = (X-a)^r R$, alors, d'après la formule de Leibniz,

$$Q^{(r)} = \sum_{k=0}^{r} \binom{r}{k} \frac{r!}{(r-k)!}(X-a)^{r-k} R^{(r-k)}.$$

En évaluant en a, il ne reste de la somme que le terme correspondant à $k = r$ donc $Q^{(r)}(a) = r!R(a)$. ■

Exemple

Décomposons la fraction $\frac{1}{X^n-1}$. Elle est de degré strictement négatif donc sa partie entière est nulle. Ses pôles sont les racines $n^{\text{ième}}$ de l'unité et sont simples. D'après la formule précédente, le coefficient associé au pôle $\omega \in \mathbb{U}^n$ est $\frac{1}{n\omega^{n-1}} = \frac{1}{n}\omega$. Ainsi,

$$\frac{1}{X^n-1} = \frac{1}{n}\sum_{\omega \in \mathbb{U}_n} \frac{\omega}{X-\omega}.$$

Exemple

Factorisons la fraction rationnelle dont la décomposition en éléments simples est

$$\sum_{\omega\in\mathbb{U}_n} \frac{\omega^2}{X-\omega}.$$

Le plus petit dénominateur commun est alors $Q = X^n - 1$. Notons P le numérateur associé; $\deg P \leq n-1$ car la fraction considérée est de degré strictement négatif. D'après la formule précédente,

$$\forall \omega \in \mathbb{U}_n, \qquad \frac{P(\omega)}{Q'(\omega)} = \omega^2.$$

Par ailleurs, pour tout $\omega \in \mathbb{U}_n$, $Q'(\omega) = n\omega^{n-1} = \frac{n}{\omega}$ donc $P(\omega) = n\omega$. Ainsi, le polynôme $P - nX$ admet n racines distinctes et son degré est au plus $n-1$ donc est nul. En conclusion, $P = nX$ et

$$\sum_{\omega\in\mathbb{U}_n} \frac{\omega^2}{X-\omega} = \frac{nX}{X^n-1}.$$

Pour déterminer l'ensemble des coefficients on peut aussi chercher des relations simples entre les coefficients inconnus avec les méthodes suivantes.

▹ On utilise les limites.

Exemple

Décomposons la fraction $\frac{X^3}{(X-2)^2(X-1)^2}$. D'après le résultat général, la décomposition en éléments simples est de la forme

$$\frac{a}{(X-2)^2} + \frac{b}{(X-2)} + \frac{c}{(X-1)^2} + \frac{d}{X-1}.$$

La méthode précédente donne $a = 8$ et $c = 1$. Si l'on multiplie par X et que l'on fait tendre x vers $+\infty$ dans la fonction rationnelle correspondante, on obtient une première relation $b + d = 1$.

▹ On évalue en des valeurs particulières (simples) différentes des pôles.

Exemple

Avec la fraction précédente, l'évaluation en 0 donne :

$$0 = \frac{a}{4} - \frac{b}{2} + c - d,$$

ce qui, compte-tenu des informations dont nous disposons déjà donne $b = -4$ et $d = 5$ soit au final :

$$\frac{X^3}{(X-2)^2(X-1)^2} = \frac{8}{(X-2)^2} + \frac{-4}{(X-2)} + \frac{1}{(X-1)^2} + \frac{5}{X-1}.$$

▹ On utilise la parité éventuelle de la fraction.

Exemple

Décomposons la fraction $\frac{X^2}{(X^2-1)^2}$ de degré strictement négatif. D'après le résultat général, la

décomposition en éléments simples est de la forme :

$$\frac{a}{(X-1)^2}+\frac{b}{X-1}+\frac{c}{(X+1)^2}+\frac{d}{X+1}.$$

Or, en changeant X en $-X$, on ne change pas la fraction rationnelle (qui est paire) mais on intervertit les coefficients a et c et b et $-d$. Par unicité de la décomposition, on obtient $a=c$ et $b=-d$. Or, d'après la première méthode, $a=c=\frac{1}{4}$ et en évaluant en 0, on obtient $a-b+c+d=0$ d'où $b=\frac{1}{4}$ et $d=-\frac{1}{4}$.

On peut bien entendu adapter cette méthode à des symétries comme l'invariance lorsque l'on remplace X par ωX où ω est une racine $n^{\text{ième}}$ de l'unité.

2.3. Quelques applications élémentaires

Donnons quelques applications de ces décompositions en éléments simples.

Exemple

Calculons $\sum_{k=1}^{n}\frac{1}{k(k+1)}$. Pour cela remarquons la décomposition en éléments simples suivante :

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

Ainsi,

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

Le dernier résultat étant obtenu par télescopage.

Exemple

Calculons la dérivée $n^{\text{ième}}$ de $f:x\mapsto\frac{1}{x^2-3x+2}$.
Commençons par décomposer la fraction rationnelle correspondante en éléments simples :

$$\frac{1}{X^2-3X+2}=\frac{1}{X-2}-\frac{1}{X-1}.$$

Il suffit désormais d'utiliser que la dérivée $n^{\text{ième}}$ de $x\mapsto\frac{1}{x-a}$ est $x\mapsto\frac{(-1)^n n!}{(x-a)^{n+1}}$ pour obtenir que :

$$f^{(n)}(x)=(-1)^n n!\left(\frac{1}{(x-2)^{n+1}}-\frac{1}{(x-1)^{n+1}}\right).$$

3. Décomposition en éléments simples (cas général)

On considère ici le cas général où le dénominateur n'est plus forcément scindé. On peut alors adapter la démonstration du cas précédent pour obtenir la proposition suivante.

Proposition 13.17.

Soit $P \in \mathbb{R}[X]$, $Q_1, \dots, Q_q$ des polynômes unitaires irréductibles deux à deux distincts et des entiers $\alpha_1, \dots, \alpha_q \in \mathbb{N} \setminus \{0\}$.
Il existe un unique $E \in \mathbb{R}[X]$ et une unique famille $(R_{i,j})_{i \leq q, j \leq \alpha_i}$ de polynômes tels que

$$\forall i \in [\![1, q]\!],\ \forall j \in [\![1, \alpha_i]\!], \qquad \deg R_{i,j} < \deg Q_i,$$

et

$$\frac{P}{\prod_{i=1}^{q} Q_i^{\alpha_i}} = E + \sum_{i=1}^{q} \sum_{j=1}^{\alpha_i} \frac{R_{i,j}}{Q_i^j}.$$

Remarque

Ainsi, dans $\mathbb{R}(X)$, chacun des éléments simples est de la forme $\frac{c}{(X-a)^\alpha}$ ou $\frac{cX+d}{(X^2+aX+b)^\alpha}$ avec $X^2 + aX + b$ à discriminant strictement négatif.

Exemple

Donnons la forme de quelques décompositions en éléments simples.
▹ Comme $X^3 - 1 = (X-1)(X^2+X+1)$,

$$\frac{1}{(X^3-1)^2} = \frac{a}{(X-1)^2} + \frac{b}{X-1} + \frac{cX+d}{(X^2+X+1)^2} + \frac{eX+f}{X^2+X+1}.$$

▹ Comme $X^4 + 1 = (X^2 - \sqrt{2}X + 1)(X^2 + \sqrt{2}X + 1)$,

$$\frac{1}{(X^4+1)^2} = \frac{aX+b}{(X^2-\sqrt{2}X+1)^2} + \frac{cX+d}{X^2-\sqrt{2}X+1} + \frac{eX+f}{(X^2+\sqrt{2}X+1)^2} + \frac{gX+h}{X^2+\sqrt{2}X+1}.$$

Exemple

Décomposons en éléments simples la fraction $\frac{X^2-2X-1}{(X^2+1)^2}$. D'après le résultat général, la décomposition en éléments simples est de la forme :

$$\frac{aX+b}{(X^2+1)^2} + \frac{cX+d}{X^2+1}.$$

Commençons par multiplier par $(X^2+1)^2$ puis par évaluer en i. On obtient que $ai + b = -2 - 2i$ d'où $a = -2$ et $b = -2$.
En multipliant par X et en étudiant la limite en $+\infty$, on trouve $c = 0$.
En évaluant en 0, on obtient $-1 = b + d$ d'où $d = 1$.
En conclusion,

$$\frac{X^2-2X-1}{(X^2+1)^2} = -\frac{2X+2}{(X^2+1)^2} + \frac{1}{X^2+1}.$$

Exemple

Nous avons déjà montré que la fraction rationnelle $\frac{1}{X^n-1}$ admettait la décomposition en éléments simples dans $\mathbb{C}(X)$ suivante :

$$\frac{1}{X^n-1}=\frac{1}{n}\sum_{k=0}^{n-1}\frac{e^{\frac{2ik\pi}{n}}}{X-e^{\frac{2ik\pi}{n}}}.$$

Rassemblons les termes deux à deux conjugués pour en déduire la décomposition en éléments simples dans $\mathbb{R}(X)$.

▹ Si $n=2p$, alors il y a deux pôles réels simples et $p-1$ paires de pôles conjugués.

$$\begin{aligned}\frac{1}{X^n-1}&=\frac{1}{n}\frac{1}{X-1}+\frac{1}{n}\frac{-1}{X+1}+\frac{1}{n}\sum_{k=1}^{p-1}\left(\frac{e^{\frac{2ik\pi}{n}}}{X-e^{\frac{2ik\pi}{n}}}+\frac{e^{\frac{-2ik\pi}{n}}}{X-e^{\frac{-2ik\pi}{n}}}\right)\\&=\frac{1}{n}\frac{1}{X-1}+\frac{1}{n}\frac{-1}{X+1}+\frac{1}{n}\sum_{k=1}^{p-1}\frac{2\cos(\frac{2k\pi}{n})X-2}{X^2-2\cos(\frac{2k\pi}{n})X+1}.\end{aligned}$$

▹ Si $n=2p+1$, alors il y a un pôle réel simple et p paires de pôles conjugués.

$$\begin{aligned}\frac{1}{X^n-1}&=\frac{1}{n}\frac{1}{X-1}+\frac{1}{n}\sum_{k=1}^{p}\left(\frac{e^{\frac{2ik\pi}{n}}}{X-e^{\frac{2ik\pi}{n}}}+\frac{e^{\frac{-2ik\pi}{n}}}{X-e^{\frac{-2ik\pi}{n}}}\right)\\&=\frac{1}{n}\frac{1}{X-1}+\frac{1}{n}\sum_{k=1}^{p}\frac{2\cos(\frac{2k\pi}{n})X-2}{X^2-2\cos(\frac{2k\pi}{n})X+1}.\end{aligned}$$

Exemple

Décomposons en éléments simples dans $\mathbb{C}(X)$ la fraction $\frac{X}{(X^2+X+1)^2}$. D'après le résultat général, la décomposition en éléments simples est de la forme :

$$\frac{a}{(X-j)^2}+\frac{b}{(X-j)}+\frac{c}{(X-\overline{j})^2}+\frac{d}{(X-\overline{j})}.$$

En conjuguant et en utilisant l'unicité de la décomposition, on obtient $\overline{a}=c$ et $\overline{b}=d$.
On calcule $a=-\frac{1}{3}j$ en multipliant par $(X-j)^2$ et en évaluant en j.
En multipliant par X et en étudiant la limite en $+\infty$, on trouve $b=-d$ puis en évaluant en -1, $d=-\frac{1}{3j(j-1)}$ donc :

$$\frac{X}{(X^2+X+1)^2}=\frac{-\frac{1}{3}j}{(X-j)^2}+\frac{\frac{1}{3j(j-1)}}{(X-j)}+\frac{-\frac{1}{3}\overline{j}}{(X-\overline{j})^2}+\frac{-\frac{1}{3j(j-1)}}{(X-\overline{j})}.$$

Or, le premier terme est déjà décomposé en éléments simples dans $\mathbb{R}(X)$.

FICHE

SYNTHÈSE

Les fractions rationnelles sont essentiellement les quotients de deux polynômes. La décomposition en éléments simples permet de manipuler simplement les fractions rationnelles. Voici l'énoncé dans le cas d'un dénominateur scindé.

Décomposition des fractions rationnelles

Soit $P \in \mathbb{K}[X]$, $a_1, \dots, a_q \in \mathbb{K}$ et $\alpha_1, \dots, \alpha_q \in \mathbb{N} \setminus \{0\}$. Il existe un unique $E \in \mathbb{K}[X]$ et une unique famille $(c_{i,j})_{i \leq q, j \leq \alpha_i}$ de scalaires tels que

$$\frac{P}{\prod_{i=1}^{q}(X-a_i)^{\alpha_i}} = E + \sum_{i=1}^{q}\sum_{j=1}^{\alpha_i} \frac{c_{i,j}}{(X-a_i)^j}.$$

Il est important de connaître la forme de la décomposition en éléments simples et de connaître des méthodes simples de calcul de la partie entière et des coefficients.

Décomposition des fractions rationnelles

Soit $P \in \mathbb{R}[X]$, $Q_1, \dots, Q_q$ des polynômes unitaires irréductibles deux à deux distincts et $\alpha_1, \dots, \alpha_q \in \mathbb{N} \setminus \{0\}$.
Il existe un unique $E \in \mathbb{R}[X]$ et une unique famille $(R_{i,j})_{i \leq q, j \leq \alpha_i}$ de polynômes tels que

$$\forall i \in [\![1, q]\!],\ \forall j \in [\![1, \alpha_i]\!], \qquad \deg R_{i,j} < \deg Q_i,$$

et

$$\frac{P}{\prod_{i=1}^{q} Q_i^{\alpha_i}} = E + \sum_{i=1}^{q}\sum_{j=1}^{\alpha_i} \frac{R_{i,j}}{Q_i^j}.$$

Parmi les exemples classiques qu'il convient de retenir, celui de la « dérivée logarithmique » est à part.

Soit $P = \prod_{i=1}^{n}(X-a_i)^{\alpha_i}$ un polynôme scindé avec les racines $a_1, \dots, a_n$ deux à deux distincts. Alors,

$$\frac{P'}{P} = \sum_{i=1}^{n} \frac{\alpha_i}{X-a_i}.$$

EXERCICES

Vrai ou faux ?

		Vrai	Faux
a)	Si une fraction R est de degré $n \in \mathbb{Z}$, alors le degré de $R\left(\frac{1}{X}\right)$ est $-n$.	☐	☐
b)	Si une fraction R est de degré $n \in \mathbb{Z}$, alors le degré de $\frac{1}{R}$ est $-n$.	☐	☐
c)	Une fraction de degré positif est un polynôme.	☐	☐
d)	Un pôle de $\frac{P}{Q}$ est une racine de P.	☐	☐
e)	Les fractions de degré strictement négatif sont les fractions de partie entière nulle.	☐	☐
f)	La partie entière de la fraction $\frac{P}{Q}$ est de degré $\deg P - \deg Q$.	☐	☐
g)	$\frac{1}{X^2+2X-3}$ est un élément simple dans $\mathbb{R}(X)$.	☐	☐
h)	$\frac{1}{X^4+1}$ est un élément simple dans $\mathbb{R}(X)$.	☐	☐
i)	$\frac{X+1}{(X^2-X+1)^2}$ est un élément simple dans $\mathbb{R}(X)$.	☐	☐
j)	$\frac{X+1}{(X-2)^2}$ est un élément simple dans $\mathbb{R}(X)$.	☐	☐
k)	$\frac{X^2(X^2+2)}{X^4-1} = \frac{3/4}{X-1} - \frac{3/4}{X+1} + \frac{1/2}{X^2+1}$.	☐	☐
l)	Si P et Q sont des polynômes tels qu'il existe une décomposition $\frac{P}{Q} = 1 + \sum_{k=1}^{n} \frac{a_k}{X-a_k}$, alors $\deg P = \deg Q \leq n$.	☐	☐

Exercices d'application du chapitre 13

●○○ **Exercice 1**

Décomposer en éléments simples les fractions rationnelles suivantes dans $\mathbb{C}(X)$

1. $\frac{X^4+1}{X^4-1}$
2. $\frac{X^2+1}{X^2(X+1)^2}$

●○○ **Exercice 2**

Décomposer $\frac{1}{X(X-1)\ldots(X-n)}$ en éléments simples.

●○○ **Exercice 3**

Déterminer les réels a, b et c tels que la fraction $\frac{aX^2+bX+c}{(X^2-1)^2}$ n'admet pas de terme de degré -1 dans sa décomposition en éléments simples.

●●○ **Exercice 4**

1. Déterminer un développement limité en 0 à l'ordre 7 de $x \mapsto \frac{x^3+1}{x^2+1}$.
2. En déduire la décomposition en éléments simples dans $\mathbb{R}[X]$ la fraction rationnelle $\frac{X^3+1}{X^8(X^2+1)}$.

●●○ **Exercice 5**

Soit $\omega = e^{\frac{2i\pi}{n}}$ et $P(X) = \sum_{k=0}^{n-1} X^k$. Déterminer la décomposition en éléments simples de la fraction $\frac{P'}{P}$ puis en déduire la valeur de la somme $\sum_{k=1}^{n-1} \frac{1}{1-\omega^k}$.

●●○ **Exercice 6**

1. Montrer que, pour tout $n \in \mathbb{N}$, il existe un polynôme $P_n \in \mathbb{R}[X]$ tel que
$$\forall \theta \in \mathbb{R}, \qquad \sin((2n+1)\theta) = P_n(\sin(\theta)).$$
2. Décomposer en éléments simples la fraction rationnelle $\frac{1}{P_n}$.

●●○ **Exercice 7**

Soit P, Q deux polynômes à coefficients réels tels que Q est de degré $n > 1$, admet n racines réelles distinctes notées $x_1, \dots, x_n$ et que P est de degré strictement inférieur à n.

1. Justifier que $Q'(x_k) \neq 0$ pour tout $k \leq n$.
2. Montrer que $\frac{P(x)}{Q(x)} =_{k=1}^{n} \frac{P(x_k)}{Q'(x_k)(x-x_k)}$.
3. a. En déduire que, si aucune des racines x_k n'est nulle, alors $\sum_{k=1}^{n} \frac{1}{x_k Q'(x_k)} = -\frac{1}{Q(0)}$.
   b. Calculer $\sum_{k=1}^{n} \frac{1}{Q'(x_k)}$.

Exercices d'approfondissement du chapitre 13

●●○ **Exercice A**

Soit $a_1, a_2, \dots, a_n \in \mathbb{C}$ deux à deux distincts, $\alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{C}$ eux aussi deux à deux distincts, tels que
$$\forall i, j \in [\![1, n]\!], \qquad a_i + \alpha_j \neq 0.$$
Considérons $x_1, x_2, \dots, x_n$ tels que
$$\forall i \in [\![1, n]\!], \qquad \sum_{j=1}^{n} \frac{x_j}{a_j + \alpha_i} = 1.$$
Posons $R(X) = 1 - \sum_{j=1}^{n} \frac{x_j}{X+a_j}$.

1. Écrivons R sous la forme
$$\frac{P(X)}{\prod_{j=1}^{n}(X+a_j)}.$$
   a. Majorer $\deg P$.
   b. Calculer les expressions possibles pour le polynôme P.
2. En déduire une expression de $x_1, x_2, \dots, x_n$.

EXERCICES

Exercice B Théorème de Gauss-Lucas

Définition 13.18.

Soit $z_1, \dots, z_n \in \mathbb{C}$. L'enveloppe convexe de ces complexes est l'ensemble des

$$\sum_{k=1}^{n} \lambda_k z_k$$

avec $\lambda_1, \dots, \lambda_n \in \mathbb{R}_+$ de somme 1.

1. Soit $P \in \mathbb{C}[X]$ non constant. Montrer que les racines de P' appartiennent à l'enveloppe convexe des racines de P.
2. Soit $P \in \mathbb{C}[X]$ un polynôme non constant et H est un demi-plan de $\mathbb{C}$ tels que $0 \in P'(H)$ (c'est-à-dire que P' admet une racine dans H).
   Montrer que $P(H) = \mathbb{C}$, c'est-à-dire que la fonction polynomiale associée à P est surjective de H dans $\mathbb{C}$.

Problèmes du chapitre 13

Problème 1

Dans tout ce problème, on admet, pour tout $k \in \mathbb{N}$, l'existence d'un unique polynôme T_k de degré k tel que

$$\forall \theta \in \mathbb{R}, \qquad T_k(\cos\theta) = \cos(k\theta).$$

Fixons $n \in \mathbb{N} \setminus \{0\}$.

Partie A – Une décomposition en élément simple

1. a. Montrer que, pour tout $\theta \in \mathbb{R}$, $T_{2n}(\cos\theta) = T_{2n}(-\cos\theta)$.
   b. En déduire que T_{2n} est pair.

   Posons, pour tout $k \in [\![0, 2n-1]\!]$, $\theta_k = \frac{2k+1}{4n}\pi$.
2. a. Justifier que les $\cos\theta_k$ pour $k \in [\![0, 2n-1]\!]$ sont deux à deux distincts.
   b. En déduire les racines de T_{2n} et préciser leur multiplicité.
3. Soit $P \in \mathbb{R}_{2n-1}[X]$. Déterminer la décomposition en éléments simples de la fraction $\frac{P}{T_{2n}}$ en fonctions des valeurs $P(\cos\theta_k)$.
4. Montrer que, pour tout $m \leq n$,

$$\frac{T'_{2m}(\cos\theta)}{T_{2n}(\cos\theta)} = \sum_{k=0}^{2n-1} (-1)^k \frac{m}{n} \frac{\sin 2m\theta_k \cos\theta}{\cos^2\theta - \cos^2\theta_k}.$$

5. En déduire que, pour tout $m \leq n$,

$$\sum_{k=0}^{2n-1} (-1)^k \frac{\sin 2m\theta_k}{\sin^2\theta_k} = 4mn.$$

Partie B – Une identité de Riesz

On cherche à établir l'identité $(\star)$ suivante : pour tout $P \in \mathbb{R}_n[X]$, la fonction $\Phi : \theta \mapsto P(e^{i\theta})$ vérifie

$$\Phi'(\theta) = \frac{1}{2n} \sum_{k=0}^{2n-1} (-1)^k \frac{\Phi(\theta + 2\theta_k)}{2\sin^2 \theta_k}.$$

1. Justifier qu'il suffit d'établir l'identité $(\star)$ pour les polynômes $P = X^m$ pour tout $m \leq n$.
2. Montrer que l'identité $(\star)$ pour $P = X^m$ est équivalente à
$$\sum_{k=0}^{2n-1} (-1)^k \frac{e^{i2m\theta_k}}{\sin^2 \theta_k} = 4imn.$$
3. Calculer, pour tout $k \in [\![0, 2n-1]\!]$, $\theta_{2n-1-k} + \theta_k$.
4. Conclure.

●●○ Problème 2

Les différentes parties de ce problème sont essentiellement indépendantes. La partie A demande de redémontrer des résultats du cours.

Partie A – Interpolation de Lagrange

Soit $n \geq 1$, $a_1, \ldots, a_n \in \mathbb{C}$ deux à deux distincts et $b_1, \ldots, b_n \in \mathbb{C}$. Un polynôme d'interpolation associé à ces points est un polynôme $P \in \mathbb{C}_{n-1}[X]$ tel que

$$\forall k \in [\![1, n]\!], \qquad P(a_k) = b_k.$$

1. Montrer l'unicité du polynôme d'interpolation.
2. a. Soit $j \in [\![1, n]\!]$. Montrer l'existence du polynôme d'interpolation L_j dans le cas particulier
$$\forall i \in [\![1, n]\!], \qquad b_i = \delta_{i,j}.$$
   b. Montrer l'existence du polynôme interpolateur dans le cas général.

Partie B – Interpolation de Cauchy

Soit $n \geq 1$, $p \in [\![0, n]\!]$, $a_1, \ldots, a_n \in \mathbb{C}$ deux à deux distincts et $b_1, \ldots, b_n \in \mathbb{C}$. Une fraction rationnelle d'interpolation d'ordre p associée à ces points est une fraction $\frac{P}{Q} \in \mathbb{C}(X)$ telle que

$$\forall k \in [\![1, n]\!], \qquad \frac{P(a_k)}{Q(a_k)} = b_k$$

et

$$\deg P < p, \qquad \deg Q \leq n - p.$$

Notons $M = \prod_{k=1}^{n} (X - a_k)$ et L le polynôme d'interpolation associé à ces valeurs d'après la partie A.

1. Montrer l'existence de la fraction rationnelle d'interpolation d'ordre $p = n$.
2. Supposons, dans cette question, $b_k \neq 0$ pour tout $k \in [\![1, n]\!]$.
   a. Justifier que $M \wedge L = 1$.
   Posons $R_0 = M$, $R_1 = L$, $U_0 = 0$, $U_1 = 1$, $V_0 = 1$, $V_1 = 0$ puis, pour tout i, Q_{i+2} le quotient dans la division euclidienne de R_i par R_{i+1} et
$$\begin{aligned} R_{i+2} &= R_i - Q_{i+2} R_{i+1} \\ U_{i+2} &= U_i - Q_{i+2} U_{i+1} \\ V_{i+2} &= V_i - Q_{i+2} V_{i+1} \end{aligned}$$
   b. Justifier qu'il existe j tel que $\deg R_j \leq p$.
   On suppose dorénavant j minimal parmi les entiers vérifiant cette propriété.

PROBLÈMES

c. Montrer que, pour tout i, $R_i = U_i L + V_i M$.
d. Montrer que, pour tout $i \in [\![1, j-1]\!]$,

$$\deg U_{i+1} = \sum_{k=2}^{i+1} \deg Q_k.$$

e. En déduire que $\deg U_j \leq n-p$.
f. Conclure.

3. Montrer qu'il existe des paramètres d'interpolation pour lesquels il n'existe pas de fraction rationnelle d'interpolation.

Partie C – Interpolation de Hermite

Soit $n \geq 1$, $\alpha_1, \ldots, \alpha_n \in \mathbb{N} \setminus \{0\}$, $a_1, \ldots, a_n \in \mathbb{C}$ deux à deux distincts et des complexes $(b_{k,l})_{k \leq n, 0 \leq l < \alpha_k}$. Un polynôme d'interpolation associé à ces paramètres est un polynôme $P \in \mathbb{C}[X]$ tel que

$$\forall k \in [\![1, n]\!],\ \forall l < \alpha_k, \qquad P^{(l)}(a_k) = b_{k,l}$$

et

$$\deg P < \sum_{k=1}^{n} \alpha_k.$$

1. Montrer l'unicité du polynôme d'interpolation associé à ces paramètres.
2. Notons $L_{k,l}$ la partie régulière du développement limité en a_k à l'ordre $\alpha_k - l$ de la fonction

$$x \mapsto \prod_{\substack{j=1 \\ j \neq k}}^{n} \frac{1}{(x-a_j)^{\alpha_j}}.$$

Posons enfin

$$H_{k,l}(X) = \frac{1}{l!} L_{k,l}(X)(X-a_k)^l \prod_{\substack{j=1 \\ j \neq k}}^{n} (X-a_j)^{\alpha_j}.$$

a. Calculer, pour tout $i \neq k$ et $j < \alpha_i$, $H_{k,l}^{(j)}(a_i)$.
b. Calculer, pour tout $j < \alpha_k$, $H_{k,l}^{(j)}(a_k)$.

3. Montrer l'existence du polynôme d'interpolation associé à ces paramètres.
4. Déterminer le polynôme d'interpolation pour les paramètres

$$n = 2,\ a_1 = 0,\ \alpha_1 = 2,\ a_2 = 1,\ \alpha_2 = 1,\ b_{1,0} = 1,\ b_{1,1} = 0,\ b_{2,0} = 3.$$

CORRIGÉS

Corrigés des Vrai/Faux

a) Faux, par exemple $\frac{1}{X+1}$.

b) Vrai

c) Faux, par exemple $\frac{X^2}{X+1}$.

d) Faux, rien n'assure que ce représentant est irréductible.

e) Vrai.

f) Faux quand la quantité $\deg P - \deg Q$ est strictement négative.

g) Faux car le dénominateur admet une racine réelle donc n'est pas irréductible.

h) Faux car le dénominateur est de degré 4 donc n'est pas irréductible.

i) Vrai.

j) Faux le degré du numérateur doit être strictement inférieur au degré du polynôme irréductible utilisé au dénominateur (ici, $X-2$).

k) Faux, la partie entière est oubliée.

l) Faux pour la dernière inégalité, car $\frac{X^k P}{X^k Q} = \frac{P}{Q}$.

Corrigés des exercices d'application du chapitre 13

Exercice 1

1. Cette fraction est de degré 0 et la partie entière est 1. On a donc

$$\frac{X^4+1}{X^4-1} = 1 + \frac{2}{X^4-1}.$$

Les pôles sont simples et valent 1, i, -1 et $-i$. L'unique coefficient associé à un pôle z est $\frac{2}{4z^3} = \frac{1}{2z^3} = \frac{z}{2}$ (car $z \in \mathbb{U}_4$). On en déduit immédiatement la décomposition en éléments simples

$$\frac{X^4+1}{X^4-1} = 1 + \frac{\frac{1}{2}}{X-1} + \frac{\frac{i}{2}}{X-i} + \frac{\frac{-1}{2}}{X+1} + \frac{\frac{-i}{2}}{X+i}$$

2. Cette fraction est de degré strictement négatif donc la partie entière est nulle. Les pôles sont 0 et -1 avec multiplicité 2. La décomposition en éléments simples est donc de la forme

$$\frac{X^2+1}{X^2(X+1)^2} = \frac{a}{X^2} + \frac{b}{X} + \frac{c}{(X+1)^2} + \frac{d}{X+1},$$

avec a, b, c, $d \in \mathbb{R}$. En multipliant par X^2 et en évaluant en 0, on obtient que $a = 1$; en multipliant par $(X+1)^2$ et en évaluant en -1, que $c = 2$. En multipliant par X et en étudiant la limite en $+\infty$ que $b+d = 0$. Enfin, en évaluant en 1,

$$\frac{1}{2} = a + b + \frac{c}{4} + \frac{d}{2}.$$

Ainsi, $b = -2$ et $d = 2$ soit

$$\frac{X^2+1}{X^2(X+1)^2} = \frac{1}{X^2} + \frac{-2}{X} + \frac{2}{(X+1)^2} + \frac{2}{X+1}.$$

Exercice 2

Avec les formules du cours la décomposition en éléments simples est

$$\frac{1}{X(X-1)\dots(X-n)} = \sum_{k=0}^{n} \frac{a_k}{X-k},$$

avec pour tout $k \in [\![0, n]\!]$,

$$a_k = \frac{1}{\prod_{j=0}^{k-1}(k-j) \prod_{j=k+1}^{n}(k-j)} = \frac{(-1)^{n-k}}{k!(n-k)!}.$$

Exercice 3

Commençons par écrire la forme de la décomposition en éléments simples de cette fraction : comme le degré est strictement négatif et que le dénominateur est scindé, il existe des réels α , β, γ et δ tels que

$$\frac{aX^2+bX+c}{(X^2-1)^2} = \frac{\alpha}{(X-1)^2} + \frac{\beta}{X-1} + \frac{\gamma}{(X+1)^2} + \frac{\delta}{X+1}.$$

Quitte à multiplier par $(X-1)^2$ puis à évaluer en 1, on obtient $\alpha = \frac{1}{4}(a+b+c)$. De même, on trouve $\gamma = \frac{1}{4}(a-b+c)$.

Revenons à l'énoncé. La fraction considérée n'admet pas de terme de degré -1 dans sa décomposition si, et seulement si,

$$\frac{aX^2+bX+c}{(X^2-1)^2} = \frac{\alpha}{(X-1)^2} + \frac{\gamma}{(X+1)^2}$$

c'est-à-dire en mettant au même dénominateur

$$\begin{aligned} aX^2+bX+c &= \alpha(X+1)^2 + \gamma(X-1)^2 \\ &= \frac{1}{2}(a+c)X^2 + bX + \frac{1}{2}(a+c). \end{aligned}$$

Cette dernière égalité est vérifiée si, et seulement si, $a = c$.

Exercice 4

1. Calculons ce développement limité

$$\begin{aligned} \frac{x^3+1}{x^2+1} &\underset{0}{=} (1+x^3)(1-x^2+x^4-x^6+o(x^7)) \\ &\underset{0}{=} 1-x^2+x^3+x^4-x^5-x^6+x^7+o(x^7). \end{aligned}$$

2. Passons au calcul de la décomposition en éléments simples qui est de la forme

$$\frac{X^3+1}{X^8(X^2+1)} = \frac{aX+b}{X^2+1} + \sum_{k=1}^{8} \frac{a_k}{X^k}.$$

En multipliant par X^2+1 et en évaluant en i, on obtient que $a=-1$ et $b=1$. Ensuite, en multipliant par X^8, on obtient l'identité entre fonctions

$$\begin{aligned}\frac{x^3+1}{x^2+1} &= \frac{-x+1}{x^2+1}x^8 + \sum_{k=1}^{8} a_k x^{8-k}\\ &\underset{0}{=} o(x^7) + \sum_{k=1}^{8} a_k x^{8-k}.\end{aligned}$$

On identifie les coefficients de ce développement limité avec le calcul précédent et on en déduit

$$\frac{X^3+1}{X^8(X^2+1)} = \frac{-X+1}{X^2+1} + \frac{1}{X^8} - \frac{1}{X^6} + \frac{1}{X^5} + \frac{1}{X^4} - \frac{1}{X^3} - \frac{1}{X^2} + \frac{1}{X}.$$

Exercice 5

Remarquons que les racines de P sont simples et qu'elles sont les racines $n^{\text{ièmes}}$ de l'unité différentes de 1, à savoir les ω^k pour $\in [\![1,n]\!]$. Ainsi, la décomposition en éléments simples est $\frac{P'}{P} = \sum_{k=1}^{n-1} \frac{1}{X-\omega^k}$. En évaluant en 1, on obtient

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

Exercice 6

1. L'unicité est élémentaire car si deux polynômes P_n et Q_n vérifient cette propriété, alors $P_n - Q_n$ admet tous les réels de la forme $\sin\theta$ comme racines donc est nul.
Pour l'existence, utilisons la formule de Moivre. Pour tout $\theta \in \mathbb{R}$,

$$\begin{aligned}\sin(2n+1)\theta &= \mathfrak{Im}(e^{(2n+1)i\theta}) = \mathfrak{Im}((\cos\theta + i\sin\theta)^{2n+1})\\ &= \mathfrak{Im}\left(\sum_{k=0}^{2n+1}\binom{2n+1}{k}(i\sin\theta)^k(\cos\theta)^{2n+1-k}\right)\\ &= \mathfrak{Im}\left(\sum_{l=0}^{n}\binom{2n+1}{2l+1}(i\sin\theta)^{2l+1}(\cos\theta)^{2(n-l)}\right)\\ &= \sum_{l=0}^{n}\binom{2n+1}{2l+1}(-1)^l(\sin\theta)^{2l+1}(1-\sin^2\theta)^{n-l}.\end{aligned}$$

On a ainsi exhibé un polynôme de degré au plus $2n+1$ qui convient.
2. Il est évident que les $2n+1$ réels $\sin\frac{k\pi}{2n+1}$ avec $k \in [\![0,2n]\!]$ sont racines de P_n. Comme $\deg P_n \leq 2n+1$, on a ainsi toutes les racines et celles-ci sont simples. La décomposition recherchée est alors

$$\frac{1}{P_n} = \sum_{k=0}^{2n} \frac{a_k}{X - \sin\frac{k\pi}{2n+1}},$$

avec pour tout $k \in [\![0,2n]\!]$, $a_k = \frac{1}{P_n'(\sin\frac{k\pi}{2n+1})}$. Or, en dérivant la relation définissant P_n, on obtient

$$\forall \theta \in \mathbb{R}, \qquad \cos\theta\, P_n'(\sin\theta) = (2n+1)\cos(2n+1)\theta.$$

Ainsi, $a_k = \frac{(-1)^k}{2n+1}\cos\frac{k\pi}{2n+1}$ pour tout $k \in [\![0,2n]\!]$.

Exercice 7

1. S'il existe k tel que $Q'(x_k) = 0$, alors x_k est une racine double de Q ce qui contredit le fait que Q admette n racines distinctes (donc simples car $\deg Q = n$).
2. Les pôles sont simples et en calculant les coefficients associés avec la formule du cours, on obtient directement le résultat.
3. a. Posons $P = 1$ et évaluons la formule précédente en $x = 0$,

$$\sum_{k=1}^{n} \frac{1}{x_k Q'(x_k)} = -\frac{1}{Q(0)}.$$

b. Posons $P = X$ et évaluons en $x = 0$,

$$\sum_{k=1}^{n} \frac{1}{Q'(x_k)} = 0.$$

Corrigés des exercices d'approfondissement du chapitre 13

Exercice A

1. a. La fraction R est de degré au plus 0 (comme somme de fractions de degré 0 ou -1) donc $\deg P \leq n$.
b. Les complexes $\alpha_1, \alpha_2, \ldots, \alpha_n$ sont des racines de R donc de P. Ainsi,

$$P(X) = c\prod_{j=1}^{n}(X - \alpha_j)$$

avec $c \in \mathbb{C}$.
2. En décomposant R en éléments simples, on obtient $c = 1$ (partie entière) et, pour tout k,

$$x_k = (-1)^n \frac{\prod_{j=1}^{n}(a_k + \alpha_j)}{\prod_{\substack{j=1\\ j\neq k}}^{n}(a_j - a_k)}.$$

Exercice B

1. ▹ Le résultat est évident pour les racines de P' qui sont aussi racines de P (c'est-à-dire pour les racines multiples de P).
▹ Soit b une racine de P' qui n'est pas racine de P.
Décomposons en éléments simples la fraction rationnelle

$$\frac{P'}{P} = \sum_{k=1}^{r} \frac{\alpha_k}{X - a_k}.$$

Appliquons ensuite cette relation en b et multiplions chaque fraction par la quantité conjuguée du dénominateur

$$\sum_{k=1}^{r} \frac{\alpha_k}{|b-a_k|^2}(\overline{b-a_k}) = 0,$$

qui devient alors

$$\left(\sum_{k=1}^{r} \frac{\alpha_k}{|b-a_k|^2}\right).b = \sum_{k=1}^{r} \frac{\alpha_k}{|b-a_k|^2}.a_k.$$

Ainsi b est une combinaison linéaire à coefficients positifs et de somme 1 des complexes a_k.
2. Raisonnons par l'absurde en utilisant la première question. Soit $z \in \mathbb{C}$. Si le polynôme $P-z$ n'admet aucune racine dans H, alors P' n'admet pas de racine dans H (car les racines de P' sont dans l'enveloppe convexe des racines de P donc dans le complémentaire de H ce qui contredit l'hypothèse $0 \in P'(H)$. Par conséquent, $P-z$ admet une racine dans H pour tout $z \in \mathbb{C}$ et donc $P(H) = \mathbb{C}$.

Corrigés des problèmes du chapitre 13

Problème 1

Partie A – Une décomposition en élément simple

1. a. Pour tout $\theta \in \mathbb{R}$,

$$T_{2n}(-\cos\theta) = T_{2n}(\cos(\theta+\pi) = \cos(2n(\theta+\pi)) = \cos(2n\theta) = T_{2n}(\cos\theta).$$

b. Les polynômes T_{2n} et $T_{2n}(-X)$ coïncident sur une infinité de valeurs (celles de $[-1,1]$) donc sont égaux : T_{2n} est donc pair.
2. a. La fonction cos est strictement décroissante sur l'intervalle $[0,\pi]$ et tous les θ_k appartiennent à cet ensemble dans tous les $\cos\theta_k$ pour $k \in [\![0, 2n-1]\!]$ sont distincts.
b. Pour tout $k \in [\![0, 2n-1]\!]$, $T_{2n}(\cos(\theta_k)) = \cos(2n\theta_k) = \cos\frac{2k+1}{2}\pi = 0$. On a ainsi exhibé $2n$ racines du polynôme T_{2n} qui est de degré $2n$: on a donc toutes les racines et celles-ci sont simples.
3. La fraction $\frac{P}{T_{2n}}$ est de degré strictement négatif donc de partie entière nulle. Le dénominateur est scindé à racines simples donc la décomposition est de la forme

$$\frac{P}{T_{2n}} = \sum_{k=0}^{2n-1} \frac{a_k}{X-\cos\theta_k}.$$

D'après la formule du cours, $a_k = \frac{P(\cos\theta_k)}{T'_{2n}(\cos\theta_k)}$. Comme $T_{2n}(\cos\theta) = \cos(2n\theta)$ pour tout θ, on obtient en dérivant $T'_{2n}(\cos\theta)\sin\theta = 2n\sin 2n\theta$ donc

$$a_k = \frac{P(\cos\theta_k)\sin\theta_k}{2n\sin 2n\theta_k} = (-1)^k \frac{P(\cos\theta_k)\sin\theta_k}{2n}.$$

4. Comme précédemment $T'_{2m}(\cos\theta_k)\sin\theta_k = 2m\sin 2m\theta_k$, d'où

$$\frac{T'_{2m}}{T_{2n}} = \sum_{k=0}^{2n-1} (-1)^k \frac{m\sin 2m\theta_k}{n} \frac{1}{X-\cos\theta_k}.$$

CORRIGÉS

Or, la fraction $\frac{T'_{2m}}{T_{2n}}$ est impaire (question 1) donc

$$\begin{aligned}
2\frac{T'_{2m}}{T_{2n}} &= \frac{T'_{2m}}{T_{2n}}(X) - \frac{T'_{2m}}{T_{2n}}(-X) \\
&= \sum_{k=0}^{2n-1}(-1)^k \frac{m\sin 2m\theta_k}{n}\left(\frac{1}{X-\cos\theta_k} - \frac{1}{-X-\cos\theta_k}\right) \\
&= \sum_{k=0}^{2n-1}(-1)^k \frac{m\sin 2m\theta_k}{n}\frac{2X}{X^2-\cos^2\theta_k}.
\end{aligned}$$

On conclut en évaluant en $\cos\theta$.

5. La somme considérée est égale au second membre de l'identité précédente évaluée en $\theta = 0$. D'où

$$\sum_{k=0}^{2n-1}(-1)^k \frac{\sin 2m\theta_k}{\sin^2\theta_k} = \frac{n}{m}\lim_{\theta\to 0}\frac{T'_{2m}(\cos\theta)}{T_{2n}(\cos\theta)}.$$

Or, d'après l'expression de la dérivée

$$T'_{2m}(\cos\theta) = 2m\frac{\sin 2m\theta}{\sin\theta} \sim 4m^2.$$

Par conséquent,

$$\sum_{k=0}^{2n-1}(-1)^k \frac{\sin 2m\theta_k}{\sin^2\theta_k} = 4mn.$$

Partie B – Une identité de Riesz

1. L'équation $(\star)$ est linéaire en Φ donc en P : il suffit d'établir cette relation pour la base canonique de $\mathbb{R}_n[X]$, c'est-à-dire les polynômes $P = X^m$ pour tout $m \leq n$.
2. Remplaçons $P = X^m$ dans l'identité $(\star)$:

$$ime^{im\theta} = \frac{1}{2n}\sum_{k=0}^{2n-1}(-1)^k \frac{e^{im(\theta+2\theta_k)}}{2\sin^2\theta_k}.$$

En simplifiant par $e^{im\theta}$, on obtient

$$\sum_{k=0}^{2n-1}(-1)^k \frac{e^{i2m\theta_k}}{\sin^2\theta_k} = 4imn.$$

3. Pour tout $k \in [\![0, 2n-1]\!]$, $\theta_{2n-1-k} + \theta_k = \frac{2(2n-1-k)+1}{4n}\pi + \frac{2k+1}{4n}\pi = \pi$.
4. Notons $S = \sum_{k=0}^{2n-1}(-1)^k \frac{e^{i2m\theta_k}}{\sin^2\theta_k}$. Avec le changement de variables $k = 2n-1-l$, on obtient

$$S = -\sum_{l=0}^{2n-1}(-1)^l \frac{e^{-i2m\theta_l}}{\sin^2\theta_l}.$$

En sommant les deux expressions pour S,

$$2S = \sum_{k=0}^{2n-1}(-1)^k \frac{2i\sin(2m\theta_k)}{\sin^2\theta_k} = 8imn,$$

en utilisant le dernier résultat de la partie A.
Par conséquent, on a établi la formule de la question 2 donc la formule de Riesz.

Problème 2

Partie A – Interpolation de Lagrange

Voir le cours.

Partie B – Interpolation de Cauchy

1. Lorsque $p = n$, la fraction rationnelle est un polynôme et nous sommes ramenés à la partie A.
2. a. Le polynôme M est scindé et aucune de ses racines n'est racine de L (hypothèse sur les b_k) donc $M \wedge L = 1$.
b. Par sa définition comme suite de restes dans une division euclidienne, la suite $(\deg R_i)_i$ est strictement décroissante (jusqu'à atteindre la valeur $-\infty$) donc elle atteint des valeurs inférieures à p.
c. Procédons par récurrence double.
▹ $U_0 L + V_0 M = 0L + 1M = M = R_0$ et $U_1 L + V_1 M = 1L + 0M = L = R_1$.
▹ Soit i tel que l'égalité est vraie au rang i et $i+1$. Alors,

$$\begin{aligned} U_{i+2}L + V_{i+2}M &= (U_i - Q_{i+2}U_{i+1})L + (V_i - Q_{i+2}V_{i+1})M \\ &= (U_i L + V_i M) - Q_{i+2}(U_{i+1}L + V_{i+1}M) \\ &= R_i - Q_{i+2}R_{i+1} = R_{i+2}. \end{aligned}$$

d. Montrons par récurrence sur i que $\deg U_{i+1} \geq \deg U_i$ et

$$\deg U_{i+1} = \sum_{k=2}^{i+1} \deg Q_k.$$

▹ Pour $i = 1$, $\deg U_2 = \deg -Q_2 = 1$ et $\deg U_1 = 0$ ce qui établit les deux propriétés.
▹ Soit i tel que les propriétés soient vérifiées au rang i. Alors, $\deg U_{i+2} = \deg(U_i - Q_{i+2}U_{i+1}) = \deg Q_{i+2}U_{i+1}$ car $\deg U_{i+1} \geq \deg U_i$. Par conséquent,

$$\deg U_{i+2} = \deg Q_{i+2} + \deg U_{i+1} \geq \deg U_{i+1}$$

$$\deg U_{i+2} = \deg Q_{i+2} + \deg U_{i+1} = \deg Q_{i+2} + \sum_{k=2}^{i+1} \deg Q_k.$$

e. Remarquons que pour tout $i \leq j$, $\deg Q_i = \deg R_{i-2} - \deg R_{i-1}$. Ainsi, par télescopage,

$$\deg U_j = \deg R_0 - \deg R_{j-1} \leq n - p$$

car $\deg R_{j-1} > p$ par minimalité de j.
f. D'après ce qui précède la fraction R_j/U_j vérifie les hypothèses recherchées : la fraction rationnelle d'interpolation existe.
3. Supposons b_1 nul et b_2 non nul et prenons $p = 0$. Alors, le polynôme au numérateur d'une fraction rationnelle d'interpolation est constant non nul mais admet b_1 comme racine : contradiction. Dans ce cas, il n'existe pas de fraction rationnelle d'interpolation.

Partie C – Interpolation de Hermite

1. La différence de deux polynômes d'interpolation est de degré strictement inférieur à $\sum_{k=1}^{n} \alpha_k$ et est divisible par $\prod_{k=1}^{n} (X - a_k)^{\alpha_k}$ donc est nulle. Il y a ainsi unicité du polynôme interpolateur.
2. a. Soit $i \neq k$. Comme $L_{k,l}$ est un polynôme, a_i est racine de $H_{k,l}$ avec multiplicité au moins α_i. Par conséquent, pour tout $j < \alpha_i$, $H_{k,l}^{(j)}(a_i) = 0$.
b. Par définition

$$L_{k,l}(x) \underset{a_k}{=} \prod_{\substack{j=1 \\ j \neq k}}^{n} \frac{1}{(x - a_j)^{\alpha_j}} + o((x - a_k)^{\alpha_k - l})$$

donc

$$H_{k,l}(x) \underset{a_k}{=} \frac{1}{l!}(x-a_k)^l + o((x-a_k)^{\alpha_k}).$$

Par conséquent, pour tout $j < \alpha_k$, $H_{k,l}^{(j)}(a_k) = \delta_{l,j}$.
3. D'après la question précédente, un polynôme d'interpolation associé à ces paramètres est obtenu avec la combinaison linéaire

$$\sum_{k=1}^{n} \sum_{l=0}^{\alpha_k - 1} b_{k,l} H_{k,l}.$$

4. Le polynôme d'interpolation pour les paramètres proposés est de degré au plus 2 donc de la forme $P = aX^2 + bX + c$.
Comme $P(0) = 1$, $c = 1$. Comme $P'(0) = 0$, $b = 0$. Enfin, comme $P(1) = 3$, $a = 2$. En conclusion, le polynôme recherché est $2X^2 + 1$.

Chapitre 14

Espaces vectoriels

1. Structure d'espace vectoriel — **2.** Applications linéaires — **3.** Endomorphismes remarquables

Objectifs et compétences du programme

- Démontrer la structure algébrique d'espace vectoriel.
- Comprendre le sous-espace vectoriel engendré par une partie.
- Montrer la linéarité d'une application.
- Passer de la définition algébrique à la définition géométrique des projecteurs et symétries.

William Rowan Hamilton, mathématicien et physicien irlandais (1805-1865).
Jeune prodige, il parvient à effectuer des recherches scientifiques de haut-niveau en optique, en mécanique et en mathématiques. Sa découverte la plus célèbre est indéniablement les quaternions, une extension non commutative du corps des nombres complexes. Hamilton raconte que lors d'une promenade avec on épouse le long du canal Royal à Dublin le 16 octobre, il a une révélation des relations $i^2 = j^2 = k^2 = ijk = -1$ définissant le produit sur les quaternions ; il les écrit alors sur la pierre du *Broom Bridge*. Depuis 1958, une plaque sur ce pont commémore cette découverte, un des rare eurekas mathématiques à être aussi bien localisée dans l'espace et dans le temps.
C'est pour étudier les applications linéaires sur les quaternions qu'il démontre un cas particulier de ce qui est désormais appelé le théorème de Cayley-Hamilton.

COURS

Dans ce chapitre, on introduit une nouvelle structure algébrique, les espaces vectoriels, qui permet de généraliser le calcul vectoriel déjà rencontré dans $\mathbb{R}^2$ ou $\mathbb{R}^3$. On met particulièrement en évidence la propriété de linéarité (qui est la propriété de morphisme associée à la structure) et l'importance des combinaisons linéaires. Le chapitre se termine avec la reprise des notions de projections, de symétries et d'homothéties.

1. Structure d'espace vectoriel

1.1. Espaces vectoriels

Dans tout ce chapitre, $\mathbb{K}$ désigne le corps $\mathbb{R}$ ou le corps $\mathbb{C}$.

Définition 14.1.

Un $\mathbb{K}$-espace vectoriel est un ensemble E muni d'une loi de composition interne $+$ et d'une application $\left\{\begin{array}{ccc} \mathbb{K}\times E & \longrightarrow & E \\ (\lambda, x) & \longmapsto & \lambda.x \end{array}\right.$ telles que

- $(E,+)$ est un groupe abélien ;
- la loi externe est distributive sur les additions de E et de $\mathbb{K}$:

$$\forall(\lambda_1,\lambda_2)\in\mathbb{K}^2,\ \forall(x_1,x_2)\in E^2,\qquad \begin{array}{rcl}(\lambda_1+\lambda_2).x_1 & = & \lambda_1.x_1+\lambda_2.x_1\\ \lambda_1.(x_1+x_2) & = & \lambda_1.x_1+\lambda_1.x_2;\end{array}$$

- pour tout $x\in E$, $1.x=x$;
- $\forall(\lambda_1,\lambda_2)\in\mathbb{K}^2,\ \forall x\in E,\qquad \lambda_1.(\lambda_2.x)=(\lambda_1\lambda_2).x.$

Remarque

Les éléments d'un espace vectoriel sont appelés vecteurs et les éléments du corps de base sont appelés scalaires. On note toujours les produits dans l'ordre scalaire puis vecteur.

Exemple

Le corps $\mathbb{K}$ est un $\mathbb{K}$-espace vectoriel. Pour tout entier $n\in\mathbb{N}\setminus\{0\}$, $\mathbb{K}^n$ est un $\mathbb{K}$-espace vectoriel. Cet exemple est un bon prototype pour l'intuition d'espace vectoriel. Lorsque l'on voudra illustrer des propriétés des espaces vectoriels, on se placera dans $\mathbb{R}^2$ ou $\mathbb{R}^3$.

Cet exemple se généralise à tout produit d'espaces vectoriels.

Proposition 14.2.

Tout produit cartésien de $\mathbb{K}$-espaces vectoriels est un $\mathbb{K}$-espace vectoriel (avec les opérations terme à terme).

Exemple

L'ensemble des applications entre un ensemble A et un $\mathbb{K}$-espace vectoriel est aussi un $\mathbb{K}$-espace vectoriel. Ainsi, les suites réelles et les fonctions réelles forment des $\mathbb{R}$-espaces vectoriels.

Exemple

L'ensemble des polynômes à coefficients dans un corps $\mathbb{K}$ est un $\mathbb{K}$-espace vectoriel.

Proposition 14.3.

Soit $\lambda \in \mathbb{K}$ et x dans un $\mathbb{K}$-espace vectoriel E. Alors,
▹ $\lambda.0_E = 0_E$
▹ $0.x = 0_E$
▹ $(-1).x = -x$
▹ $\lambda.x = 0_E$ entraîne $\lambda = 0$ ou $x = 0_E$.

Montrons le dernier point.

Démonstration
▹ Comme $0_E + 0_E = 0_E$, on obtient

$$\lambda.0_E = \lambda.(0_E + 0_E) = \lambda.0_E + \lambda.0_E.$$

En ajoutant l'opposé de $\lambda.0_E$ (qui existe d'après la propriété de groupe additif), on obtient $0_E = \lambda.0_E$.
▹ On procède de même en partant de $0 + 0 = 0$ (dans le corps $\mathbb{K}$) et on obtient

$$0.x = (0+0).x = 0.x + 0.x$$

donc $0.x = 0_E$.
▹ On part cette fois-ci de $0 = 1 + (-1)$ et on obtient

$$0_E = 0.x = (1 + (-1)).x = 1.x + (-1).x = x + (-1).x$$

En ajoutant l'opposé $-x$ de x, on obtient bien $-x = (-1).x$.
▹ Supposons $\lambda.x = 0_E$. Si $\lambda \neq 0$, alors λ est inversible et

$$0_E = \lambda^{-1}.0_E = \lambda^{-1}.(\lambda.x) = (\lambda^{-1}\lambda).x = 1.x = x.$$

■

1.2. Sous-espaces vectoriels

Définition 14.4.

Une partie F d'un $\mathbb{K}$-espace vectoriel E est un sous-espace vectoriel de E si
- F est stable par $+$ et $.$
- muni de ces lois, F est un $\mathbb{K}$-espace vectoriel.

Exemple

Si E est un $\mathbb{K}$-espace vectoriel, E et $\{0_E\}$ sont deux sous-espaces vectoriels de E.
Attention, l'ensemble vide n'est pas un sous-espace vectoriel de E car il ne contient pas l'élément neutre de l'addition 0_E donc n'est pas un groupe additif.

Pour caractériser les sous-espaces vectoriels, nous avons besoin de la notion de combinaison linéaire.

Définition 14.5.

Soit $(x_i)_{i\in I}$ une famille de vecteurs d'un $\mathbb{K}$-espace vectoriel E. Une combinaison linéaire de $(x_i)_{i\in I}$ est un vecteur de la forme $\sum_{i\in I}\lambda_i x_i$ où $(\lambda_i)_{i\in I}$ est une famille presque nulle de scalaires.

Exemple

▹ Les fonctions sh et ch sont des combinaisons linéaires réelles des fonctions $x\mapsto e^x$ et $x\mapsto e^{-x}$.
▹ Considérons $f_1 : x\mapsto \cos x$, $f_2 : x\mapsto \sin x$, $f_3 : x\mapsto \sin(x+1)$ et $g : x\mapsto \sin(x-1)$. On a alors

$$g = \cos(1)f_2 - \sin(1)f_1 = \cos(1)\sin(2)f_1 - \sin(1)\sin(2)f_2 + \cos(2)f_3.$$

Ainsi, g est combinaison linéaire des fonctions f_1, f_2 et f_3 : il n'y a pas unicité de cette combinaison linéaire.
▹ Un polynôme est une combinaison linéaire de la famille $(X^n)_{n\in\mathbb{N}}$.

Avec la notion de combinaison linéaire, on obtient une caractérisation maniable des sous-espaces vectoriels.

Proposition 14.6.

Une partie non vide d'un $\mathbb{K}$-espace vectoriel est un sous-espace vectoriel si, et seulement si, elle est stable par combinaison linéaire de deux vecteurs.

Démonstration

▹ Le sens direct est évident par récurrence sur le nombre de coefficients non nuls.

▹ Considérons F une partie non vide d'un $\mathbb{K}$-espace vectoriel E stable par combinaison linéaire de deux vecteurs. Ainsi, pour tous x, $y\in F$, $x-y\in F$ donc F est un sous groupe additif de E. De plus, la loi . est bien définie de $\mathbb{K}\times F$ dans F et les autres propriétés sont héritées des propriétés correspondantes sur E.

■

Conseils méthodologiques

On utilise souvent cette caractérisation de sous-espace vectoriel pour montrer qu'un ensemble est un espace vectoriel à partir d'un « grand » espace vectoriel. En pratique, pour montrer qu'un ensemble A est un espace vectoriel, on montre successivement

- que A est incluse dans un (grand) espace vectoriel E,
- que A est non vide,
- que A est stable par combinaison linéaire.

Exemple

On vérifie rapidement que les exemples suivants satisfont la caractérisation ci-dessus.
▹ La partie $\mathbb{R}\times\{0_E\}$ est un sous-espace vectoriel de $\mathbb{R}^2$.

▹ L'ensemble des vecteurs $\overrightarrow{OM}$ où M décrit une droite passant par l'origine est un sous-espace vectoriel du plan.
▹ L'ensemble des fonctions polynomiales réelles est un sous-espace vectoriel de $\mathbb{R}^{\mathbb{R}}$.
▹ L'ensemble des fonctions continues (respectivement dérivables, $\mathscr{C}^p$, $\mathscr{C}^\infty$) est un sous-espace vectoriel de $\mathbb{R}^{\mathbb{R}}$.

▹ L'ensemble des solutions d'une équation différentielle linéaire homogène est un sous-espace vectoriel de $\mathbb{R}^{\mathbb{R}}$.
▹ Soit $a, b, c \in \mathbb{R}$. L'ensemble des suites $(u_n)_n$ vérifiant :

$$\forall n \in \mathbb{N}, \quad a\,u_{n+2} + b\,u_{n+1} + c\,u_n = 0$$

est un sous-espace vectoriel de $\mathbb{R}^{\mathbb{N}}$.
▹ L'ensemble des n-uplets vérifiant un système linéaire sans second membre est un sous-espace vectoriel de $\mathbb{K}^n$.
▹ L'ensemble $\mathbb{K}_n[X]$ est un sous-espace vectoriel de $\mathbb{K}[X]$. En revanche, l'ensemble des polynômes de degré au moins n n'est pas un sous-espace vectoriel (il ne contient pas 0 et n'est pas stable par addition : $(X^n+1)+(-X^n)=1$).

1.3. Sous-espaces vectoriels engendrés

Voici un exemple intéressant d'application directe de la caractérisation des sous-espaces vectoriels.

Proposition 14.7.

L'intersection d'une famille non vide de sous-espaces vectoriels est encore un sous-espace vectoriel.

On peut ainsi obtenir le plus petit sous-espace contenant une partie A d'un espace vectoriel E comme intersection de tous les sous-espaces contenant A.

Définition 14.8.

Soit A une partie d'un $\mathbb{K}$-espace vectoriel E. Le sous-espace engendré par A est le plus petit (au sens de l'inclusion) sous-espace vectoriel de E contenant A ; il est noté $\mathrm{Vect}(A)$. En particulier, $\mathrm{Vect}(\varnothing) = \{0_E\}$.

Notation

Si $A = \{x\}$, on note « abusivement » $\mathrm{Vect}(x)$ pour $\mathrm{Vect}(\{x\})$; plus généralement, si $A = \{x_i,\ i \in I\}$, on préfère la notation plus simple $\mathrm{Vect}(x_i)_{i\in I}$ à la rigoureuse $\mathrm{Vect}(\{x_i,\ i \in I\})$.

Remarque

Par construction, tout sous-espace vectoriel contenant A contient $\mathrm{Vect}(A)$.

Exemple

Soit E un $\mathbb{K}$-espace vectoriel et x un vecteur de E non nul. La droite vectorielle engendrée par x est le sous-espace vectoriel $\mathrm{Vect}(x) = \{\lambda.x,\ \lambda \in \mathbb{K}\}$ formé des vecteurs colinéaires à x.

On déduit aisément de cette définition les propriétés suivantes.

Proposition 14.9.

Soit A, B deux parties d'un espace vectoriel E.
▹ $A = \mathrm{Vect}(A)$ si, et seulement si, A est un sous-espace vectoriel de E.
▹ Si $A \subset B$, alors $\mathrm{Vect}(A) \subset \mathrm{Vect}(B)$.

Remarque

La réciproque au deuxième point est fausse. Exhibons quelques exemples de parties $A \not\subset B$ telles que $\operatorname{Vect} A \subset \operatorname{Vect} B$.

▹ Posons $A = \{(1,1)\}$ et $B = \{(2,2)\}$. Par définition des droites vectorielles, $\operatorname{Vect} A = \operatorname{Vect} B$ et pourtant on a $A \not\subset B$.

▹ Donnons un exemple où $\operatorname{Vect} A \neq \operatorname{Vect} B$. Posons A l'ensemble des triplets de la forme $(x,x,0)$ avec $x \in \mathbb{R}$ et B l'ensemble des triplets (x,y,z) tels que $x = 0$ ou ($y = 0$ et $z = 0$). A n'est pas inclus dans B et pourtant $\operatorname{Vect}(A) = A \subset \mathbb{R}^3 = \operatorname{Vect}(B)$.

Proposition 14.10.

Soit $(x_i)_{i\in I}$ une famille de vecteurs d'un $\mathbb{K}$-espace vectoriel E. Le sous-espace vectoriel $\operatorname{Vect}(x_i)_{i\in I}$ est l'ensemble des combinaisons linéaires des vecteurs $(x_i)_{i\in I}$.

Démonstration

▹ Comme $\operatorname{Vect}(x_i)_{i\in I}$ est un espace vectoriel, il contient toutes les combinaisons linéaires de ses éléments.

▹ Pour obtenir l'autre inclusion, il suffit de vérifier que l'ensemble des combinaisons linéaires des vecteurs $(x_i)_{i\in I}$ est bien un sous-espace de E qui contient $(x_i)_{i\in I}$, ce qui découle directement de la caractérisation des sous-espaces vectoriels. ■

Exemple

Dans $\mathbb{R}^3$, $\operatorname{Vect}((1,2,-3),(-3,2,1),(1,-2,1)) = \operatorname{Vect}((1,-1,0),(0,1,-1))$ car chacun des vecteurs d'une des familles peut s'exprimer comme combinaison linéaire des vecteurs de l'autre famille. Par exemple,

$$(1,2,-3) = 1(1,-1,0) + 3(0,1,-1)$$
$$(1,-1,0) = -\frac{1}{8}(1,2,-3) - \frac{3}{8}(-3,2,1) + 0(1,-2,1)$$

Définition 14.11.

Une partie A est génératrice d'un sous-espace vectoriel F de E si $F = \operatorname{Vect}(A)$.

Exemple

D'après la résolution des équations différentielles linéaires à coefficients constants, l'espace des solutions $\{y \in \mathscr{C}^2(\mathbb{R},\mathbb{R}),\ y'' + 2y' + y = 0\}$ est engendré par $x \mapsto e^{-x}$ et $x \mapsto xe^{-x}$.

Exemple

Considérons E, l'ensemble des quadruplets $(x,y,z,t) \in \mathbb{R}^4$ vérifiant le système linéaire homogène

$$\left\{\begin{array}{rcrcrcrcl} x & + & y & + & z & + & t & = & 0 \\ 2x & - & y & - & z & + & 3t & = & 0 \\ x & - & 5y & - & 5z & + & 3t & = & 0 \end{array}\right.$$

Alors, $(x,y,z,t) \in E$ si, et seulement si,

$$\left\{\begin{array}{rcrcrcrcll} x & + & y & + & z & + & t & = & 0 & \\ & - & 3y & - & 3z & + & t & = & 0 & L_2 \leftarrow L_2 - 2L_1 \\ & - & 6y & - & 6z & + & 2t & = & 0 & L_3 \leftarrow L_3 - L_1 \end{array}\right.$$

soit, en remarquant que les deux dernières équations sont proportionnelles,

$$\begin{cases} x + y + z + t = 0 \\ \quad - 3y - 3z + t = 0 \end{cases}$$

Ainsi, $(x, y, z, t) \in E$ si, et seulement si,

$$\begin{aligned}(x, y, z, t) &= (-4y - 4z, y, z, 3y + 3z) \\ &= y(-4, 1, 0, 3) + z(-4, 0, 1, 3).\end{aligned}$$

En conclusion, E est égal à $\mathrm{Vect}((-4, 1, 0, 3), (-4, 0, 1, 3))$, c'est-à-dire la famille composée des vecteurs $(-4, 1, 0, 3)$, $(-4, 0, 1, 3)$ est génératrice de E.

Proposition 14.12.

Soit $e_1, \dots, e_p$ des vecteurs d'un espace vectoriel E.
▹ Pour tous $\lambda \in \mathbb{K}^*$ et $i \in [\![1, p]\!]$, $\mathrm{Vect}(e_1, \dots, e_p) = \mathrm{Vect}(e_1, \dots, \lambda e_i, \dots, e_p)$.
▹ Pour tous $\lambda \in \mathbb{K}$ et $i \neq j \in [\![1, p]\!]$, $\mathrm{Vect}(e_1, \dots, e_p) = \mathrm{Vect}(e_1, \dots, e_i + \lambda e_j, \dots, e_p)$.

Démonstration

▹ Soit $\lambda \in \mathbb{K}^*$ et $i \in [\![1, p]\!]$.

$$\begin{aligned} x \in \mathrm{Vect}(e_1, \dots, e_p) &\Leftrightarrow \exists (x_1, \dots, x_p) \in \mathbb{K}^p, \quad x = \sum_{k=1}^{p} x_k e_k \\ &\Leftrightarrow \exists (x_1, \dots, x_p) \in \mathbb{K}^p, \quad x = \sum_{\substack{k=1 \\ k \neq i}}^{p} x_k e_k + (x_i \lambda^{-1}) \lambda e_i \\ &\Leftrightarrow \exists (y_1, \dots, y_p) \in \mathbb{K}^p, \quad x = \sum_{\substack{k=1 \\ k \neq i}}^{p} y_k e_k + y_i \lambda e_i \\ &\Leftrightarrow x \in \mathrm{Vect}(e_1, \dots, \lambda e_i, \dots, e_p). \end{aligned}$$

▹ Soit $\lambda \in \mathbb{K}$ et $i \neq j \in [\![1, p]\!]$.

$$\begin{aligned} x \in \mathrm{Vect}(e_1, \dots, e_p) &\Leftrightarrow \exists (x_1, \dots, x_p) \in \mathbb{K}^p, \quad x = \sum_{k=1}^{p} x_k e_k \\ &\Leftrightarrow \exists (x_1, \dots, x_p) \in \mathbb{K}^p, \\ &\quad x = \sum_{\substack{k=1 \\ k \neq i,j}}^{p} x_k e_k + x_i (e_i + \lambda e_j) + (x_j - \lambda x_i) e_j \\ &\Leftrightarrow \exists (y_1, \dots, y_p) \in \mathbb{K}^p, \quad x = \sum_{\substack{k=1 \\ k \neq i,j}}^{p} y_k e_k + y_i (e_i + \lambda e_j) + y_j e_j \\ &\Leftrightarrow x \in \mathrm{Vect}(e_1, \dots, e_i + \lambda e_j, \dots, e_p). \end{aligned}$$

■

1.4. Sommes et sommes directes de sous-espaces

Définition 14.13.

Soit E un $\mathbb{K}$-espace vectoriel et $(F_i)_{i\in[1,p]}$ des sous-espaces vectoriels de E.
▹ La somme des sous-espaces $(F_i)_{i\in[1,p]}$ est le sous-espace vectoriel

$$F_1+\ldots+F_p=\{x_1+\ldots+x_p,\ (x_1,\ldots,x_p)\in F_1\times\ldots\times F_p\}.$$

▹ La somme directe des sous-espaces $(F_i)_{i\in[1,p]}$ est le sous-espace $F_1+\ldots+F_p$ lorsque pour tout vecteur $x\in F_1+\ldots+F_p$, il existe un unique $(x_1,\ldots,x_p)\in F_1\times\ldots\times F_p$ tel que $x=x_1+\ldots+x_p$.
La somme est alors notée $F_1\oplus\ldots\oplus F_p$.

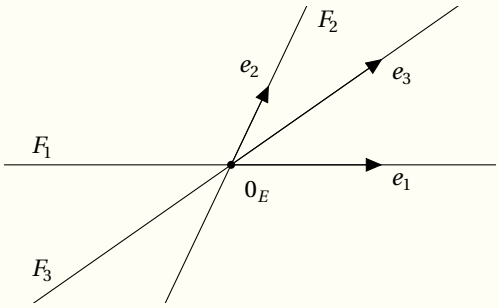
Une somme est donc directe si chaque vecteur de celle-ci admet une unique décomposition comme somme de vecteurs des sous-espaces considérés. Dans le cas de deux sous-espaces seulement, on obtient la caractérisation suivante plus maniable.

Proposition 14.14.

La somme de deux sous-espaces vectoriels est directe si, et seulement si, leur intersection est égale à $\{0_E\}$.

Remarque

Ce résultat est faux pour plus de deux sous-espaces comme on peut le voir avec trois droites F_1, F_2 et F_3 du plan passant par l'origine et deux à deux distinctes.



En effet, si on note e_1, e_2 et e_3 des vecteurs directeurs de chacune de ces trois droites, alors (e_1,e_2) est une base du plan ; ainsi, il existe $\lambda_1,\lambda_2\in\mathbb{R}$ tels que $e_3=\lambda_1e_1+\lambda_2e_2$. Il n'y a donc pas unicité de la décomposition sur la somme $\mathrm{Vect}(e_1)+\mathrm{Vect}(e_2)+\mathrm{Vect}(e_3)$ car $0e_1+0e_2+1e_3=\lambda_1e_1+\lambda_2e_2+0e_3$.

Démonstration

▹ Supposons que F_1 et F_2 soient en somme directe. Alors, tout $x\in F_1\cap F_2$ s'écrit

$$\underbrace{x}_{\in F_1}+\underbrace{0_E}_{\in F_2}=\underbrace{0_E}_{\in F_1}+\underbrace{x}_{\in F_2}.$$

Par unicité de la décomposition, $x=0_E$.
▹ Réciproquement, supposons $F_1\cap F_2=\{0_E\}$. Soit $x_1,y_1\in F_1$ et $x_2,y_2\in F_2$ tels que $x_1+x_2=y_1+y_2$. Alors, $x_1-y_1=y_2-x_2\in F_1\cap F_2=\{0_E\}$ d'où $x_1=y_1$ et $x_2=y_2$: il y a unicité de l'écriture comme somme d'un vecteur de F_1 et d'un vecteur de F_2, la somme est directe. ■

Exemple

Deux droites vectorielles sont soit confondues, soit en somme directe. En effet, si $\mathrm{Vect}(x)$ et $\mathrm{Vect}(y)$ (avec x et y des vecteurs non nuls) ne sont pas en somme directe, il existe $z \in \mathrm{Vect}(x) \cap \mathrm{Vect}(y)$ non nul : il existe donc λ et μ des scalaires non nuls tels que $z = \lambda x = \mu y$; alors, $x = \frac{\mu}{\lambda} y \in \mathrm{Vect}(y)$ et $y = \frac{\lambda}{\mu} x \in \mathrm{Vect}(x)$. En conclusion, $\mathrm{Vect}(x) = \mathrm{Vect}(z) = \mathrm{Vect}(y)$.

Proposition 14.15.

Soit A, B deux parties d'un espace vectoriel E. Alors, $\mathrm{Vect}(A \cup B) = \mathrm{Vect}(A) + \mathrm{Vect}(B)$.
En particulier, pour deux sous-espaces vectoriels F_1 et F_2, $\mathrm{Vect}(F_1 \cup F_2) = F_1 + F_2$.

Remarque

Comme pour les sous-groupes, la réunion de deux sous-espaces vectoriels n'est en général pas un sous-espace vectoriel sauf si l'un est inclus dans l'autre.

Démonstration

L'ensemble $\mathrm{Vect}(A) + \mathrm{Vect}(B)$ est un sous-espace vectoriel contenant $A \cup B$ donc il contient $\mathrm{Vect}(A \cup B)$.

Réciproquement, $\mathrm{Vect}(A \cup B)$ est un sous-espace vectoriel contenant A donc $\mathrm{Vect}(A)$. Il contient de même $\mathrm{Vect}(B)$ donc contient $\mathrm{Vect}(A) + \mathrm{Vect}(B)$. ■

Définition 14.16.

Deux sous-espaces F et G d'un espace vectoriel E sont supplémentaires si $F \oplus G = E$.

Remarque

Attention, il n'y a pas unicité du supplémentaire : on dit donc « un » supplémentaire et non « le » supplémentaire.

Il faut aussi veiller à ne pas confondre un supplémentaire et le complémentaire d'un sous-espace vectoriel (qui n'est jamais un sous-espace vectoriel puisqu'il ne contient pas 0_E par exemple).

Conseils méthodologiques

Pour montrer que deux sous-espaces F et G sont supplémentaires, on utilise souvent un raisonnement par analyse-synthèse (condition nécessaire-condition suffisante) :

- Dans un premier temps, on considère un élément x qui s'écrit $x_F + x_G$ avec $x_F \in F$ et $x_G \in G$ puis on trouve l'expression de x_F et x_G en fonction de x et des données du problème.
  Autrement dit, on suppose l'existence de la décomposition et on établit l'unicité.
- Dans un second temps, on prend un vecteur x quelconque et on utilise les expressions trouvées à l'étape précédente pour x_F et x_G ; on vérifie alors que $x_F \in F$, $x_G \in G$ et $x = x_F + x_G$.
  Autrement dit, on vérifie que tout vecteur admet bien une décomposition.

Voici quelques exemples de ce raisonnement.

Exemple

Considérons l'espace vectoriel $E = \mathbb{R}^{\mathbb{R}}$ des fonctions réelles, F_0 le sous-espace des fonctions de E nulles en 0, g une fonction non-nulle en 0 et $G = \operatorname{Vect} g$. Montrons par analyse-synthèse que les sous-espaces F_0 et G sont supplémentaires dans E.

- Supposons qu'une fonction $f \in E$ soit la somme d'une fonction h qui s'annule en 0 et d'une fonction de la forme $\lambda.g$ avec $\lambda \in \mathbb{R}$. Alors, pour tout $x \in \mathbb{R}$,
$$f(0) = h(0) + \lambda g(0) = \lambda g(0).$$
Ainsi, $\lambda = \frac{f(0)}{g(0)}$ et $h = f - \lambda g$: on a donc l'unicité d'une telle décomposition pour f (c'est-à-dire F_0 et G sont en somme directe).
- Soit $f \in E$ une fonction réelle quelconque. Posons $\lambda = \frac{f(0)}{g(0)}$ et $h = f - \lambda g$; on vérifie que $f = h + \lambda g$, $h(0) = 0$: on a donc l'existence d'une décomposition de f sur la somme $F_0 + G$.

Exemple

Considérons l'espace vectoriel $E = \mathbb{R}^{\mathbb{R}}$ des fonctions réelles et les deux sous-espaces $\mathscr{I}$ et $\mathscr{P}$ des fonctions impaires et paires respectivement. Montrons par analyse-synthèse que les sous-espaces $\mathscr{I}$ et $\mathscr{P}$ sont supplémentaires dans E.

- Supposons qu'une fonction $f \in E$ soit la somme d'une fonction paire p et d'une fonction impaire i. Alors, pour tout $x \in \mathbb{R}$,
$$f(x) = p(x) + i(x)$$
$$f(-x) = p(x) - i(x).$$
Ainsi, $p(x) = \frac{1}{2}(f(x) + f(-x))$ et $i(x) = \frac{1}{2}(f(x) - f(-x))$ et on a donc l'unicité d'une telle décomposition pour f.
- Soit $f \in E$ une fonction réelle quelconque. Posons, pour tout $x \in \mathbb{R}$, $p(x) = \frac{1}{2}(f(x) + f(-x))$ et $i(x) = \frac{1}{2}(f(x) - f(-x))$ et vérifions que p est paire, i impaire et $f = p + i$.

Illustrons l'intérêt de ce résultat en résolvant l'équation $y''(-x) = y(x)$. Une solution y admet une unique décomposition $p + i$ où p est paire et i est impaire (elles sont dérivables d'après l'expression trouvée durant la phase d'analyse). Ainsi $p''(x) - i''(x) = p(x) + i(x)$. Par unicité de cette décomposition, $p''(x) = p(x)$ et $i''(x) = -i(x)$. En ajoutant les conditions de parité, on obtient $p = A.\operatorname{ch}$ et $i = B.\sin$ avec $A, B \in \mathbb{R}$.

2. Applications linéaires

2.1. Définitions

Définition 14.17.

▹ Une application u entre deux $\mathbb{K}$-espaces vectoriels E et F est linéaire si
$$\forall (x,y) \in E^2,\ \forall (\lambda,\mu) \in \mathbb{K}^2, \qquad u(\lambda x + \mu y) = \lambda u(x) + \mu u(y).$$
L'ensemble des applications linéaires entre E et F est noté $\mathscr{L}(E,F)$.
▹ Un endomorphisme de E est une application linéaire de E dans E. L'ensemble des endomorphismes de E est noté $\mathscr{L}(E)$.
▹ Un isomorphisme entre deux $\mathbb{K}$-espaces vectoriels E et F est une application linéaire bijective entre E et F.

▹ Un automorphisme de E est un endomorphisme bijectif de E. L'ensemble des automorphismes de E est noté $\mathrm{GL}(E)$ et appelé groupe linéaire de E.
▹ Une forme linéaire sur E est une application linéaire de E dans le corps $\mathbb{K}$. L'ensemble des formes linéaires sur E est noté E^* et appelé espace dual de E.

Remarque

En fait, une application u entre deux $\mathbb{K}$-espaces vectoriels E et F est linéaire si, et seulement si,

$$\forall (x,y)\in E^2,\ \forall \lambda\in\mathbb{K},\qquad u(\lambda x+y)=\lambda u(x)+u(y).$$

En effet, cette propriété entraîne $u(0_E)=0_E$ (avec $x=y=0_E$), puis $u(\lambda x)=\lambda u(x)$ (avec $y=0$) et la propriété générale en remplaçant y par μy.

Conseils méthodologiques

On a, en passant, établi qu'une application linéaire u entre E et F vérifie $u(0_E)=0_F$. Par contraposition, on peut utiliser cette propriété pour montrer qu'une application telle que $u(0_E)\neq 0_F$ n'est pas linéaire.

Exemple

Soit E un $\mathbb{K}$-espace vectoriel. L'identité Id et l'application identiquement égale à 0_E sont des endomorphismes de E.

Exemple

▹ Les applications $\mathfrak{Im}$ et $\mathfrak{Re}$ sont des formes linéaires sur le $\mathbb{R}$-espace vectoriel $\mathbb{C}$. En effet, pour tous $z,z'\in\mathbb{C}$ et tout $\lambda\in\mathbb{R}$,

$$\mathfrak{Re}(z+\lambda z')=\mathfrak{Re}(z)+\lambda\mathfrak{Re}(z'),\qquad \mathfrak{Im}(z+\lambda z')=\mathfrak{Im}(z)+\lambda\mathfrak{Im}(z').$$

▹ La conjugaison $z\mapsto\overline{z}$ est un endomorphisme du $\mathbb{R}$-espace vectoriel $\mathbb{C}$ mais n'est pas linéaire sur le $\mathbb{C}$-espace vectoriel $\mathbb{C}$. En effet, pour tous $z,z'\in\mathbb{C}$ et tout $\lambda\in\mathbb{R}$,

$$\overline{z+\lambda z'}=\overline{z}+\lambda\overline{z'}.$$

Exemple

Soit $(a,b,c)\in\mathbb{R}^*\times\mathbb{R}^2$. Notons S le $\mathbb{R}$-espace vectoriel des solutions y de $a\,y''+b\,y'+c\,y=0$.
L'application $\left\{\begin{array}{ccc} S & \to & \mathbb{R}^2 \\ y & \mapsto & (y(0),y'(0)) \end{array}\right.$ est un isomorphisme (la linéarité se vérifie immédiatement et le caractère bijectif provient de la résolution du problème de Cauchy).

Munissons les ensembles d'applications linéaires d'opérations et regardons leur structure.

Proposition 14.18.

Soit E et F deux $\mathbb{K}$-espaces vectoriels.
▹ Une combinaison linéaire d'applications linéaires de E dans F est une application linéaire.
Ainsi, $\mathscr{L}(E,F)$ et E^* sont des $\mathbb{K}$-espaces vectoriels.
▹ La composition de deux endomorphismes de E est un endomorphisme de E.

Ainsi, $\mathscr{L}(E)$ est un anneau (la deuxième loi de composition interne étant la composition).
▹ La composition de deux automorphismes de E est un automorphisme de E. La bijection réciproque d'un automorphisme de E est un automorphisme de E.
Ainsi, $\mathrm{GL}(E)$ est un groupe (c'est le groupe des inversibles de l'anneau $\mathscr{L}(E)$).

Démonstration

- Soit $u,\ v \in \mathscr{L}(E,F)$, $\lambda \in \mathbb{K}$. Montrons que $u+\lambda v$ est une application linéaire. Pour tous $x, y \in E$, $\mu \in \mathbb{K}$,

$$\begin{aligned}(u+\lambda v)(x+\mu y) &= u(x+\mu y)+\lambda v(x+\mu y)\\ &= u(x)+\mu u(y)+\lambda v(x)+\lambda\mu v(y) \text{ car } u,\ v \text{ sont linéaires,}\\ &= (u+\lambda v)(x)+\mu(u+\lambda v)(y).\end{aligned}$$

- Soit $u,\ v \in \mathscr{L}(E)$. Pour tous $x, y \in E$, $\mu \in \mathbb{K}$,

$$\begin{aligned}u\circ v(x+\mu y) &= u(v(x)+\mu v(y)) && \text{car } v \text{ est linéaire,}\\ &= u\circ v(x)+\mu u\circ v(y) && \text{car } u \text{ est linéaire,}\end{aligned}$$

- Soit $u \in \mathrm{GL}(E)$. Pour tous $x,\ y \in E$, posons $x' = u^{-1}(x)$ et $y' = u^{-1}(y)$; pour tout $\mu \in \mathbb{K}$,

$$\begin{aligned}u^{-1}(x+\mu y) &= u^{-1}(u(x')+\mu u(y'))\\ &= u^{-1}(u(x'+\mu y'))\\ &= x'+\mu y' = u^{-1}(x)+\mu u^{-1}(y).\end{aligned}$$

Le reste des affirmations provient des caractérisations de sous-espace vectoriel (avec l'inclusion $\mathscr{L}(E,F)\subset F^E$), de sous-anneau ($\mathscr{L}(E)\subset E^E$) et de sous-groupe ($\mathrm{GL}(E)$ est inclus dans le groupe des bijections de E dans E). ■

Remarque

Il est souvent intéressant de calculer les puissances de composition d'un endomorphisme (on note $u^n = u\circ\ldots\circ u$ la composée n fois). Rappelons que dans l'anneau $\mathscr{L}(E)$, on peut appliquer la formule du binôme de Newton dès que deux éléments commutent.

Par exemple, considérons deux endomorphismes de $\mathbb{R}^2$: $u : (x,y)\mapsto(3x+y,3y)$ et $v : (x,y)\mapsto(y,0)$. Alors, $u = 3\mathrm{Id}+v$ et v et Id commutent donc, pour tout $n\in\mathbb{N}$, la formule du binôme donne

$$u^n = \sum_{k=0}^{n}\binom{n}{k}3^{n-k}v^k = 3^n\mathrm{Id}+n3^{n-1}v,$$

car v^2 est l'application linéaire nulle.

Montrons l'intérêt de ce calcul dans l'étude des suites $(x_n)_n$ et $(y_n)_n$ vérifiant, pour tout $n\in\mathbb{N}$, les relations de récurrence :

$$\begin{cases} x_{n+1} &= 3x_n + y_n\\ y_{n+1} &= 3y_n.\end{cases}$$

Pour tout $n\in\mathbb{N}$, $(x_{n+1},y_{n+1}) = u(x_n,y_n)$ donc $(x_n,y_n) = u^n(x_0,y_0) = (3^n x_0+n3^{n-1}y_0, 3^n y_0)$.

Notation

Le programme suggère la notation uv pour la composition $u\circ v$; toutefois, par souci de clarté pédagogique, on a conservé le symbole $\circ$ dans ce livre.

2.2. Images réciproques et directes de sous-espaces

Proposition 14.19.

Soit u une application linéaire entre deux $\mathbb{K}$-espaces vectoriels E et F.
▹ L'image directe par u d'un sous-espace vectoriel de E est un sous-espace vectoriel de F.
▹ L'image réciproque par u d'un sous-espace vectoriel de F est un sous-espace vectoriel de E.

Démonstration
▹ Soit E' un sous-espace vectoriel de E. Remarquons tout d'abord que $u(E')$ est une partie non vide de F car $0_F = u(0_E) \in u(E')$. Ensuite, pour tous $x, y \in E'$, $\lambda \in \mathbb{K}$,

$$u(x) + \lambda u(y) = u(x + \lambda y) \in u(E'),$$

car E' est stable par combinaison linéaire. Ainsi, $u(E')$ est un sous-espace vectoriel de F.
▹ Soit F' un sous-espace vectoriel de F. Remarquons tout d'abord que $u^{-1}(F')$ est une partie non vide de E car $u(0_E) = 0_F \in F'$. Ensuite, pour tous $x, y \in u^{-1}(F')$, $\lambda \in \mathbb{K}$,

$$u(x + \lambda y) = u(x) + \lambda u(y) \in F'$$

donc $u^{-1}(F')$ est stable par combinaison linéaire : c'est un sous-espace vectoriel de E. ■

Forts de cette proposition, nous pouvons introduire le noyau et l'image d'une application linéaire.

Définition 14.20.

Soit $u \in \mathscr{L}(E, F)$.
▹ Le noyau de u est le sous-espace vectoriel de E défini par

$$\operatorname{Ker} u = u^{-1}(\{0_F\}) = \{x \in E,\ u(x) = 0_F\}.$$

▹ L'image de u est le sous-espace vectoriel de F défini par $\operatorname{Im} u = u(E)$.

Proposition 14.21.

Soit $u \in \mathscr{L}(E, F)$.
▹ L'application u est injective si, et seulement si, $\operatorname{Ker} u = \{0_E\}$.
▹ L'application u est surjective si, et seulement si, $\operatorname{Im} u = F$.

La propriété est évidente pour la surjectivité. Montrons la caractérisation de l'injectivité.

Démonstration
Soit $x, y \in E$. Alors, $u(x) = u(y)$ si, et seulement si, $u(x - y) = 0_F$, c'est-à-dire $x - y \in \operatorname{Ker} u$. Ainsi, u est injective si, et seulement si, $\operatorname{Ker} u = \{0_E\}$. ■

Voici une application bien utile dans le chapitre suivant.

Proposition 14.22.

Soit $u \in \mathscr{L}(E, F)$ et E' un supplémentaire du noyau $\operatorname{Ker} u$. L'application u réalise un isomorphisme entre E' et $\operatorname{Im} u$.

Démonstration
Considérons l'application induite par u entre E' et $\operatorname{Im} u$:

$$\tilde{u} : \left\{ \begin{array}{ccc} E' & \longrightarrow & \operatorname{Im} u \\ x & \longmapsto & u(x) \end{array} \right.$$

Cette application linéaire est injective car $\operatorname{Ker} \tilde{u} = \operatorname{Ker} u \cap E' = \{0_E\}$ puisque E' est en somme directe avec $\operatorname{Ker} u$. Elle est aussi surjective car, pour tout $y \in \operatorname{Im} u$, il existe $x \in E$ tel que $y = u(x)$. Or, ce vecteur x s'écrit comme somme d'un élément du noyau et d'un élément de E' donc y est l'image d'un élément de E' : ainsi, $\operatorname{Im} \tilde{u} = \operatorname{Im} u$. ■

Une autre utilisation des noyaux accompagne la notion de commutation.

Proposition 14.23.

Soit u et v deux endomorphismes d'un $\mathbb{K}$-espace vectoriel E tels que $u \circ v = v \circ u$. Alors, $u(\operatorname{Ker} v) \subset \operatorname{Ker} v$ et $u(\operatorname{Im} v) \subset \operatorname{Im} v$.
Autrement dit, les sous-espaces $\operatorname{Ker} v$ et $\operatorname{Im} v$ sont stables par u.

Démonstration
▹ Pour tout $x \in \operatorname{Ker} v$, $v(u(x)) = u(v(x)) = u(0_E) = 0_E$ donc $u(x) \in \operatorname{Ker} v$.
▹ Pour tout $y \in \operatorname{Im} v$, il existe $x \in E$ tel que $y = v(x)$. Ainsi,

$$u(y) = u(v(x)) = v(u(x)) \in \operatorname{Im} v.$$

■

Proposition 14.24.

Soit u et v deux endomorphismes d'un $\mathbb{K}$-espace vectoriel E. Alors, $u \circ v = 0$ si, et seulement si, $\operatorname{Im} v \subset \operatorname{Ker} u$.

Démonstration
Séparons les deux implications.
▹ Si $u \circ v = 0$, alors pour tout $x \in E$, $u(v(x)) = 0_E$ donc $v(x) \in \operatorname{Ker} u$: $\operatorname{Im} v \subset \operatorname{Ker} u$.
▹ Si $\operatorname{Im} v \subset \operatorname{Ker} u$, alors, pour tout $x \in E$, $v(x) \in \operatorname{Im} v \subset \operatorname{Ker} u$ donc $u(v(x)) = 0_E$. ■

2.3. Équations linéaires avec second membre

Définition 14.25.

La translation de vecteur $y \in E$ est une application de E dans E définie par $x \longmapsto x + y$.
L'image d'un sous-espace vectoriel F par la translation de vecteur y est appelée sous-espace affine de direction F et notée $y + F$.

Notation
Les éléments d'un sous-espace affine (donc tous les éléments de E puisque E est un sous-espace affine de E) sont aussi appelés points et quelquefois notés avec des lettres majuscules pour conserver l'analogie avec la géométrie élémentaire. Ainsi, si $\mathscr{F} = A + F$, alors $M \in \mathscr{F}$ si, et seulement si, $\overrightarrow{AM} = M - A \in F$.

Ayant choisi un point Ω appelé origine, on peut identifier le point M et le vecteur $\overrightarrow{\Omega M}$.

Proposition 14.26.

Soit F, G deux sous-espaces vectoriels de E, x, $y \in E$. Si $x+F=y+G$, alors $F=G$.

Démonstration

Montrons une inclusion, l'autre suit par symétrie.

▹ Remarquons tout d'abord que $x \in x+F=y+G$; ainsi, il existe $z \in G$ tel que $x=y+z$.

▹ Pour tout $f \in F$, $x+f \in y+G$ donc il existe $g \in G$ tel que $x+f=y+g$ donc, avec le point précédent,

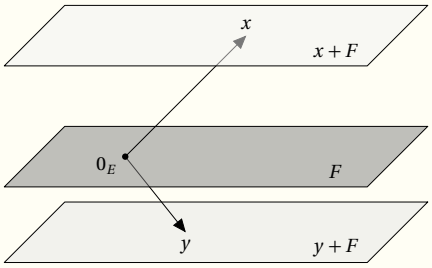
$$f=(x+f)-x=(y+g)-(y+z)=g-z \in G.$$

Ainsi, $F \subset G$. Par symétrie, $F=G$. ■

Remarque

D'après la proposition, la direction d'un sous-espace affine est bien définie comme l'unique sous-espace dont il est le translaté.

Voici un sous-espace vectoriel F et deux sous-espaces affines dont F est direction.



Voici un exemple essentiel de sous-espace affine.

Proposition 14.27.

Soit $u \in \mathscr{L}(E,F)$. L'équation $u(x)=b$ d'inconnue le vecteur x admet des solutions si, et seulement si, $b \in \operatorname{Im} u$; dans ce cas, l'ensemble des solutions est un sous-espace affine de direction $\operatorname{Ker} u$.

Démonstration

La première affirmation relève de la définition de l'image de u. Pour la seconde, supposons qu'il existe $x_0 \in E$ tel que $u(x_0)=b$; alors, $u(x)=b$ si, et seulement si, $u(x-x_0)=0_F$ soit $x-x_0 \in \operatorname{Ker} u$. ■

On retrouve les résultats de structure obtenus pour les équations différentielles linéaires...

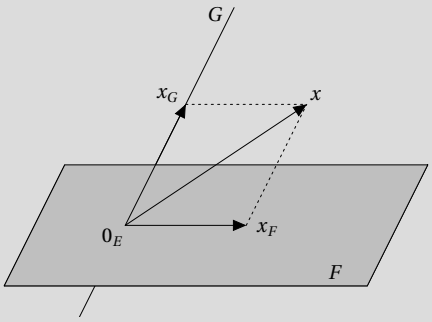
Remarque
L'intersection de sous-espaces affines est soit vide, soit un sous-espace affine. Plus précisément, si $\mathscr{F}$ et $\mathscr{G}$ sont deux sous-espaces affines de directions respectives F et G tels que $\mathscr{F}\cap\mathscr{G}\neq\varnothing$, alors $\mathscr{F}\cap\mathscr{G}$ est un sous-espace affine de direction $F\cap G$.
Montrons, pour tout $a\in\mathscr{F}\cap\mathscr{G}$, la double inclusion entre $a+F\cap G$ et $\mathscr{F}\cap\mathscr{G}$.
▹ Pour tout $x\in a+F\cap G$, $x\in a+F=\mathscr{F}$ et $x\in a+G=\mathscr{G}$ donc $x\in\mathscr{F}\cap\mathscr{G}$.
▹ Pour tout $x\in\mathscr{F}\cap\mathscr{G}$, il existe $u\in F$ et $v\in G$ tel que $x=a+u=a+v$ d'où $u=v\in F\cap G$ et $x\in a+F\cap G$.

3. Endomorphismes remarquables

3.1. Projecteurs

Définition 14.28.

Soit E un $\mathbb{K}$-espace vectoriel.
▹ Soit F et G deux sous-espaces vectoriels supplémentaires. La projection sur F parallèlement à G est l'application qui, au vecteur $x=x_F+x_G$ avec $(x_F,x_G)\in F\times G$, associe le vecteur x_F.



▹ Un projecteur de E est un endomorphisme p de E tel que $p\circ p=p$ (c'est-à-dire $p^2=p$ avec la notation des puissances pour la composition).

Montrons l'équivalence entre ces deux notions (l'une plus géométrique, l'autre plus algébrique).

Proposition 14.29.

Soit F et G deux sous-espaces vectoriels supplémentaires. La projection sur F parallèlement à G est un projecteur.

Démonstration
La linéarité est évidente par définition.
Pour tous $x_F\in F$ et $x_G\in G$,

$$p^2(x_F+x_G)=p(p(x_F+x_G))=p(x_F)=x_F=p(x_F+x_G).$$

Or, tout vecteur de E s'écrit comme somme d'un vecteur de F et d'un vecteur de G donc $p^2=p$. ■

Proposition 14.30.

Soit p un projecteur d'un espace vectoriel E.
▹ $\operatorname{Ker} p$ et $\operatorname{Im} p$ sont supplémentaires.
▹ p est la projection sur $\operatorname{Im} p$ parallèlement à $\operatorname{Ker} p$.

Démonstration

▹ Procédons par analyse-synthèse.

- Si $x = a + p(b)$ avec $a \in \operatorname{Ker} p$ et $b \in E$, alors $p(x) = p^2(b) = p(b)$ et $a = x - p(x)$.
- Réciproquement, pour tout $x \in E$, posons $a = x - p(x)$. On vérifie immédiatement que $a \in \operatorname{Ker} p$, $p(x) \in \operatorname{Im} p$ et $x = (x - p(x)) + p(x)$.

D'où $\operatorname{Ker} p$ et $\operatorname{Im} p$ sont supplémentaires.

▹ On a montré que si $x = a + p(b)$ avec $a \in \operatorname{Ker} p$ et $b \in E$, alors $p(x) = p^2(b) = p(b)$ donc p est la projection sur $\operatorname{Im} p$ parallèlement à $\operatorname{Ker} p$. ■

Proposition 14.31.

Soit p un projecteur de E. L'endomorphisme $q = \mathrm{Id} - p$ est alors un projecteur tel que
▹ $\operatorname{Ker} p = \operatorname{Im} q$ et $\operatorname{Im} p = \operatorname{Ker} q$;
▹ $p + q = \mathrm{Id}$;
▹ $p \circ q = q \circ p = 0$.
Le projecteur q est quelquefois appelé « projecteur complémentaire » de p.

Démonstration

Comme $p^2 = p$, $(\mathrm{Id} - p)^2 = \mathrm{Id} - 2p + p^2 = \mathrm{Id} - p$. Ainsi, q est un projecteur. La seule propriété à établir est $\operatorname{Ker} q = \operatorname{Im} p$ (l'autre égalité se déduit par symétrie).

- Si $x \in \operatorname{Ker}(\mathrm{Id} - p)$, alors $p(x) = x$ donc $x \in \operatorname{Im} p$.
- Si $y \in \operatorname{Im} p$, alors il existe $x \in E$ tel que $y = p(x)$ et $(p - \mathrm{Id})(y) = p^2(x) - p(x) = 0$: $y \in \operatorname{Ker}(\mathrm{Id} - p)$.

■

Remarque

D'après la proposition précédente, pour tout projecteur p, $\operatorname{Im} p = \operatorname{Ker}(p - \mathrm{Id})$. Cette égalité est intuitive car $\operatorname{Ker}(p - \mathrm{Id})$ est l'ensemble des éléments invariants par p (ce qui correspond bien à l'intuition : les éléments de l'espace sur lequel on projette sont invariants par la projection).

Exemple

Les espaces $\mathscr{P}$ et $\mathscr{I}$ des fonctions paires et impaires de $\mathbb{R}$ dans $\mathbb{R}$ sont supplémentaires. D'après les calculs de l'exemple 1.4., la projection sur $\mathscr{P}$ parallèlement à $\mathscr{I}$ est l'application

$$f \mapsto \left(x \mapsto \frac{f(x) + f(-x)}{2}\right).$$

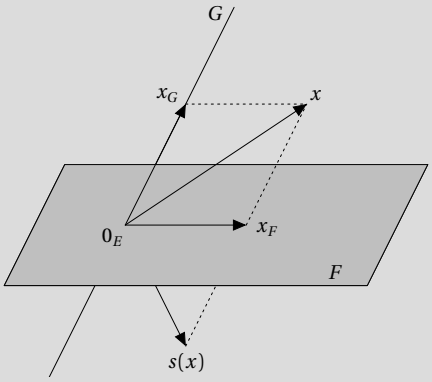
Remarque

Pour tout projecteur p, $p \circ (p - \mathrm{Id}) = 0$; par conséquent, le seul projecteur inversible vérifie $p - \mathrm{Id} = 0$, c'est-à-dire $p = \mathrm{Id}$: c'est la projection sur E parallèlement à $\{0_E\}$; le projecteur complémentaire est l'application nulle.

3.2. Symétries

Définition 14.32.

Soit E un $\mathbb{K}$-espace vectoriel.
▹ Soit F et G deux sous-espaces vectoriels supplémentaires. La symétrie par rapport à F parallèlement à G est l'application qui, au vecteur $x = x_F + x_G$ avec $(x_F, x_G) \in F \times G$, associe le vecteur $x_F - x_G$.



▹ Une symétrie vectorielle de E est un endomorphisme s de E tel que $s \circ s = \mathrm{Id}$ (on dit que s est involutive).

Comme précédemment avec les projecteurs, il y a équivalence entre les deux définitions.

Proposition 14.33.

Toute symétrie est une symétrie vectorielle.
Réciproquement, si s est une symétrie vectorielle, alors $\mathrm{Ker}(s-\mathrm{Id})$ et $\mathrm{Ker}(s+\mathrm{Id})$ sont supplémentaires et s est la symétrie par rapport à $\mathrm{Ker}(s-\mathrm{Id})$ parallèlement à $\mathrm{Ker}(s+\mathrm{Id})$.

Remarque

Il est facile de comprendre géométriquement pourquoi les espaces associés à une symétrie s sont $\mathrm{Ker}(s-\mathrm{Id})$ et $\mathrm{Ker}(s+\mathrm{Id})$. L'espace $\mathrm{Ker}(s-\mathrm{Id})$ est l'ensemble des vecteurs invariants par s ; l'espace $\mathrm{Ker}(s+\mathrm{Id})$ est l'ensemble des vecteurs changés en leurs opposés par s.

On aurait aussi pu retrouver ce résultat en explorant le lien entre symétrie et projection.

Proposition 14.34.

Soit E un $\mathbb{K}$-espace vectoriel et F et G deux sous-espaces vectoriels supplémentaires. L'endomorphisme p est la projection sur F parallèlement à G si, et seulement si, $2p-\mathrm{Id}$ est la symétrie par rapport à F parallèlement à G.

Démonstration

Commençons par calculer $(2p-\mathrm{Id})^2 = 4p^2 - 4p + \mathrm{Id}$. Ainsi, $(2p-\mathrm{Id})^2 = \mathrm{Id}$ si, et seulement si, $p^2 = p$. La

propriété sur les sous-espaces provient des remarques suivantes, pour tout $x \in E$,

$$p(x) = x \Leftrightarrow (2p - \mathrm{Id})(x) = x$$
$$p(x) = 0_E \Leftrightarrow (2p - \mathrm{Id})(x) = -x$$

■

Exemple

L'endomorphisme

$$s : \left\{ \begin{array}{ccc} \mathbb{R}^{\mathbb{R}} & \to & \mathbb{R}^{\mathbb{R}} \\ f & \mapsto & (x \mapsto f(-x)) \end{array} \right.$$

est la symétrie par rapport au sous-espace $\mathrm{Ker}(s - \mathrm{Id})$ des fonctions paires parallèlement au sous-espace $\mathrm{Ker}(s - \mathrm{Id})$ des fonctions impaires.

Exemple

L'endomorphisme

$$s : \left\{ \begin{array}{ccc} \mathbb{R}_n[X] & \to & \mathbb{R}_n[X] \\ P(X) & \mapsto & X^n P(\frac{1}{X}) \end{array} \right.$$

est une symétrie car linéaire et involutive (vérification immédiate). Déterminons les sous-espaces $\mathrm{Ker}(s - \mathrm{Id})$ et $\mathrm{Ker}(s + \mathrm{Id})$.

▹ $\sum_{k=0}^{n} a_k X^k \in \mathrm{Ker}(s - \mathrm{Id})$ si, et seulement si,

$$\sum_{k=0}^{n} a_k X^{n-k} = \sum_{k=0}^{n} a_k X^k,$$

c'est-à-dire si $a_k = a_{n-k}$ pour tout $k \in [\![0, n]\!]$.

▹ $\sum_{k=0}^{n} a_k X^k \in \mathrm{Ker}(s + \mathrm{Id})$ si, et seulement si,

$$\sum_{k=0}^{n} a_k X^{n-k} = -\sum_{k=0}^{n} a_k X^k,$$

c'est-à-dire si $a_k = -a_{n-k}$ pour tout $k \in [\![0, n]\!]$.

3.3. Homothéties

Définition 14.35.

Soit E un $\mathbb{K}$-espace vectoriel et $\lambda \in \mathbb{K}$. L'homothétie de rapport λ est l'endomorphisme de E défini par $x \mapsto \lambda x$.

On limite quelquefois l'appellation homothéties aux seules homothéties de rapport non nul.
Donnons quelques caractérisations classiques des homothéties.

Proposition 14.36.

Soit E un $\mathbb{K}$-espace vectoriel et u un endomorphisme de E tel que, pour tout $x \in E$, il existe $\lambda_x \in \mathbb{K}$ tel que $u(x) = \lambda_x . x$. Alors, u est une homothétie.

Démonstration
Soit x et y deux vecteurs de E. Alors, par linéarité de u,

$$\lambda_x.x + \lambda_y.y = \lambda_{x+y}.(x+y).$$

▹ Si $\mathrm{Vect}(x) = \mathrm{Vect}(y) \neq \{0_E\}$, alors x et y sont colinéaires non nuls et $\lambda_x = \lambda_y$.
▹ Si $\mathrm{Vect}(x)$ et $\mathrm{Vect}(y)$ sont en somme directe, alors, par unicité de l'écriture,

$$\lambda_x = \lambda_{x+y} = \lambda_y.$$

■

Remarque
En modifiant cette démonstration, on montre que les seuls endomorphismes $u \in \mathscr{L}(E)$ tels que les vecteurs x, $u(x)$ soient colinéaires pour tout $x \in E$ sont les homothéties.

Proposition 14.37.

Les endomorphismes d'un $\mathbb{K}$-espace vectoriel E commutant avec tous les endomorphismes de E sont les homothéties.

Démonstration
Soit u un endomorphisme commutant avec tous les endomorphismes de E. Pour tout $x \neq 0_E$, considérons p une projection sur $\mathrm{Vect}(x)$ (dont on suppose l'existence). Comme u commute avec p, u laisse stable $\mathrm{Im}\, p = \mathrm{Vect}(x)$. Ainsi, pour tout $x \in E$, il existe $\lambda_x \in \mathbb{K}$ tel que $u(x) = \lambda_x.x$. On conclut avec la proposition précédente. ■

FICHE

SYNTHÈSE

L'objet de ce chapitre est l'étude d'une nouvelle structure algébrique : les espaces vectoriels sur le corps $\mathbb{R}$ ou $\mathbb{C}$.

Espaces et sous-espaces vectoriels

▹ Un $\mathbb{K}$-espace vectoriel est un ensemble E muni d'une loi de composition interne $+$ et d'une application $\left\{\begin{array}{ccc}\mathbb{K}\times E & \longrightarrow & E\\ (\lambda,x) & \longmapsto & \lambda.x\end{array}\right.$ telles que

- $(E,+)$ est un groupe abélien ;
- la loi externe est distributive sur les additions de E et de $\mathbb{K}$:

$$\forall(\lambda_1,\lambda_2)\in\mathbb{K}^2,\ \forall(x_1,x_2)\in E^2,\qquad \begin{array}{ll}(\lambda_1+\lambda_2).x_1 & =\lambda_1.x_1+\lambda_2.x_1\\ \lambda_1.(x_1+x_2) & =\lambda_1.x_1+\lambda_1.x_2;\end{array}$$

- pour tout $x\in E$, $1.x=x$;
- $\forall(\lambda_1,\lambda_2)\in\mathbb{K}^2,\ \forall x\in E,\qquad \lambda_1.(\lambda_2.x)=(\lambda_1\lambda_2).x$.

▹ Une partie d'un espace vectoriel est un sous-espace vectoriel si elle est non vide et stable par combinaison linéaire.
▹ Le sous-espace engendré par une partie A est le plus petit sous-espace vectoriel contenant A ; c'est l'ensemble des combinaisons linéaires de vecteurs de A.
▹ Soit E un $\mathbb{K}$-espace vectoriel et $(E_i)_{i\in[1,p]}$ des sous-espaces vectoriels de E.

- La somme des sous-espaces $(F_i)_{i\in[1,p]}$ est le sous-espace vectoriel

$$F_1+\ldots+F_p=\{x_1+\ldots+x_p,\ (x_1,\ldots,x_p)\in F_1\times\ldots\times F_p\}$$

- La somme directe des sous-espaces $(F_i)_{i\in[1,p]}$ est le sous-espace $F_1+\ldots+F_p$ lorsque pour tout $x\in F_1+\ldots+F_p$, il existe un unique $(x_1,\ldots,x_p)\in F_1\times\ldots\times F_p$ tel que $x=x_1+\ldots+x_p$. La somme est alors notée $F_1\oplus\ldots\oplus F_p$.

Le « transport » de cette structure se fait avec les applications linéaires.

Application linéaire

▹ Une application $u:E\longrightarrow F$ est linéaire si l'image d'une combinaison linéaire est la combinaison linéaire des images avec les mêmes coefficients.
▹ L'injectivité et la surjectivité d'une application linéaire sont étudiées avec le noyau

$$\operatorname{Ker} u=\{x\in E,\ u(x)=0_F\}$$

et l'image $\operatorname{Im} u=u(E)$ respectivement.
▹ Un endomorphisme est une application linéaire d'un espace dans lui-même, une forme linéaire une application linéaire d'un espace dans son corps de base.

Tout au long du chapitre, on a défini des propriétés d'existence et d'unicité. Regroupons-les (en attendant d'y ajouter au chapitre suivant les notions de famille génératrice et de famille libre).

Existence et unicité

▹ Existence

- Une application linéaire $u : E \longrightarrow F$ est surjective si pour tout $y \in F$, il existe $x \in E$ tel que $y = u(x)$, c'est-à-dire $\operatorname{Im} u = F$.
- Un espace E est la somme des sous-espaces $F_1, \ldots, F_n$ si, pour tout $x \in E$, il existe $x_1 \in F_1$, $\ldots$, $x_n \in F_n$ tel que $x = \sum_{k=1}^{n} x_k$. C'est la définition précédente avec

$$u : \left\{ \begin{array}{ccc} F_1 \times \ldots \times F_n & \longrightarrow & E \\ (x_1, \ldots, x_n) & \longmapsto & \sum_{k=1}^{n} x_k \end{array} \right.$$

▹ Unicité

- Une application linéaire $u : E \longrightarrow F$ est injective si pour tout $y \in F$, il existe au plus un $x \in E$ tel que $y = u(x)$, c'est-à-dire $\operatorname{Ker} u = \{0_E\}$.
- Des sous-espaces $F_1, \ldots, F_n$ sont en somme directe si, pour tout $x \in E$, il existe au plus un $(x_1, \ldots, x_n) \in F_1 \times \ldots \times F_n$ tel que $x = \sum_{k=1}^{n} x_k$.

On retient aussi quelques endomorphismes particuliers qui vérifient une propriété algébrique qui s'interprète géométriquement.

Endomorphismes remarquables

▹ Un endomorphisme p est une projection si $p^2 = p$.
Les sous-espaces $\operatorname{Ker} p$ et $\operatorname{Im} p = \operatorname{Ker}(p - \mathrm{Id})$ sont supplémentaires et p est la projection géométrique sur le deuxième parallèlement au premier.
▹ Un endomorphisme s est une symétrie si $s^2 = \mathrm{Id}$.
Les sous-espaces $\operatorname{Ker}(s + \mathrm{Id})$ et $\operatorname{Ker}(s - \mathrm{Id})$ sont supplémentaires et s est la symétrie géométrique par rapport au deuxième parallèlement au premier.
▹ Un endomorphisme h est une homothétie s'il existe $\lambda \in \mathbb{K}$ tel que $h = \lambda \mathrm{Id}$.

EXERCICES

Vrai ou faux ?

		Vrai	Faux
a)	Soit F, G deux sous-espaces vectoriels de E tels que $F+G=F\cap G$. Alors, $F=G$.	☐	☐
b)	Soit F, G, H trois sous-espaces vectoriels de E en somme directe. Alors, $F\cap G=F\cap H=G\cap H=\{0_E\}$.	☐	☐
c)	Soit A, B deux parties de E. Alors, $\mathrm{Vect}(A\cap B)=\mathrm{Vect}\,A\cap\mathrm{Vect}\,B$.	☐	☐
d)	Soit $A\subset E$ et $u\in\mathscr{L}(E)$. Alors, $u(\mathrm{Vect}(A))=\mathrm{Vect}\,u(A)$.	☐	☐
e)	Si G et H sont supplémentaires dans E et F est un sous-espace de E, alors $G\cap F$ et $H\cap F$ sont des supplémentaires dans F.	☐	☐
f)	Si $u\in\mathscr{L}(E)$, alors $\mathrm{Im}\,u$ et $\mathrm{Ker}\,u$ sont supplémentaires.	☐	☐
g)	Si $(u,v)\in(\mathscr{L}(E))^2$, alors $\mathrm{Im}(u+v)=\mathrm{Im}(u)+\mathrm{Im}(v)$.	☐	☐
h)	Si $u\in\mathscr{L}(E)$ et G, H sont deux sous-espaces de E, alors $u(G+H)=u(G)+u(H)$.	☐	☐
i)	Soit $(u,v)\in(\mathscr{L}(E))^2$ tel que $u\circ v=0$. Alors, $u=0$ ou $v=0$.	☐	☐
j)	L'application $P\mapsto\frac{1}{2}(P(X)+P(-X))$ est le projecteur sur l'espace des polynômes pairs, parallèlement à l'espace des polynômes impairs.	☐	☐
k)	La somme de deux projecteurs est un projecteur.	☐	☐
l)	p est un projecteur si, et seulement si, $\mathrm{Id}-p$ est un projecteur.	☐	☐
m)	Si p est un projecteur, alors $p-2\mathrm{Id}$ est une symétrie.	☐	☐
n)	Si p est un projecteur, alors $\mathrm{Im}\,p=\mathrm{Ker}(p-\mathrm{Id})$.	☐	☐
o)	Si s est une symétrie, alors $\mathrm{Im}\,s=\mathrm{Ker}(s-\mathrm{Id})$.	☐	☐

Exercices d'application du chapitre 14

●○○ **Exercice 1**

1. L'ensemble des suites réelles bornées est-il un sous-espace vectoriel de $\mathbb{R}^{\mathbb{N}}$?
2. L'ensemble des suites réelles monotones est-il un sous-espace vectoriel de $\mathbb{R}^{\mathbb{N}}$?
3. L'ensemble des suites réelles somme d'une suite croissante et d'une suite décroissante est-il un sous-espace vectoriel de $\mathbb{R}^{\mathbb{N}}$?

●○○ **Exercice 2**

Déterminer parmi les ensembles suivants ceux qui sont des sous-espaces vectoriels de $\mathbb{R}^{\mathbb{R}}$.

1. l'ensemble des fonctions réelles 1-lipschitziennes ;
2. l'ensemble des fonctions réelles f telles que
$$\exists k \in \mathbb{R},\ \forall x \in \mathbb{R}, \qquad |f(x)| \leq k|x|,$$
3. l'ensemble des fonctions réelles f telles que
$$\exists k \in \mathbb{R},\ \forall x \in \mathbb{R}, \qquad |f(x)| \geq k|x|.$$

●○○ **Exercice 3**

Soit $E = \mathscr{C}^0(\mathbb{R},\mathbb{R})$ l'espace vectoriel des fonctions continues de $\mathbb{R}$ dans $\mathbb{R}$. Pour tout $a \in \mathbb{R}$, posons $E_a = \{f \in E,\ f(a) = 0\}$.

1. Montrer, que pour tout $a \in \mathbb{R}$, E_a est un sous-espace vectoriel de E.
2. Soit $a \neq b$. Montrer que $E = E_a + E_b$.
3. La somme de E_a et de E_b peut-elle être directe ?

●○○ **Exercice 4**

Soit F et G deux sous-espaces vectoriels d'un $\mathbb{K}$-espace vectoriel E tels que $F + G = E$. Notons F' un supplémentaire de $F \cap G$ dans F. Montrer que $E = F' \oplus G$.

●○○ **Exercice 5**

Soit E un espace vectoriel et F, G et H trois sous-espaces vectoriels de E. Montrer que la somme $F+G+H$ est directe si, et seulement si, $F \cap G = \{0_E\}$ et $(F+G) \cap H = \{0_E\}$.

●○○ **Exercice 6**

Soit F un sous-espace vectoriel strict de E.

1. Déterminer $\mathrm{Vect}(E \setminus F)$.
2. Déterminer tous les endomorphismes nuls sur le complémentaire de F.

●●○ **Exercice 7**

Soit $u \in \mathscr{L}(E,F)$, E_1 et E_2 des sous-espaces de E en somme directe.

1. Montrer que si u est injective, alors $u(E_1)$ et $u(E_2)$ sont en somme directe.
2. Étudier la réciproque.

●○○ **Exercice 8**

Soit E, F deux espaces vectoriels et $u \in \mathscr{L}(E,F)$. Montrer que $\Phi : (x,y) \mapsto (x, y - u(x))$ est un automorphisme de $E \times F$.

●○○ **Exercice 9**

Soit E un espace vectoriel, F un sous-espace de E et $u \in \mathscr{L}(E)$.

1. Montrer que $u^{-1}(u(F)) = F + \mathrm{Ker}\, u$.
2. Déterminer $u(u^{-1}(F))$.

●●○ **Exercice 10**

Soit u et $v \in \mathscr{L}(E)$ tels que $u \circ v \circ u = u$ et $v \circ u \circ v = v$.

1. Montrer que $E = \mathrm{Ker}\, v \oplus \mathrm{Im}\, u$.
2. Montrer que $u(\mathrm{Im}\, v) = \mathrm{Im}\, u$.

●●● **Exercice 11**

Soit E un espace vectoriel et $u \in \mathscr{L}(E)$.

1. Montrer que la suite des noyaux itérés $(\operatorname{Ker} u^n)_{n\in\mathbb{N}}$ est croissante pour l'inclusion.
2. Montrer que la suite des images itérées $(\operatorname{Im} u^n)_{n\in\mathbb{N}}$ est décroissante pour l'inclusion.
3. Établir l'équivalence entre $\operatorname{Ker} u = \operatorname{Ker} u^2$ et $\operatorname{Im} u \cap \operatorname{Ker} u = \{0_E\}$.
4. Établir l'équivalence entre $\operatorname{Im} u = \operatorname{Im} u^2$ et $E = \operatorname{Ker} u + \operatorname{Im} u$.

●●● **Exercice 12**

Soit u un endomorphisme de E, $\lambda \neq \mu$ deux scalaires.

Montrer que les sous-espaces $\operatorname{Ker}(u-\lambda\operatorname{Id})^2$ et $\operatorname{Ker}(u-\mu\operatorname{Id})^2$ sont en somme directe.

●○○ **Exercice 13**

Soit p un projecteur de E. Définissons les sous-espaces

$$F_1 = \{f \in \mathscr{L}(E),\ \exists u \in \mathscr{L}(E),\ f = u \circ p\}, \qquad F_2 = \{f \in \mathscr{L}(E),\ \exists u \in \mathscr{L}(E),\ f = u \circ (\operatorname{Id} - p)\}.$$

Montrer que ces deux sous-espaces sont supplémentaires de $\mathscr{L}(E)$.

●●○ **Exercice 14**

Soit E, F deux espaces vectoriels, $u \in \mathscr{L}(E,F)$ et $v \in \mathscr{L}(F,E)$ tels que $v \circ u = \operatorname{Id}_E$.

Montrer que $\operatorname{Ker} v \oplus \operatorname{Im} u = F$.

●○○ **Exercice 15**

Soit E, F deux $\mathbb{R}$-espaces vectoriels, $u \in \mathscr{L}(E,F)$ et V, W deux sous-espaces vectoriels de E. Montrer que $u(V) \subset u(W)$ si, et seulement si, $V + \operatorname{Ker} u \subset W + \operatorname{Ker} u$.

●●● **Exercice 16**

Soit E et F deux $\mathbb{K}$-espaces vectoriels, $\phi \in \mathscr{L}(E)$, et $u \in \mathscr{L}(E,F)$. Montrer qu'il existe $v \in \mathscr{L}(E,F)$ tel que $u = v \circ \phi$ si, et seulement si, $\operatorname{Ker}\phi \subset \operatorname{Ker} u$.

On admettra ici que tout sous-espace vectoriel de E et de F admet un supplémentaire.

●●○ **Exercice 17**

Soit a et b deux scalaires distincts et $u \in \mathscr{L}(E)$ tels que $(u - a\operatorname{Id}) \circ (u - b\operatorname{Id}) = 0$.

1. Montrer que $\operatorname{Ker}(u - a\operatorname{Id})$ et $\operatorname{Ker}(u - b\operatorname{Id})$ sont supplémentaires.
2. Déterminer une expression simple de la projection p sur $\operatorname{Ker}(u - a\operatorname{Id})$ parallèlement à $\operatorname{Ker}(u - b\operatorname{Id})$
3. Montrer que $u = ap + b(\operatorname{Id} - p)$. En déduire l'expression de u^n pour tout $n \in \mathbb{N}$.

●●○ **Exercice 18**

Considérons deux projections p et q sur le même sous-espace G (mais de directions différentes) et $\lambda \in \mathbb{R}$. Montrer que $\lambda p + (1-\lambda)q$ est une projection sur G.

●○○ **Exercice 19**

Soit E un $\mathbb{K}$-espace vectoriel, p et q deux projecteurs de E qui commutent. Montrer que

$$\operatorname{Ker}(p+q) = \operatorname{Ker} p \cap \operatorname{Ker} q.$$

EXERCICES

Exercices d'approfondissement du chapitre 14

●○○ **Exercice A**

Soit x, y deux vecteurs d'un $\mathbb{K}$-espace vectoriel E et F un sous-espace vectoriel de E. Montrer que $F + \mathrm{Vect}(x) = F + \mathrm{Vect}(y)$ si, et seulement s'il existe $z \in F$ et $(\alpha, \beta) \in \mathbb{K}^2$ tels que $\alpha\beta \neq 0$ et $z + \alpha x + \beta y = 0_E$.

●●● **Exercice B**

Soit E un $\mathbb{R}$-espace vectoriel et $F_1, \ldots, F_n$ des sous-espaces vectoriels de E strictement inclus dans E. Supposons que $F_n \not\subset F_1 \cup \ldots \cup F_{n-1}$ et choisissons $x \in F_n \setminus (F_1 \cup \ldots \cup F_{n-1})$ et $y \in E \setminus F_n$.

1. Montrer que $\lambda x + y \notin F_n$ pour tout $\lambda \in \mathbb{R}$.
2. Montrer que, pour tout $i \in [\![1, n-1]\!]$, il existe au plus un $\lambda \in \mathbb{R}$ tel que $\lambda x + y \in F_i$.
3. Conclure que $\bigcup_{k \in [1,n]} F_k \neq E$.

●●○ **Exercice C**

Soit u, v et $w \in \mathscr{L}(E)$ tels que $u \circ v = w$, $v \circ w = u$ et $w \circ u = v$.

1. Montrer que $\mathrm{Ker}\, u = \mathrm{Ker}\, v = \mathrm{Ker}\, w$ et $\mathrm{Im}\, u = \mathrm{Im}\, v = \mathrm{Im}\, w$.
2. Montrer que $u^5 = u$.
3. Vérifier que $E = \mathrm{Ker}\, u \oplus \mathrm{Im}\, u$.

●●● **Exercice D**

1. Soit $\Phi \in \mathscr{L}(\mathbb{C}[X])$ tel que $\Phi(P)' = \Phi(P')$ pour tout $P \in \mathbb{C}[X]$. Montrer qu'il existe une suite $(a_n)_n$ telle que
$$\forall P \in \mathbb{C}[X], \qquad \Phi(P) = \sum_{k=0}^{\infty} a_k P^{(k)}.$$
2. Trouver la suite $(a_n)_n$ pour $\Phi : P(X) \mapsto P(X+1)$.

●●○ **Exercice E**

Soit E un $\mathbb{K}$-espace vectoriel, p et q deux projecteurs de E. Montrer que $p + q$ est un projecteur si, et seulement si, $p \circ q = q \circ p = 0$.

●●○ **Exercice F**

Soit E un espace vectoriel, F, G et H des sous-espaces de E tels que $E = F \oplus G = F \oplus H$. Montrer que G et H sont isomorphes.

Problèmes du chapitre 14

●○○ **Problème 1**

Dans ce court problème, on cherche à étudier des sous-espaces qui admettent un supplémentaire commun.

Partie A – Premiers exemples

1. Soit $E = \mathbb{R}[X]$ l'ensemble des polynômes réels, P et $Q \in E$ de degré p et q, E_P et E_Q les ensembles des multiples de P et Q respectivement.
   a. Montrer que E_P est un espace vectoriel.

b. Justifier que $E = E_P \oplus \mathbb{R}_{p-1}[X]$.
c. En déduire que si $p = q$, alors E_P et E_Q admettent un supplémentaire commun.

2. Soit E un espace vectoriel.
   a. Soit F un sous-espace vectoriel et $x \in E$ non nul tels que $E = F \oplus \mathrm{Vect}(x)$. Montrer que pour tout $y \notin F$, $E = F \oplus \mathrm{Vect}(y)$.
   b. Soit F_1, F_2 des sous-espaces vectoriels et $x_1 \in F_2 \setminus F_1$, $x_2 \in F_1 \setminus F_2$ tels que $E = F_1 \oplus \mathrm{Vect}(x_1) = F_2 \oplus \mathrm{Vect}(x_2)$. Déterminer une droite vectorielle qui est un supplémentaire commun à F_1 et F_2.

Partie B – Exemple avec projecteur

Notons $E = \mathbb{R}[X]$ l'ensemble des polynômes réels, $\mathscr{P}$ et $\mathscr{I}$ les sous-espaces vectoriels des polynômes pairs et impairs respectivement.

1. Montrer que $\mathscr{I}$ est un supplémentaire de $\mathscr{P}$ dans E.
2. Soit l'application linéaire

$$\varphi : \left\{ \begin{array}{ccc} E & \longrightarrow & E \\ P & \longmapsto & \frac{P(X)+P(-X)}{2} + X\,\frac{P(X)-P(-X)}{2} \end{array} \right.$$

   a. Déterminer $\mathrm{Im}\,\varphi$ puis établir que

$$\mathrm{Ker}\,\varphi = \{(1-X)P(X),\ P \in \mathscr{I}\}.$$

   b. Montrer que φ est un projecteur de E.
   c. En déduire que $\mathrm{Ker}\,\varphi$ est un supplémentaire de $\mathscr{P}$.

Partie C – Condition suffisante

Soit E un espace vectoriel, F_1 et F_2 deux sous-espaces vectoriels de E

1. Supposons, dans cette question, que F_1 et F_2 sont supplémentaires dans E et qu'il existe un isomorphisme $u : F_1 \to F_2$.
   Montrer que $G = \{x - u(x), x \in F_1\}$ est un espace vectoriel puis qu'il est un supplémentaire commun à F_1 et F_2.
2. Réciproquement supposons dans cette question que F_1 et F_2 admettent un supplémentaire commun G.
   a. Montrer que F_1 et F_2 sont isomorphes.
   b. Exhiber des contre-exemples pour chacune des deux propriétés suivantes
      - $F_1 \cap F_2 = \{0\}$,
      - $F_1 + F_2 = E$.

PROBLÈMES

CORRIGÉS

Corrigés des Vrai/Faux

a) Vrai.

b) Vrai.

c) Faux, avec par exemple deux parties disjointes : $A = \{x\}$, $B = \{2x\}$.

d) Vrai.

e) Faux, par exemple pour $F \cap G = F \cap H = \{0_E\}$.

f) Faux.

g) Faux, si $u = -v \neq 0$, alors $\mathrm{Im}(u+v) = \{0_E\}$ et $\mathrm{Im}\, u + \mathrm{Im}\, v = \mathrm{Im}\, u \neq \{0\}$.

h) Vrai.

i) Faux.

j) Vrai.

k) Faux, par exemple avec Id et Id.

l) Vrai.

m) Faux, $(p - 2\mathrm{Id})^2 = 4\mathrm{Id} - p$.

n) Vrai.

o) Faux, $\mathrm{Im}\, s$ est tout l'espace car s est bijective alors que $\mathrm{Ker}(s - \mathrm{Id})$ est l'espace des vecteurs invariants par s. L'égalité n'a lieu que pour $s = \mathrm{Id}$.

Corrigés des exercices d'application du chapitre 14

Exercice 1

1. L'ensemble est évidemment non vide. Soit $(u_n)_n$ et $(v_n)_n$ des suites bornées par M_u et M_v respectivement, $\lambda, \mu \in \mathbb{R}$. Alors, pour tout n, $|\lambda u_n + \mu v_n| \leq |\lambda| M_u + |\mu| M_v$. Ainsi, toute combinaison linéaire de suites bornées est bornée. L'ensemble est un espace vectoriel.
2. Considérons $(u_n)_n$ et $(v_n)_n$ définies par

$$u_0 = 0,\ u_1 = 1,\ u_n = 3 \text{ pour tout } n \geq 2,$$
$$v_0 = 0,\ v_n = -2 \text{ pour tout } n \geq 1.$$

Ces deux suites sont monotones et pourtant $u_0 + v_0 = 0$, $u_1 + v_1 = -1$, $u_2 + v_2 = 1$ donc $(u_n + v_n)_n$ n'est pas monotone : cet ensemble n'est pas un espace vectoriel.
3. Oui : cet ensemble est non vide et une combinaison linéaire de telles suites est bien dans l'ensemble (on regroupe les suites croissantes et décroissantes en discutant selon le signe des scalaires).

Exercice 2

1. Non, 2Id n'est pas dans l'ensemble alors que Id y appartient.
2. Cet ensemble est non vide (fonction nulle). Pour tous f, g dans cet ensemble, il existe k et k' tels que, pour tout $x \in \mathbb{R}$,

$$|f(x)| \leq k|x|, \qquad |g(x)| \leq k'|x|.$$

Alors, pour tous $\lambda, \mu \in \mathbb{R}$,

$$\forall x \in \mathbb{R}, \qquad |\lambda f(x) + \mu g(x)| \leq (|\lambda|k + |\mu|k')|x|.$$

La fonction $\lambda f + \mu g$ appartient à l'ensemble qui est donc un espace vectoriel.
3. Cet espace est $\mathbb{R}^{\mathbb{R}}$: toutes les fonctions vérifient la propriété pour $k = 0$.

Exercice 3

1. Soit $a \in \mathbb{R}$. La partie E_a de E est non vide car elle contient la fonction nulle. Par ailleurs, pour tous $f, g \in E_a$ et tout $\lambda \in \mathbb{R}$, $(\lambda f + g)(a) = \lambda f(a) + g(a) = 0$ donc $\lambda f + g \in E_a$. D'après la caractérisation des sous-espaces vectoriels, E_a est un sous-espace vectoriel de E.
2. Pour tout $f \in E$,

$$f(x) = \underbrace{\frac{x-a}{b-a} f(x)}_{\in E_a} + \underbrace{\frac{b-x}{b-a} f(x)}_{\in E_b}.$$

D'où $E \subset E_a + E_b$ l'autre inclusion étant évidente puisque E_a et E_b sont des parties de E.
3. La somme de E_a et E_b pour $a \neq b$ ne peut être directe car l'intersection contient des fonctions non nulles : par exemple $x \mapsto (x-a)(x-b)$.

Exercice 4

▹ Soit $x \in E$; considérons $f \in F$ et $g \in G$ tels que $x = f + g$. Comme $f \in F = F' \oplus (F \cap G)$, il existe $f' \in F'$ et $g' \in F \cap G$ tels que $f = f' + g'$ et donc

$$x = f' + (g' + g) \in F' + G.$$

Par conséquent, $E \subset F' + G$ l'autre inclusion étant évidente car F' et G sont des parties de E.
▹ Par ailleurs, $F' \cap G = (F' \cap F) \cap G = F' \cap (F \cap G) = \{0_E\}$ donc F' et G sont en somme directe.
En combinant les deux points, on obtient $E = F' \oplus G$.

Exercice 5

($\Rightarrow$) Supposons que la somme $F + G + H$ est directe.

- Si $x \in F \cap G$, alors

$$\underset{\in F}{x} + \underset{\in G}{0_E} + \underset{\in H}{0_E} = \underset{\in F}{0_E} + \underset{\in G}{x} + \underset{\in H}{0_E},$$

donc $x = 0_E$ par unicité de la décomposition. Ainsi $F \cap G = \{0_E\}$.
- Soit $f \in F$ et $g \in G$ tels que $f + g \in (F + G) \cap H$; alors

$$\underset{\in F}{f} + \underset{\in G}{g} + \underset{\in H}{0_E} = \underset{\in F}{0_E} + \underset{\in G}{0_E} + \underset{\in H}{(f+g)},$$

donc $f = g = 0_E$ par unicité de la décomposition. Ainsi $(F + G) \cap H = \{0_E\}$.
($\Leftarrow$) Supposons $F \cap G = \{0_E\}$ et $(F + G) \cap H = \{0_E\}$ et considérons $f_1, f_2 \in F$, $g_1, g_2 \in G$, $h_1, h_2 \in H$ tels que $f_1 + g_1 + h_1 = f_2 + g_2 + h_2$. Alors, le vecteur $(f_1 - f_2) + (g_1 - g_2) = h_2 - h_1 \in (F + G) \cap H$ donc est nul. Par suite, le vecteur $f_1 - f_2 = g_2 - g_1 \in F \cap G$ est nul aussi. En conclusion, $f_1 = f_2$, $g_1 = g_2$ et $h_1 = h_2$: il y a unicité de la décomposition donc la somme $F + G + H$ est directe.

Exercice 6

1. Montrons que $\mathrm{Vect}(E \setminus F) = E$. Comme $E \setminus F \subset \mathrm{Vect}(E \setminus F)$, il suffit de montrer que $F \subset \mathrm{Vect}(E \setminus F)$. Considérons $x \in F$ et $y \notin F$. Alors, $x = (x-y)+y$ et les vecteurs $x-y$ et y n'appartiennent pas à F donc $x \in \mathrm{Vect}(E \setminus F)$ ce qui termine la preuve.
2. Si un endomorphisme u est nul sur $E \setminus F$, il l'est également sur $\mathrm{Vect}(E \setminus F) = E$: c'est l'endomorphisme nul.

Exercice 7

1. Utilisons la caractérisation de somme-directe de deux sous-espaces par l'intersection nulle.
Soit $y \in u(E_1) \cap u(E_2)$. Il existe $x_1 \in E_1$ et $x_2 \in E_2$ tels que $y = u(x_1) = u(x_2)$. Par injectivité de u, $x_1 = x_2 \in E_1 \cap E_2$. Comme $E_1 \cap E_2 = \{0_E\}$, $x_1 = 0_E$ donc $y = 0_F$.
2. La réciproque est fausse en général. On peut construire un contre-exemple dans $\mathbb{R}^3$ avec un endomorphisme u défini par $u((x,y,z)) = (x,y,0)$.
Cet endomorphisme n'est pas injectif car $\mathrm{Ker}\, u = \mathrm{Vect}((0,0,1))$.
En revanche, les droites $E_1 = \mathrm{Vect}((1,0,0))$ et $E_2 = \mathrm{Vect}((0,1,0))$ sont en somme directe ; par ailleurs, $u(E_1) = E_1$ et $u(E_2) = E_2$ donc $u(E_1)$ et $u(E_2)$ sont en somme directe.

Exercice 8

▹ Le caractère linéaire de Φ est une simple vérification.
▹ Fixons un couple $(x', y') \in E \times F$ et cherchons ses antécédents $(x,y) \in E \times F$ par Φ. La condition $(x', y') = \Phi((x,y))$ se réécrit $x' = x$ et $y' = y - u(x)$ soit encore $x = x'$ et $y = y' + u(x')$: il y a donc un et un seul antécédent par Φ. Ainsi, Φ est bijective donc un automorphisme.

Exercice 9

1. Soit $x \in E$. Alors,

$$\begin{aligned} x \in u^{-1}(u(F)) &\Leftrightarrow u(x) \in u(F) \\ &\Leftrightarrow \exists f \in F, \quad u(x) = u(f) \\ &\Leftrightarrow \exists f \in F, \quad u(x-f) = 0_E \\ &\Leftrightarrow \exists f \in F, \quad x - f \in \mathrm{Ker}\, u \\ &\Leftrightarrow x \in F + \mathrm{Ker}\, u \end{aligned}$$

2. Montrons que $u(u^{-1}(F)) = F \cap \mathrm{Im}\, u$.
▹ Soit $y \in u(u^{-1}(F))$. Il existe $x \in u^{-1}(F)$ tel que $y = u(x)$ donc $y \in \mathrm{Im}\, u$ et $y \in F$.
▹ Soit $y \in F \cap \mathrm{Im}\, u$. Comme $y \in \mathrm{Im}\, u$, il existe $x \in E$ tel que $y = u(x)$ et, comme $y \in F$, $x \in u^{-1}(F)$. Ainsi, $y \in u(u^{-1}(F))$.

Exercice 10

1. Procédons par analyse-synthèse.
▹ Soit $x \in E$ tel que $x = y + z$ avec $y \in \mathrm{Ker}\, v$ et $z \in \mathrm{Im}\, u$. Alors, $u \circ v(x) = u \circ v(z) = z$ car $u \circ v \circ u = u$. Ainsi, $z = u \circ v(x)$ et $y = x - z$.
▹ Soit $x \in E$. Posons $z = u \circ v(x)$ et $y = x - z$. Alors, $z \in \mathrm{Im}\, u$, $y + z = x$ et $v(y) = v(x) - v \circ u \circ v(x) = 0_E$.
En conclusion, $E = \mathrm{Ker}\, v \oplus \mathrm{Im}\, u$.
2. Par définition de l'image $u(\mathrm{Im}\, v) \subset u(E) = \mathrm{Im}\, u$. Par ailleurs, la relation $u \circ v \circ u = u$ entraîne que $\mathrm{Im}\, u = u \circ v(\mathrm{Im}\, u)$ donc

$$\mathrm{Im}\, u \subset u \circ v(E) = u(\mathrm{Im}\, v).$$

En conclusion, $u(\mathrm{Im}\, v) = \mathrm{Im}\, u$.

Exercice 11

1. Soit $n \in \mathbb{N}$. Pour tout $x \in \operatorname{Ker} u^n$, $u^{n+1}(x) = u(u^n(x)) = u(0_E) = 0_E$ donc $x \in \operatorname{Ker} u^{n+1}$. Par conséquent, $\operatorname{Ker} u^n \subset \operatorname{Ker} u^{n+1}$.
2. Soit $n \in \mathbb{N}$. Pour tout $y \in \operatorname{Im} u^{n+1}$, il existe $x \in E$ tel que $y = u^{n+1}(x) = u^n(u(x)) \in \operatorname{Im} u^n$. Par conséquent, $\operatorname{Im} u^{n+1} \subset \operatorname{Im} u^n$.
3. ($\Rightarrow$) Soit $y \in \operatorname{Im} u \cap \operatorname{Ker} u$. Il existe $x \in E$ tel que $y = u(x)$ et $u(y) = 0_E$. Par conséquent, $u^2(x) = 0_E$ donc $x \in \operatorname{Ker} u^2 = \operatorname{Ker} u$ et ainsi, $y = 0_E$.
Par conséquent, $\operatorname{Im} u \cap \operatorname{Ker} u = \{0_E\}$.
($\Leftarrow$) Soit $x \in E$.

$$\begin{aligned} x \in \operatorname{Ker} u^2 &\Leftrightarrow u^2(x) = 0_E \\ &\Leftrightarrow u(x) \in \operatorname{Im} u \cap \operatorname{Ker} u \\ &\Leftrightarrow u(x) = 0_E \\ &\Leftrightarrow x \in \operatorname{Ker} u. \end{aligned}$$

Par conséquent, $\operatorname{Ker} u = \operatorname{Ker} u^2$.
4. ($\Rightarrow$) Soit $x \in E$. Alors, $u(x) \in \operatorname{Im} u = \operatorname{Im} u^2$ et il existe $t \in E$ tel que $u(x) = u^2(t)$, c'est-à-dire $x - u(t) \in \operatorname{Ker} u$.

$$x = (x - u(t)) + u(t) \in \operatorname{Ker} u + \operatorname{Im} u.$$

Par conséquent, $E \subset \operatorname{Ker} u + \operatorname{Im} u$ et l'autre inclusion est évidente.
($\Leftarrow$) Soit $y \in \operatorname{Im} u$ et $x \in E$ tel que $y = u(x)$. Soit $x' \in \operatorname{Ker} u$ et $t \in E$ tel que $x = x' + u(t)$. Alors,

$$y = u(x) = u(x') + u^2(t) = u^2(t) \in \operatorname{Im} u^2.$$

Par conséquent, $\operatorname{Im} u = \operatorname{Im} u^2$.

Exercice 12

Utilisons la caractérisation de somme directe pour deux sous-espaces.
Soit $x \in \operatorname{Ker}(u - \lambda \operatorname{Id})^2 \cap \operatorname{Ker}(u - \mu \operatorname{Id})^2$. Alors,

$$\begin{aligned} 0_E &= (u - \lambda \operatorname{Id})^2(x) = u^2(x) - 2\lambda u(x) + x, \\ 0_E &= (u - \mu \operatorname{Id})^2(x) = u^2(x) - 2\mu u(x) + x. \end{aligned}$$

En effectuant la différence, on obtient $2(\lambda - \mu)u(x) = 0_E$ donc $u(x) = 0_E$ car $\lambda \neq \mu$. En reportant dans l'une des équations, on en déduit $x = 0_E$.
En conclusion, $\operatorname{Ker}(u - \lambda \operatorname{Id})^2 \cap \operatorname{Ker}(u - \mu \operatorname{Id})^2 = \{0_E\}$ donc les sous-espaces sont en somme directe.

Exercice 13

De manière évidente, pour tout $f \in \mathscr{L}(E)$, $f = f \circ p + f \circ (\operatorname{Id} - p)$. Il ne reste à montrer que la propriété de somme directe. Soit $g \in F_1 \cap F_2$. Alors, il existe $u, v \in \mathscr{L}(E)$ tels que

$$g = u \circ p = v \circ (\operatorname{Id} - p).$$

Par conséquent, $\operatorname{Ker} p \subset \operatorname{Ker} g$ et $\operatorname{Ker}(\operatorname{Id} - p) \subset \operatorname{Ker} g$ et donc $E = \operatorname{Ker} p \oplus \operatorname{Ker}(\operatorname{Id} - p) \subset \operatorname{Ker} g$: g est l'endomorphisme nul.
En conclusion, $\mathscr{L}(E) = F_1 \oplus F_2$.

Exercice 14

Montrons ce résultat élémentaire par analyse-synthèse.

▹ Soit $x = y + u(z)$ avec $y \in \operatorname{Ker} v$ et $z \in E$. Alors, en composant par v, on obtient

$$v(x) = v(y) + v \circ u(z) = z.$$

▹ Soit $x \in E$. Posons $z = v(x)$ et $y = x - u \circ v(x)$. Par construction,

- $x = y + u(z)$.
- $u(z) \in \operatorname{Im} u$.
- $v(y) = v(x) - v \circ u(v(x)) = v(x) - v(x) = 0_E$ donc $y \in \operatorname{Ker} v$.

En conclusion, $E = \operatorname{Ker} v \oplus \operatorname{Im} u$.

Exercice 15

($\Rightarrow$) Soit $v \in V$. Alors, $u(v) \in u(V) \subset u(W)$. Ainsi, il existe $w \in W$ tel que $u(v) = u(w)$, c'est-à-dire $v - w \in \operatorname{Ker} u$. Ainsi, $v = w + (v - w) \in W + \operatorname{Ker} u$.

Par conséquent, $V \subset W + \operatorname{Ker} u$ et comme $\operatorname{Ker} u \subset V + \operatorname{Ker} u$, on a bien $V + \operatorname{Ker} u \subset W + \operatorname{Ker} u$.

($\Leftarrow$) Soit $v \in V \subset W + \operatorname{Ker} u$. Considérons $w \in W$ et $x \in \operatorname{Ker} u$ tels que $v = w + x$. Alors, $u(v) = u(w) \in u(W)$. Ainsi, $u(V) \subset u(W)$.

Exercice 16

($\Rightarrow$) Le sens direct est évident. En effet, si $x \in \operatorname{Ker} \phi$, alors, pour tout $v \in \mathscr{L}(E, F)$, $v \circ \phi(x) = v(0_E) = 0_F$ donc $x \in \operatorname{Ker} v \circ \phi$.

($\Leftarrow$) Soit E' un supplémentaire de $\operatorname{Ker} \phi$ dans E et E'' un supplémentaire de $\operatorname{Im} \phi$ dans E. Alors,

$$\tilde{\phi} : \left\{ \begin{array}{ccc} E' & \to & \operatorname{Im} \phi \\ x & \mapsto & \phi(x) \end{array} \right.$$

est un isomorphisme. Considérons pour v l'unique application linéaire qui coïncide avec $u \circ \tilde{\phi}^{-1}$ sur $\operatorname{Im} \phi$ et avec l'application nulle sur E''.

Soit $x \in E$; considérons $x' \in E'$ et $t \in \operatorname{Ker} \phi$ tels que $x = x' + t$. Alors,

$$v \circ \phi(x) = v(\phi(x') + 0_E) = v(\tilde{\phi}(x')) = u \circ \tilde{\phi}^{-1}(\tilde{\phi}(x')) = u(x') = u(x),$$

car $t \in \operatorname{Ker} \phi \subset \operatorname{Ker} u$. Par conséquent, $u = v \circ \phi$.

Exercice 17

1. Remarquons que

$$\frac{1}{b-a}(u - a\operatorname{Id}) - \frac{1}{b-a}(u - b\operatorname{Id}) = \operatorname{Id}.$$

Ainsi, on a immédiatement que si $x \in \operatorname{Ker}(u - a\operatorname{Id}) \cap \operatorname{Ker}(u - b\operatorname{Id})$, alors $x = 0_E$ (en appliquant l'identité en x). Les deux sous-espaces $\operatorname{Ker}(u - a\operatorname{Id})$ et $\operatorname{Ker}(u - b\operatorname{Id})$ sont en somme directe.

Par ailleurs, pour tout $x \in E$,

$$x = \underbrace{\frac{1}{b-a}(u - a\operatorname{Id})(x)}_{\in \operatorname{Ker}(u - b\operatorname{Id})} + \underbrace{\frac{1}{a-b}(u - b\operatorname{Id})(x)}_{\in \operatorname{Ker}(u - a\operatorname{Id})} \in \operatorname{Ker}(u - b\operatorname{Id}) \oplus \operatorname{Ker}(u - a\operatorname{Id}),$$

car $(u - a\operatorname{Id}) \circ (u - b\operatorname{Id}) = 0$ et $(u - b\operatorname{Id}) \circ (u - a\operatorname{Id}) = 0$.

En conclusion, les deux sous-espaces $\operatorname{Ker}(u - a\operatorname{Id})$ et $\operatorname{Ker}(u - b\operatorname{Id})$ sont supplémentaires.

2. Dans la question précédente, on a explicité la décomposition d'un vecteur x quelconque donc déterminé la projection, c'est-à-dire l'application qui à x associe sa composante sur $\mathrm{Ker}(u-a\mathrm{Id})$:

$$p : x \mapsto \frac{1}{a-b}(u-b\mathrm{Id})(x).$$

3. La projection sur $\mathrm{Ker}(u-b\mathrm{Id})$ parallèlement à $\mathrm{Ker}(u-a\mathrm{Id})$ est $\mathrm{Id}-p$. Par conséquent, $u = u\circ\mathrm{Id} = u\circ(p+\mathrm{Id}-p) = u\circ p + u\circ(\mathrm{Id}-p)$. Or, par définition de $\mathrm{Ker}(u-a\mathrm{Id}) = \mathrm{Im}\,p$ (respectivement $\mathrm{Ker}(u-b\mathrm{Id}) = \mathrm{Im}(\mathrm{Id}-p)$), la restriction de u à ce sous-espace coïncide avec $a\mathrm{Id}$ (respectivement $b\mathrm{Id}$). Ainsi, $u = ap + b(\mathrm{Id}-p)$.
Remarquons $p\circ(\mathrm{Id}-p) = (\mathrm{Id}-p)\circ p = 0$ donc $u^n = a^n p^n + b^n(\mathrm{Id}-p)^n = (a^n-b^n)p + b^n\mathrm{Id}$.
L'ensemble de cette solution repose sur la première identité qui n'est autre qu'une traduction de l'identité de Bézout entre les polynômes $X-a$ et $X-b$.

Exercice 18

L'endomorphisme $u = \lambda p + (1-\lambda)q$ vérifie

$$\begin{aligned} u^2 - u &= \lambda^2 p^2 + (1-\lambda)^2 q^2 + \lambda(1-\lambda)p\circ q + \lambda(1-\lambda)q\circ p - \lambda p - (1-\lambda)q \\ &= \lambda(\lambda-1)p - (1-\lambda)\lambda q + 2\lambda(1-\lambda)p\circ q \text{ car } p\circ q = q\circ p \\ &= \lambda(\lambda-1)((\mathrm{Id}-q)\circ p + (\mathrm{Id}-p)\circ q). \end{aligned}$$

Or, en utilisant les projecteurs associés p et $\mathrm{Id}-p$, on a $\mathrm{Ker}(p-\mathrm{Id}) = \mathrm{Im}\,p = \mathrm{Im}\,q$ et, de même, $\mathrm{Ker}(\mathrm{Id}-q) = \mathrm{Im}\,q = \mathrm{Im}\,p$ donc $u^2 - u = 0$ et u est un projecteur.

Exercice 19

Il est clair que $\mathrm{Ker}\,p \cap \mathrm{Ker}\,q \subset \mathrm{Ker}(p+q)$. Montrons l'autre inclusion en considérant $x \in \mathrm{Ker}(p+q)$. L'identité $(p+q)^2 = p+q+2p\circ q$ donne en x, $p\circ q(x) = 0_E$. Or, comme $(p+q)(x) = 0_E$, $q(x) = -p(x)$ et la condition trouvée est $p^2(x) = 0_E$ soit encore $p(x) = 0_E$ car p est un projecteur. Ainsi, $x \in \mathrm{Ker}\,p$ et par symétrie, $x \in \mathrm{Ker}\,q$.

Corrigés des exercices d'approfondissement du chapitre 14

Exercice A

($\Rightarrow$) Par hypothèse $x \in \mathrm{Vect}(x) \subset F + \mathrm{Vect}(x) = F + \mathrm{Vect}(y)$ donc il existe $z \in F$ et $\beta \in \mathbb{K}$ tels que $x = z + \beta y$, c'est-à-dire $z - x + \beta y = 0_E$.
Si $\beta \neq 0$, on a bien exhibé une relation de la forme voulue ; sinon, $x = z \in F$ et donc $y \in F + \mathrm{Vect}(y) = F + \mathrm{Vect}(x) = F$: une relation est alors obtenue avec

$$\underbrace{x+y}_{\in F} - x - y = 0_E.$$

($\Leftarrow$) Montrons une inclusion, l'autre s'obtenant par symétrie. Avec la relation de l'énoncé, $x = -\frac{1}{\alpha}z - \frac{\beta}{\alpha}y \in F + \mathrm{Vect}(y)$ d'où $\mathrm{Vect}(x) \subset F + \mathrm{Vect}(y)$. Or, $F \subset F + \mathrm{Vect}(y)$ donc $F + \mathrm{Vect}(x) \subset F + \mathrm{Vect}(y)$.

Exercice B

1. Raisonnons par l'absurde et supposons qu'il existe $\lambda \in \mathbb{R}$ tel que $\lambda x + y \in F_n$. Alors, $y = (\lambda x + y) - \lambda x \in F_n$ car $x \in F_n$: contradiction.
2. Raisonnons encore par l'absurde et supposons qu'il existe $i \in [\![1, n-1]\!]$ et deux réels distincts λ et μ tels que $\lambda x + y \in F_i$ et $\mu x + y \in F_i$. Alors, $x = \frac{1}{\lambda - \mu}\big((\lambda x + y) - (\mu x + y)\big) \in F_i$, d'où la contradiction avec le choix de x.
3. Si $\bigcup\limits_{k \in [1,n]} F_k = E$, alors, pour tout $\lambda \in \mathbb{R}$, $\lambda x + y \in \bigcup\limits_{k \in [1,n-1]} F_k$ d'après la première question.

Or, d'après la deuxième question, il y a au plus $n-1$ réels λ tels que cette condition soit vérifiée : contradiction.

Dans cet exercice, nous avons montré qu'un $\mathbb{R}$-espace vectoriel n'est pas réunion d'un nombre fini de sous-espaces stricts.

Exercice C

1. Comme $u \circ v = w$, on a $\operatorname{Ker} v \subset \operatorname{Ker} w$. En procédant de même avec les autres égalités, on obtient

$$\operatorname{Ker} v \subset \operatorname{Ker} w \subset \operatorname{Ker} u \subset \operatorname{Ker} v$$

et donc l'égalité des trois noyaux. Le même raisonnement entraîne l'égalité des images.
2. On a successivement

$$\begin{aligned} u &= v \circ w = v \circ u \circ v = v \circ u \circ w \circ u \\ &= v \circ (v \circ w) \circ w \circ u = v \circ (w \circ u) \circ (u \circ v) \circ w \circ u \\ &= (v \circ w) \circ u \circ u \circ (v \circ w) \circ u = u^5. \end{aligned}$$

3. Procédons par analyse et synthèse.

▹ Soit $x \in E$ tel que $x = y + z$ avec $u(y) = 0_E$ et $z \in \operatorname{Im} u$. Alors, $u^4(x) = u^4(z) = z$ car $u^5 = u$. Ainsi, $z = u^4(x)$ et $y = x - z$.

▹ Soit $x \in E$. Posons $z = u^4(x)$ et $y = x - z$. Alors, $z \in \operatorname{Im} u$, $y + z = x$ et $u(y) = u(x) - u^5(x) = 0_E$.

En conclusion, $E = \operatorname{Ker} u \oplus \operatorname{Im} u$.

Exercice D

1. Construisons par récurrence sur $n \in \mathbb{N}$, la suite $(a_n)_n$.

▹ Nécessairement $a_0 = \Phi(1)$. Alors, pour tout $P \in \mathbb{C}_0[X]$, $\Phi(P) = \Phi(1).P = a_0 P$.

▹ Soit $n \in \mathbb{N}$ telle que les valeurs $a_0, \ldots, a_n$ soient définies, et donc que pour tout $P \in \mathbb{C}_n[X]$,

$$\Phi(P) = \sum_{k=0}^{n} a_k P^{(k)}.$$

Soit P de degré $n+1$. D'après l'hypothèse de récurrence appliquée à P',

$$\Phi(P)' = \Phi(P') = \sum_{k=0}^{n} a_k P^{(k+1)}.$$

En intégrant, il existe $C_P \in \mathbb{C}$ telle que $\Phi(P) = C_P \sum\limits_{k=0}^{n} a_k P^{(k)}$.

Pour deux polynômes P et Q unitaires de degré $n+1$, $C_P = C_Q$ en utilisant l'hypothèse de récurrence pour $P - Q \in \mathbb{C}_n[X]$. Notons a_{n+1} cette valeur commune de la constante. Alors, par linéarité de Φ, pour tout $P \in \mathbb{C}_{n+1}[X]$,

$$\Phi(P) = \sum_{k=0}^{n+1} a_k P^{(k)},$$

ce qui termine la récurrence.

2. Remarquons tout d'abord que Φ commute bien avec la dérivation donc il existe une suite $(a_n)_n$ telle que

$$\forall P \in \mathbb{C}[X], \qquad P(X+1) = \sum_{k=0}^{\infty} a_k P^{(k)}.$$

En considérant cette égalité pour $P = X^n$ et en évaluant en 0, on obtient que $1 = n!a_n$ donc $a_n = \frac{1}{n!}$.

Exercice E

Remarquons tout d'abord que $p+q$ est un projecteur si, et seulement si, $(p+q)^2 = p+q$ soit en développant (et en utilisant que $p^2 = p$ et $q^2 = q$) $p \circ q + q \circ p = 0$.
Montrons maintenant que cette condition entraîne bien celle de l'énoncé (la réciproque est évidente).

$$\begin{aligned} p \circ q &= p \circ p \circ q \\ &= p \circ (-q \circ p) \\ &= (-p \circ q) \circ p \\ &= q \circ p \circ p = q \circ p. \end{aligned}$$

Avec la condition $p \circ q = -q \circ p$, on obtient $p \circ q = q \circ p = 0$.

Exercice F

Considérons p la projection sur G parallèlement à F et introduisons l'application linéaire induite par p de H dans G (qui est bien définie car $G = \operatorname{Im} p$) :

$$u : \left\{ \begin{array}{ccc} H & \to & G \\ x & \mapsto & p(x) \end{array} \right.$$

▹ L'application u est injective car $\operatorname{Ker} u = \operatorname{Ker} p \cap H = F \cap H = \{0_E\}$.
▹ Montrons que u est surjective. Pour tout $y \in G = \operatorname{Im} p$, il existe $x \in E$ tel que $y = p(x)$. Décomposons x selon $F \oplus H$: il existe $f \in \mathbb{K} = \operatorname{Ker} p$ et $h \in H$ tels que $x = f + h$. Alors,

$$y = p(f+g) = p(f) + p(g) = p(g) = u(g).$$

En conclusion, $G = \operatorname{Im} u$ ce qui traduit la surjectivité.

Corrigés des problèmes du chapitre 14

Problème 1

Partie A – Premiers exemples

1. a. E_P est une partie non vide de E (car elle contient le polynôme nul) et stable par combinaison linéaire : c'est un sous-espace vectoriel de E.
b. La décomposition $E = E_P \oplus \mathbb{R}_{p-1}[X]$ n'est autre que la division euclidienne par P.
c. Si $p = q$, alors $\mathbb{R}_{p-1}[X]$ est un supplémentaire commun à E_P et E_Q.
2. a. Soit $y \notin F$; il existe $f \in F$ et $\lambda \in \mathbb{K}^*$ tel que $y = f + \lambda x$ donc $x = -\lambda^{-1} f + \lambda^{-1} y \in F + \operatorname{Vect}(y)$. Ainsi, $E = F \oplus \operatorname{Vect}(x) = F + \operatorname{Vect}(y)$.
Par ailleurs, $F \cap \operatorname{Vect}(y) = \{0_E\}$ car sinon $y \in F$.
b. On remarque que le vecteur $x = x_1 + x_2$ n'appartient ni à F_1 ni à F_2. D'après la question précédente, $\operatorname{Vect}(x)$ est un supplémentaire commun à F_1 et F_2.

Partie B – Exemple avec projecteur

1. Les deux sous-espaces sont en somme directe car $\mathscr{I} \cap \mathscr{P} = \{0\}$. Or, leur somme est égale à E car tout polynôme P admet la décomposition suivante comme somme d'un polynôme pair et d'un polynôme impair :

$$P(X) = \frac{P(X)+P(-X)}{2} + \frac{P(X)-P(-X)}{2}$$

Par conséquent, $E = \mathscr{I} \oplus \mathscr{P}$, c'est-à-dire $\mathscr{I}$ est un supplémentaire de $\mathscr{P}$.
2. a. Remarquons tout d'abord que, pour tout $P \in E$, $\varphi(P) \in \mathscr{P}$; d'où $\operatorname{Im}\varphi \subset \mathscr{P}$.
Montrons l'autre inclusion : remarquons (un peu inspiré par la question suivante) que pour tout polynôme pair P, $P = \varphi(P) \in \operatorname{Im}\varphi$. Donc $\operatorname{Im}\varphi = \mathscr{P}$.
Notons respectivement P_0 et P_1 les parties paires et impaires du polynôme P. Alors,

$$\begin{aligned} P \in \operatorname{Ker}\varphi &\Leftrightarrow (1+X)P(X)+(1-X)P(-X) = 0 \\ &\Leftrightarrow 2(P_0(X)+XP_1(X)) = 0 \\ &\Leftrightarrow P(X) = (1-X)P_1(X) \end{aligned}$$

D'où le résultat annoncé.
b. L'application φ est un endomorphisme de E et, pour tout polynôme P, $\varphi(\varphi(P)) = \varphi(P)$ donc φ est un projecteur de E.
c. Comme φ est un projecteur, son image et son noyau sont supplémentaires, c'est-à-dire $\operatorname{Ker}\varphi$ est un supplémentaire de $\operatorname{Im}\varphi = \mathscr{P}$.

Partie C – Condition suffisante

1. Montrons que G et F_1 sont supplémentaires par analyse-synthèse.
▹ Si $x \in G + F_1$, alors il existe $(y,z) \in F_1^2$ tel que $x = (y - u(y)) + z = y + z + u(-y)$. Or, F_1 et F_2 sont supplémentaires d'où $u(-y)$ est le terme de F_2 dans la décomposition de x sur $F_1 \oplus F_2$: par conséquent, $y = u^{-1}(p_{F_2}(x))$ et $z = p_{F_1}(x) - y$ avec des notations évidentes.
▹ Soit $x \in E$. Posons $y = u^{-1}(p_{F_2}(x))$ et $z = p_{F_1}(x) - y$. On vérifie successivement que $x = (y - u(y)) + z$, $y \in F_1$ et $z \in F_1$. D'où G et F_1 sont supplémentaires et on procède de même pour G et F_2.
2. a. Considérons les applications

$$u : \left\{ \begin{array}{ccc} F_1 & \to & E \\ x & \mapsto & x \end{array} \right. \qquad v : \left\{ \begin{array}{ccc} F_2 \oplus G & \to & F_2 \\ x & \mapsto & p_{F_2}(x) \end{array} \right.$$

la corestriction de la projection sur F_2 parallèlement à G. Montrons que $v \circ u$ est un isomorphisme de F_1 dans F_2.
Si $x \in \operatorname{Ker} v \circ u$, alors $u(x) \in \operatorname{Ker} v = G$ or $u(x) = x \in F_1$ d'où $x = 0$ car $F_1 \cap G = \{0\}$.
Pour tout $z \in F_2 \subset F_1 \oplus G$, il existe $(x,t) \in F_1 \times G$ tel que $z = x + t$ d'où z est le projeté sur F_2 parallèlement à G de $x + t$ donc de x. Par conséquent, $z = v \circ u(x)$.
b. • Il suffit de prendre $F_1 = F_2$.
• Dans $E = \mathbb{R}^4$, les sous-espaces

$$\begin{aligned} F_1 &= \{(x,y,z,t) \in E,\ z = t = 0\} \\ F_2 &= \{(x,y,z,t) \in E,\ y = t = 0\} \end{aligned}$$

admettent $\operatorname{Vect}((0,1,1,0),(0,0,0,1))$ comme supplémentaire commun. Toutefois,

$$F_1 + F_2 \subset \{(x,y,z,t) \in E,\ t = 0\} \subsetneq E.$$

Chapitre 15

Dimension des espaces vectoriels

1. Bases — **2.** Dimension — **3.** Applications linéaires en dimension finie

Objectifs et compétences du programme

- Utiliser la définition de base et les coordonnées dans une base.
- Calculer la dimension d'un espace vectoriel.
- Exploiter des propriétés de dimension finie afin de simplifier des raisonnements.

Hermann Günther Grassmann, mathématicien allemand (1809-1877).
Il est considéré comme l'inventeur (ou tout au moins comme l'un des inventeurs) des espaces vectoriels. Toutefois, il ne bénéficie pas immédiatement de la reconnaissance pour ces travaux. Il faut dire qu'il les mentionne une première fois dans une thèse intitulée « Théorie des flots et des marées » en 1839. Il publie ensuite un mémoire intitulé *l'enseignement de l'extension linéaire, une nouvelle branche des mathématiques* en 1844 dans lequel il détaille les prémices de la définition axiomatique de la structure d'espace vectoriel. Ce livre ne reçoit pas l'accueil espéré et ses résultats ne sont redécouverts que vingt ou trente ans plus tard.

COURS

Le chapitre débute avec l'introduction des bases : ces familles libres et génératrices permettent de généraliser les manipulations de coordonnées connues dans des espaces vectoriels naturels comme $\mathbb{R}^2$ ou $\mathbb{R}^3$. On remarque ensuite que dans un espace vectoriel admettant une base finie, toutes les bases ont le même nombre d'éléments : cela permet de définir rigoureusement la notion de dimension d'un espace vectoriel. On apprend ensuite à calculer la dimension puis à exploiter celle-ci pour simplifier les différentes caractérisations (propriété de base, de supplémentaires, d'isomorphisme...).

1. Bases

1.1. Familles libres

Définition 15.1.

Soit $(x_i)_{i\in I}$ une famille de vecteurs d'un $\mathbb{K}$-espace vectoriel E.
▹ La famille $(x_i)_{i\in I}$ est liée s'il existe une combinaison linéaire à coefficients non tous nuls donnant le vecteur nul.
▹ Dans le cas contraire, la famille $(x_i)_{i\in I}$ est libre (ou linéairement indépendante). En d'autres termes, l'unique combinaison linéaire de $(x_i)_{i\in I}$ nulle est celle dont tous les coefficients sont nuls.

Remarque

▹ Une famille qui contient le vecteur nul est forcément liée (il suffit de prendre un coefficient non nul devant le vecteur nul et des coefficients nuls partout ailleurs). Par contraposition, toute famille libre ne contient pas le vecteur nul.
▹ Toute famille contenant une famille liée est liée (il suffit de prendre les coefficients non tous nuls devant les vecteurs de la sous-famille liée de sorte à obtenir une somme nulle et des coefficients nuls partout ailleurs). Par contraposition, toute famille contenue dans une famille libre est libre.

Exemple

Deux vecteurs x et $y \in E$ forment une famille liée si, et seulement s'il existe $\lambda, \mu \in \mathbb{K}$ non tous les deux nuls tels que $\lambda x + \mu y = 0_E$.
- Si $\lambda \neq 0$, alors $x = -\frac{\mu}{\lambda} y$.
- Sinon, $\mu \neq 0$, et $y = -\frac{\lambda}{\mu} x$.

Autrement dit, si deux vecteurs forment une famille liée, alors ils sont colinéaires. La réciproque est élémentaire.

Exemple

Une famille de polynômes $(P_k)_{k\in[1,n]}$ est échelonnée en degré si

$$\forall k \in [\![1, n]\!], \qquad 0 \leq \deg P_k < \deg P_{k+1}.$$

Remarquons en particulier qu'une famille échelonnée en degré ne contient pas le polynôme nul.
Alors, toute famille de polynômes échelonnée en degré est libre.
Montrons le résultat par récurrence.
▹ Un polynôme non nul forme une famille libre.

▹ Soit $n \in \mathbb{N}\backslash\{0\}$. Supposons que le résultat est vrai pour toute famille de n polynômes échelonnée en degré et considérons une famille de $n+1$ polynômes $(P_k)_{k\in[1,n+1]}$ échelonnée en degré.
Si $\sum_{k=1}^{n+1} \lambda_k P_k = 0$, alors $\lambda_{n+1} = 0$ (car sinon le membre de gauche est de degré $\deg P_{n+1} \neq -\infty$) et on conclut avec l'hypothèse de récurrence appliquée à la famille $(P_k)_{k\in[1,n]}$.

Exemple

Il existe des familles libres de polynômes qui ne sont pas échelonnées en degré. Par exemple, les polynômes de Lagrange $(L_k)_{0\leq k\leq n}$ associés aux réels $a_0 < \ldots < a_n$ forment une famille libre.
En effet, si $\sum_{k=0}^{n} \lambda_k L_k(X) = 0$, alors, en évaluant en chacun des a_i, on trouve que $\lambda_i = 0$ pour tout $i \leq n$.

Donnons un résultat utile pour la section suivante.

Lemme 15.2.

Soit $(x_i)_{i\in I}$ une famille libre d'un $\mathbb{K}$-espace vectoriel E et $x \in E$ tel que la famille composée de x et $(x_i)_{i\in I}$ est liée. Alors, $x \in \mathrm{Vect}(x_i)_{i\in I}$.

Démonstration
Comme la famille composée de x et $(x_i)_{i\in I}$ est liée, il existe des coefficients non tous nuls μ et $(\lambda_i)_{i\in I}$ tels que

$$\mu x + \sum_{i\in I} \lambda_i . x_i = 0_E.$$

Si $\mu = 0$, alors $\sum_{i\in I} \lambda_i . x_i = 0_E$ et par liberté de la famille $(x_i)_{i\in I}$, tous les coefficients λ_i sont nuls : contradiction. Donc $\mu \neq 0$ et

$$x = \sum_{i\in I} -\frac{\lambda_i}{\mu} . x_i$$

■

1.2. Bases

Passons maintenant à la notion de base d'un espace vectoriel.

Définition 15.3.

Une base d'un $\mathbb{K}$-espace vectoriel E est une famille de E à la fois libre et génératrice.

Exemple

Voici quelques exemples de bases dites canoniques.
▹ $(1,0,\ldots,0),\ldots,(0,\ldots,0,1)$ est une base de $\mathbb{K}^n$.
▹ $(X^k)_{k\in[0,n]}$ est une base de $\mathbb{K}_n[X]$.
▹ $(X^k)_{k\in\mathbb{N}}$ est une base de $\mathbb{K}[X]$.

Proposition 15.4.

Une famille $(x_i)_{i\in I}$ est une base si, et seulement si, tout vecteur de E admet une unique écriture comme combinaison linéaire des vecteurs $(x_i)_{i\in I}$. Les scalaires de cette combinaison linéaire sont les

coefficients du vecteur dans la base. L'application qui, à un vecteur, associe la coordonnée selon x_i est une forme linéaire.

Démonstration

▹ L'existence de l'écriture sur une base correspond à la propriété de famille génératrice, l'unicité de la propriété de liberté ; en effet, s'il existe deux combinaisons linéaires égales

$$\sum_{i\in I}\lambda_i x_i = \sum_{i\in I}\mu_i x_i,$$

alors

$$\sum_{i\in I}(\lambda_i-\mu_i)x_i = 0_E.$$

D'après la liberté de la famille, pour tout $i\in I$, $\lambda_i=\mu_i$: d'où l'unicité de l'écriture.

▹ Soit $\lambda\in\mathbb{K}$, a, $b\in E$ d'écriture

$$a=\sum_{i\in I}a_i x_i, \qquad b=\sum_{i\in I}b_i x_i.$$

Alors,

$$a+\lambda b=\sum_{i\in I}a_i x_i+\lambda\sum_{i\in I}b_i x_i=\sum_{i\in I}(a_i+\lambda b_i)x_i.$$

■

Exemple

D'après la formule de Taylor pour les polynômes, la famille $((X-a)^k)_{k\in\mathbb{N}}$ est une base de $\mathbb{K}[X]$ et la forme linéaire coordonnée associée à $(X-a)^k$ est $P\mapsto \frac{1}{k!}P^{(k)}(a)$.

Exemple

D'après le théorème d'interpolation, les polynômes de Lagrange $(L_k)_{0\leq k\leq n}$ associés aux réels deux à deux distincts $a_0<\ldots<a_n$ forment une base de $\mathbb{R}_n[X]$ et la forme linéaire coordonnée associée à $L_k(X)$ est $P\mapsto P(a_k)$.

La proximité de la propriété de base et celle de somme directe est explicitée avec la proposition suivante.

Proposition 15.5.

Soit F et G deux sous-espaces supplémentaires d'un $\mathbb{K}$-espace vectoriel E. La famille obtenue par concaténation d'une base de F et d'une base de G est une base de E.

Démonstration

Soit I et J des ensembles disjoints, $(e_i)_{i\in I}$ une base de F et $(e_i)_{i\in J}$ une base de G.

▹ Soit $x\in E$. Comme $E=F+G$, il existe $y\in F$ et $z\in G$ tels que $x=y+z$. Comme $F=\mathrm{Vect}(e_i)_{i\in I}$ et $G=\mathrm{Vect}(e_i)_{i\in J}$, on en déduit que $x\in\mathrm{Vect}(e_i)_{i\in I\cup J}$ donc la famille $(e_i)_{i\in I\cup J}$ est génératrice.

▹ Soit $(\lambda_i)_{i\in I\cup J}$ une famille de scalaires presque nulle telle que

$$\sum_{i\in I\cup J}\lambda_i e_i = 0_E.$$

Alors,

$$\sum_{i\in I}\lambda_i e_i = -\sum_{i\in J}\lambda_i e_i\in F\cap G.$$

Comme F et G sont en somme directe, ce vecteur est nul.

$$\sum_{i\in I}\lambda_i e_i = \sum_{i\in J}\lambda_i e_i = 0_E.$$

Enfin, comme les familles $(e_i)_{i\in I}$ et $(e_i)_{i\in J}$ sont libres, tous les scalaires sont nuls. En conclusion, la famille $(e_i)_{i\in I\cup J}$ est libre. ■

La proposition suivante fournit la réciproque.

Proposition 15.6.

Soit $(e_i)_{i\in I}$ une base de E et $J\subset I$. Alors, les sous-espaces $F=\mathrm{Vect}(e_i)_{i\in J}$ et $G=\mathrm{Vect}(e_i)_{i\in I\setminus J}$ sont supplémentaires.

Démonstration

▹ Pour tout $x\in E$, il existe des scalaires $(x_i)_{i\in I}$ tels que

$$x=\sum_{i\in I}x_i e_i=\sum_{i\in J}x_i e_i+\sum_{i\in I\setminus J}x_i e_i$$

donc $x\in F+G$.

▹ Soit $(\lambda_i)_{i\in I}$ une famille de scalaires presque nulle telle que

$$\sum_{i\in J}\lambda_i e_i=\sum_{i\in I\setminus J}\lambda_i e_i.$$

Alors, par la liberté de la famille $(e_i)_{i\in I}$, tous les scalaires λ_i sont nuls : les sous-espaces F et G sont en somme directe. ■

1.3. Bases et applications linéaires

L'idée de cette section consiste à remarquer qu'une application linéaire est uniquement déterminée par l'image d'une base.

Proposition 15.7.

Soit E et F deux $\mathbb{K}$-espaces vectoriels, $(e_i)_{i\in I}$ une base de E et $(f_i)_{i\in I}$ une famille de F. Il existe une unique application linéaire $u\in\mathscr{L}(E,F)$ telle que

$$\forall i\in I,\qquad u(e_i)=f_i.$$

Démonstration

▹ Supposons qu'une telle application linéaire u existe. Tout vecteur x admet une unique écriture de la forme $\sum_{i\in I}\lambda_i e_i$; donc son image est $u(x)=\sum_{i\in I}\lambda_i f_i$. Ainsi, l'application linéaire est unique.

▹ Réciproquement, l'application définie par $u\left(\sum_{i\in I}\lambda_i e_i\right)=\sum_{i\in I}\lambda_i f_i$ est bien linéaire d'après la linéarité des formes linéaires coordonnées dans la base $(e_i)_{i\in I}$. ■

Exemple

Déterminons l'unique application linéaire $u:\mathbb{R}^3\to\mathbb{R}^2$ telle que $u(1,0,0)=(1,4)$, $u(0,1,0)=(0,7)$ et $u(0,0,1)=(8,9)$. Pour tout $(x,y,z)\in\mathbb{R}^3$,

$$\begin{aligned}u(x,y,z)&=u(x(1,0,0)+y(0,1,0)+z(0,0,1))\\&=x\,u(1,0,0)+y\,u(0,1,0)+z\,u(0,0,1)\\&=x(1,4)+y(0,7)+z(8,9)=(x+8z,4x+7y+9z).\end{aligned}$$

On peut rapprocher cette dernière proposition du résultat suivant (dont elle est un cas particulier avec $F_k = \mathrm{Vect}(e_k)$ pour tout $k \in [\![1, n]\!]$).

Proposition 15.8.

Soit E et F deux $\mathbb{K}$-espaces vectoriels, $F_1, \dots F_n$ des sous-espaces de E tels que $E = F_1 \oplus \dots \oplus F_n$ et $u_1 \in \mathscr{L}(F_1, F)$, ..., $u_n \in \mathscr{L}(F_n, F)$. Alors, il existe une unique application linéaire $u \in \mathscr{L}(E, F)$ telle que, pour tout $k \in [\![1, n]\!]$, $u_{|F_k} = u_k$.

Démonstration

▹ En notant, pour tout k, p_k la projection sur F_k parallèlement à la somme des autres sous-espaces, l'application $u = \sum_{k=1}^{n} u_k \circ p_k$ convient.

▹ Soit $x \in E$ et $(x_1, \dots, x_n) \in F_1 \times \dots \times F_n$ tels que $x = \sum_{k=1}^{n} x_k$. Alors, une application linéaire $u \in \mathscr{L}(E, F)$ telle que, pour tout $k \in [\![1, n]\!]$, $u_{|F_k} = u_k$, vérifie $u(x_k) = u_k(x_k)$ pour tout k donc

$$u(x) = \sum_{k=1}^{n} u_k(x_k).$$

Le vecteur $u(x)$ est uniquement défini pour chaque $x \in E$, donc u est unique. ■

Théorème 15.9.

Soit E et F deux $\mathbb{K}$-espaces vectoriels, $(e_i)_{i \in I}$ une base de E et $u \in \mathscr{L}(E, F)$.
▹ L'application u est injective si, et seulement si, la famille $(u(e_i))_{i \in I}$ est libre.
▹ L'application u est surjective si, et seulement si, la famille $(u(e_i))_{i \in I}$ est génératrice de F.
▹ L'application u est bijective si, et seulement si, la famille $(u(e_i))_{i \in I}$ est une base de F.

Démonstration

▹ L'application u n'est pas injective si, et seulement s'il existe un vecteur non nul dans le noyau. En écrivant ce vecteur dans la base $(e_i)_{i \in I}$, on obtient qu'il existe un vecteur non nul dans le noyau si, et seulement s'il existe une combinaison linéaire nulle non triviale des $(u(e_i))_{i \in I}$. En conclusion, u n'est pas injective si, et seulement si, la famille $(u(e_i))_{i \in I}$ est liée.

▹ L'application u est surjective si, et seulement si, $F = \mathrm{Im}\, u = \mathrm{Vect}(u(e_i))_{i \in I}$. ■

Exemple

L'endomorphisme u de $\mathbb{R}[X]$ défini par l'image de la base canonique $u(X^k) = X^k + 1$, pour tout entier $k \in \mathbb{N}$, est injectif car l'image d'une base est une famille libre (car échelonnée en degré).

2. Dimension

2.1. Espaces vectoriels de dimension finie

Définition 15.10.

Un espace vectoriel est de dimension finie s'il admet une famille génératrice finie.

Exemple

▹ Les espaces $\mathbb{K}^n$, $\mathbb{K}_n[X]$ sont des $\mathbb{K}$-espaces vectoriels de dimension finie.

▹ L'espace $\mathbb{K}[X]$ n'est pas de dimension finie ; en effet, s'il existait une famille génératrice finie $(P_i)_{i \leq N}$ de $\mathbb{K}[X]$, alors tout polynôme serait de degré au plus $\max\{\deg P_i,\ i \leq N\}$, ce qui est absurde.

Intéressons-nous à l'obtention d'une base pour un $\mathbb{K}$-espace vectoriel de dimension finie. On va procéder de deux manières différentes : en partant d'une famille génératrice et en extrayant une base ou en partant d'une famille libre et en la complétant en une base.

Proposition 15.11. Théorème de la base extraite

Soit E un $\mathbb{K}$-espace vectoriel de dimension finie différent de $\{0_E\}$. Toute famille génératrice finie de E admet une sous-famille qui est une base de E.

Démonstration

Soit $(e_i)_{i\in I}$ une famille génératrice finie de E. Comme E est différent de $\{0_E\}$, il existe des vecteurs non nuls donc des familles libres à un élément dans $(e_i)_{i\in I}$. Considérons une sous-famille libre $(e_i)_{i\in J}$ avec $J \subset I$ de cardinal maximal.

Par maximalité, lorsque l'on ajoute un vecteur e_j avec $j \notin J$, la famille $(e_i)_{i\in J\cup\{j\}}$ est liée donc $e_j \in \mathrm{Vect}(e_i)_{i\in J}$. Ainsi, $E = \mathrm{Vect}(e_i)_{i\in I} = \mathrm{Vect}(e_i)_{i\in J}$: la famille $(e_i)_{i\in J}$ est libre et génératrice donc c'est une base de E.

■

Corollaire 15.12.

Tout $\mathbb{K}$-espace vectoriel de dimension finie admet une base.

Proposition 15.13. Théorème de la base incomplète

Soit E un $\mathbb{K}$-espace vectoriel de dimension finie et $(e_i)_{i\in I}$ une famille génératrice finie.
Toute famille libre de E peut être complétée en une base de E avec des vecteurs de la famille $(e_i)_{i\in I}$.

Démonstration

Soit $(e_i)_{i\in I}$ une famille génératrice finie de E, $(e'_j)_{j\in J}$ une famille libre de E et considérons une sous-famille libre de la réunion de ces deux familles de vecteurs, contenant $(e'_j)_{j\in J}$ et de cardinal maximal. Cette famille existe car $(e'_j)_{j\in J}$ est une famille libre, elle est génératrice car lorsque l'on ajoute un nouveau vecteur de la famille $(e_i)_{i\in I}$, on obtient une famille liée : on conclut comme précédemment en utilisant le lemme rappelé ci-dessus. ■

2.2. Définition de la dimension

Proposition 15.14.

Toute famille de $n+1$ vecteurs d'un espace vectoriel admettant une base de cardinal n est liée.

Démonstration

Raisonnons par récurrence sur $n \in \mathbb{N} \setminus \{0\}$.

▹ Le résultat est vrai pour $n = 1$: si une base est de cardinal 1 alors tous les vecteurs sont colinéaires avec le vecteur de cette base.

▹ Soit $n \geq 1$. Supposons le résultat établi pour les espaces vectoriels admettant une base de cardinal au plus n. Soit E un espace vectoriel de base $(e_i)_{i\in[1,n+1]}$ et $(x_k)_{k\in[1,n+2]}$ une famille de $n+2$ vecteurs de E.

- Si, pour tout $k \leq n+2$, $x_k \in \mathrm{Vect}(e_i)_{i\in[1,n]}$, alors on conclut que la famille est liée en appliquant l'hypothèse de récurrence.
- Sinon, il existe k_0 tel que la coordonnée sur e_{n+1} de x_{k_0} est non nulle. Notons, pour tout i et k, $\lambda_{k,i}$

la coordonnée selon e_i de x_k. Ainsi, pour tout k,

$$x_k = \sum_{i=1}^{n+1} \lambda_{k,i}\, e_i.$$

La famille de vecteurs $\left(x_k - \frac{\lambda_{k,n+1}}{\lambda_{k_0,n+1}} x_{k_0}\right)_{k \neq k_0}$ est une famille de $n+1$ vecteurs dans l'espace $\mathrm{Vect}(e_i)_{i\in[1,n]}$ car, pour tout k,

$$\begin{aligned} x_k - \frac{\lambda_{k,n+1}}{\lambda_{k_0,n+1}} x_{k_0} &= \sum_{i=1}^{n+1} \lambda_{k,i}\, e_i - \frac{\lambda_{k,n+1}}{\lambda_{k_0,n+1}} \sum_{i=1}^{n+1} \lambda_{k_0,i}\, e_i \\ &= \sum_{i=1}^{n} \left(\lambda_{k,i} - \frac{\lambda_{k,n+1}}{\lambda_{k_0,n+1}} \lambda_{k_0,i}\right) e_i. \end{aligned}$$

D'après l'hypothèse de récurrence, cette famille est liée donc $(x_k)_{k\in[1,n+2]}$ est liée.

■

Proposition 15.15.

Soit E un $\mathbb{K}$-espace vectoriel de dimension finie. Toutes les bases de E sont finies et ont le même cardinal.

Démonstration

D'après la proposition précédente, toute famille dont le cardinal est supérieur au cardinal d'une base est liée. Comme deux bases sont deux familles libres, le cardinal de l'une est donc inférieur ou égal à celui de l'autre et réciproquement. ■

Définition 15.16.

La dimension d'un $\mathbb{K}$-espace vectoriel non nul de dimension finie est le cardinal commun de toutes ses bases, noté $\dim_{\mathbb{K}} E$ ou, plus simplement, $\dim E$ s'il n'y a pas ambiguïté sur le corps de base. Par convention, la dimension du sous-espace nul est 0.

Exemple

En dénombrant les éléments de la base canonique, on obtient que $\mathbb{K}^n$ est un $\mathbb{K}$-espace vectoriel de dimension n.

Remarque

Attention à préciser (le cas échéant) le corps de base. Par exemple, $\dim_{\mathbb{R}} \mathbb{C} = 2$ (une base est 1 et i) et $\dim_{\mathbb{C}} \mathbb{C} = 1$ (une base est 1).

Plus généralement, si E est un $\mathbb{C}$-espace vectoriel de dimension n, alors E est également un $\mathbb{R}$-espace vectoriel de dimension $2n$ (en remplaçant la base $(e_1, \dots, e_n)$ par la base $(e_1, ie_1, \dots, e_n, ie_n)$).

Exemple

L'espace $\mathbb{K}_n[X]$ est un $\mathbb{K}$-espace vectoriel de dimension $n+1$ car la base canonique $(X^k)_{k\in[0,n]}$ compte $n+1$ éléments.

Exemple

Soit $(a,b,c)\in\mathbb{R}^*\times\mathbb{R}^2$.
▹ L'espace des solutions de l'équation différentielle $a\,y''+b\,y'+c\,y=0$ est de dimension 2.
▹ L'espace des suites vérifiant la relation de récurrence $a\,u_{n+2}+b\,u_{n+1}+c\,u_n=0$ pour tout n, est de dimension 2.

Avec le théorème de la base incomplète et le principe d'extraction de base, on obtient le résultat suivant.

Proposition 15.17.

Soit E un $\mathbb{K}$-espace vectoriel de dimension n et $(e_i)_{i\in[1,p]}$ une famille de vecteurs de E.
- Si $(e_i)_{i\in[1,p]}$ est libre, alors $p\leq n$.
- Si $(e_i)_{i\in[1,p]}$ est génératrice, alors $p\geq n$.

Ceci peut également se reformuler ainsi.

Proposition 15.18.

Soit E un $\mathbb{K}$-espace vectoriel de dimension n et $(e_i)_{i\in[1,p]}$ une famille de vecteurs de E. Il y a équivalence entre
- $(e_i)_{i\in[1,p]}$ est libre et $n=p$,
- $(e_i)_{i\in[1,p]}$ est génératrice de E et $n=p$,
- $(e_i)_{i\in[1,p]}$ est une base de E.

Démonstration
Il n'y a aucun vecteur à retirer (extraction de base depuis une famille génératrice) ou à ajouter (complétion d'une famille libre en une base) pour avoir une base car la famille est déjà de cardinal n. ■

Remarque
Cette proposition est en fait une caractérisation pratique des bases (il suffit de vérifier une propriété et de dénombrer les vecteurs de la base) comme famille libre maximale ou famille génératrice minimale.

Exemple

Soit $a_1,\dots,a_{n+1}$ des réels deux à deux distincts et $L_1,\dots,L_{n+1}$ les polynômes de Lagrange associés.
La famille $(L_k)_{k\in[1,n+1]}$ est libre ; en effet, si

$$\sum_{k=1}^{n+1}\lambda_k L_k(X)=0,$$

alors, en évaluant en chaque a_j, on obtient $\lambda_j=0$ puisque $L_k(a_j)=0$ si $j\neq k$, 1 sinon.
Par conséquent, cette famille est libre, son cardinal est égal à $n+1=\dim\mathbb{R}_n[X]$ donc c'est une base de cet espace vectoriel.
On retrouve ainsi la propriété d'interpolation : la décomposition de $P\in\mathbb{R}_n[X]$ dans cette base est

$$P(X)=\sum_{k=1}^{n+1}P(a_k)L_k(X).$$

Proposition 15.19.

Deux $\mathbb{K}$-espaces vectoriels de dimension finie sont isomorphes si, et seulement s'ils ont la même dimension.

Démonstration

▹ Soit E et F deux espaces de même dimension n et de bases $(e_1,\dots,e_n)$ et $(f_1,\dots,f_n)$ respectivement. L'unique application linéaire $u : E \to F$ définie par $u(e_i) = f_i$ pour tout i est un isomorphisme car elle transforme une base en une base.

▹ Réciproquement, soit $u : E \to F$ un isomorphisme entre deux espaces de dimension finie. Comme l'image d'une base de E par u est une base de F, ces bases comportent le même nombre d'éléments, soit $\dim E = \dim F$. ■

Remarque

On définit la dimension d'un sous-espace affine comme la dimension de sa direction. Les sous-espaces affines de dimension 1 sont les droites affines, de dimension 2 les plans affines.

2.3. Calculs de dimension

Proposition 15.20.

Soit $E_1,\dots,E_p$ des espaces vectoriels de dimension finie. L'espace vectoriel $E_1 \times \dots \times E_p$ est de dimension finie et

$$\dim(E_1 \times \dots \times E_p) = \sum_{i=1}^{p} \dim E_i.$$

Démonstration

Il suffit de montrer le résultat pour deux espaces. Si $(e_1,\dots,e_p)$ est une base de E_1 et $(f_1,\dots,f_n)$ est une base de E_2, alors la famille $((e_1,0_{E_2}),\dots,(e_p,0_{E_2}),(0_{E_1},f_1),\dots,(0_{E_1},f_n))$ est une base de $E_1 \times E_2$. ■

On en déduit directement la dimension des espaces d'applications linéaires.

Proposition 15.21.

Soit E et F deux espaces vectoriels de dimension finie. Alors, $\mathscr{L}(E,F)$ est de dimension finie et

$$\dim \mathscr{L}(E,F) = \dim E \times \dim F.$$

Démonstration

Notons $(e_1,\dots,e_n)$ une base de E. L'application linéaire

$$\left\{ \begin{array}{ccc} \mathscr{L}(E,F) & \to & F^{\dim E} \\ u & \mapsto & (u(e_1),\dots,u(e_n)) \end{array} \right.$$

est un isomorphisme (car une application linéaire est uniquement déterminée par l'image de la base $(e_1,\dots,e_n)$) donc $\dim \mathscr{L}(E,F) = \dim(F^{\dim E}) = \dim F \times \dim E$. ■

En prenant F égal au corps de base, on obtient le corollaire suivant.

Corollaire 15.22.

Soit E un espace vectoriel de dimension finie. Alors, l'espace dual de E est de dimension finie et $\dim E^* = \dim E$.

Proposition 15.23.

Soit E un espace vectoriel de dimension finie et F un sous-espace vectoriel de E. Alors,
▹ F est de dimension finie et $\dim F \leq \dim E$.
▹ Si $\dim F = \dim E$, alors $F = E$.

Démonstration
Supposons $F \neq \{0_E\}$. Considérons une famille de F libre, de cardinal maximal (qui existe car les familles libres de E sont de cardinal au plus $\dim E$). Si on ajoute un autre vecteur de F, la famille obtenue est liée. D'après le lemme déjà utilisé, cette famille est génératrice de F. Ainsi, F admet une base finie et $\dim F \leq \dim E$.
Si $\dim E = \dim F$, alors une base de F est une famille libre de cardinal $\dim E$ donc une base de E : ainsi $F = E$. ■

Le théorème de la base incomplète donne également l'existence de supplémentaires.

Proposition 15.24.

Tout sous-espace vectoriel d'un espace vectoriel de dimension finie admet un supplémentaire.

Démonstration
Soit F un sous-espace vectoriel de E de base $(e_i)_{i\in[1,p]}$. Complétons la famille libre $(e_i)_{i\in[1,p]}$ en une base $(e_i)_{i\in[1,n]}$ de E. Alors, le sous-espace $\mathrm{Vect}(e_i)_{i\in[p+1,n]}$ est un supplémentaire de F. ■

Proposition 15.25.

Soit E un espace vectoriel de dimension finie et F_1, F_2 deux sous-espaces en somme directe. Alors, $F_1 \oplus F_2$ est de dimension finie et $\dim F_1 \oplus F_2 = \dim F_1 + \dim F_2$.

Démonstration
C'est la propriété établie au chapitre précédent : la famille obtenue par concaténation d'une base de F_1 et d'une base de F_2 est une base de E. ■

En appliquant ce résultat à un espace F, on obtient qu'un supplémentaire de F (dont on a établi l'existence avant) a pour dimension $\dim E - \dim F$.

Corollaire 15.26.

Soit E un espace vectoriel et F un sous-espace vectoriel de E. Tous les supplémentaires de F sont isomorphes.

Passons au cas général de la somme de deux sous-espaces.

Proposition 15.27. Formule de Grassmann

Soit E un espace vectoriel de dimension finie et F_1, F_2 deux sous-espaces. Alors, F_1+F_2 est de dimension finie et

$$\dim F_1+F_2 = \dim F_1 + \dim F_2 - \dim F_1 \cap F_2.$$

Démonstration

Si la somme est directe (c'est-à-dire si $F_1 \cap F_2 = \{0_E\}$), alors le résultat est déjà établi. Sinon, considérons G_1 un supplémentaire de $F_1 \cap F_2$ dans F_1. Alors, $F_1+F_2 = G_1 \oplus F_2$ (la propriété de somme directe est évidente car $F_2 \cap G_1 = F_2 \cap F_1 \cap G_1 = \{0_E\}$, pour le reste il suffit de décomposer la composante sur F_1 d'un élément de F_1+F_2). Par conséquent,

$$\dim F_1+F_2 = \dim G_1 + \dim F_2.$$

Or, G_1 et $F_1 \cap F_2$ sont supplémentaires dans F_1 d'où :

$$\dim F_1 = \dim G_1 + \dim F_1 \cap F_2,$$

ce qui permet de conclure. ■

Proposition 15.28.

Soit E un espace vectoriel de dimension finie et F_1, F_2 deux sous-espaces. Il y a alors équivalence entre

- $E = F_1 \oplus F_2$;
- $E = F_1 + F_2$ et $\dim E = \dim F_1 + \dim F_2$;
- $F_1 \cap F_2 = \{0_E\}$ et $\dim E = \dim F_1 + \dim F_2$.

Démonstration

Il suffit d'appliquer la formule de Grassmann pour obtenir que $E = F_1 + F_2$ et $\dim E = \dim F_1 + \dim F_2$ entraîne $F_1 \cap F_2 = \{0_E\}$ et que $F_1 \cap F_2 = \{0_E\}$ et $\dim E = \dim F_1 + \dim F_2$ entraîne $\dim F_1 + F_2 = \dim E$ et, par conséquent, $F_1 + F_2 = E$.

Les autres implications sont déjà établies. ■

De manière générale, ces résultats sur la dimension des sommes de deux sous-espaces vectoriels se généralisent à une somme de p sous-espaces.

2.4. Rang d'une famille de vecteurs

Définition 15.29.

Le rang d'une famille finie de vecteurs est la dimension de l'espace engendré par cette famille.

Remarque

Le rang d'une famille finie est donc le cardinal maximal d'une famille libre extraite de cette famille.

Remarque

On peut bien entendu définir de la même manière le rang d'une famille infinie de vecteurs si le sous-espace engendré est de dimension finie.

Pour déterminer le rang d'une famille de vecteurs, on utilise l'algorithme du rang (ou pivot de Gauss) :

1. On commence par exprimer tous les vecteurs par leurs coordonnées dans une base et l'on s'intéresse à ces vecteurs colonnes.
2. On utilise un vecteur « pivot » pour annuler une coordonnée de tous les autres vecteurs en retranchant un multiple du vecteur pivot. Cette opération élémentaire ne modifie pas l'espace engendré.
3. On recommence jusqu'à obtenir une famille de vecteurs étagée : la famille obtenue est alors libre et génératrice de l'espace engendré par la famille initiale ; c'est donc une base de ce sous-espace et son cardinal est le rang recherché.

Exemple

Considérons les vecteurs de coordonnées $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 4 \end{bmatrix}, \begin{bmatrix} 1 \\ 6 \\ 5 \end{bmatrix}, \begin{bmatrix} 4 \\ 12 \\ 15 \end{bmatrix}$ et appliquons l'algorithme ci-dessus :

$$\begin{aligned}
&\operatorname{Vect}\left(\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 4 \end{bmatrix}, \begin{bmatrix} 1 \\ 6 \\ 5 \end{bmatrix}, \begin{bmatrix} 4 \\ 12 \\ 15 \end{bmatrix}\right) \\
&= \operatorname{Vect}\left(\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ -4 \\ -2 \end{bmatrix}, \begin{bmatrix} 0 \\ 4 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 4 \\ 3 \end{bmatrix}\right) \\
&= \operatorname{Vect}\left(\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ -4 \\ -2 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right).
\end{aligned}$$

Ainsi, la famille considérée est de rang 3.

3. Applications linéaires en dimension finie

3.1. Rang d'une application linéaire

Définition 15.30.

Soit u une application linéaire entre deux sous-espaces E et F avec E de dimension finie. Le rang de u est la dimension de l'image de u. On le note $\operatorname{rg}(u)$.

Remarque

Le rang d'une application linéaire u est aussi le rang de l'image par u d'une base de E puisqu'une telle famille est génératrice de $\operatorname{Im} u$.

Remarque

On remarque qu'il n'est pas nécessaire de supposer F de dimension finie dans cette définition car $\dim \operatorname{Im} u \leq \dim E$ (car la restriction de u entre E et $\operatorname{Im} u$ est surjective).

Exemple

Une application linéaire nulle est de rang 0 ; toute application linéaire de rang 0 est nulle puisque son image est le sous-espace nul.

Exemple

Si $u \in \mathscr{L}(E,F)$ avec E de dimension finie est injective, alors $\operatorname{rg} u = \dim E$; en effet, la restriction de u entre E et $\operatorname{Im} u$ est un isomorphisme (surjective par construction et injective comme restriction d'une application injective) donc $\dim \operatorname{Im} u = \dim E$.

Rappelons un résultat démontré au chapitre précédent et écrivons la conséquence en termes de dimension.

Proposition 15.31. Formule du rang

Soit u une application linéaire entre deux sous-espaces E et F avec E de dimension finie. L'application u induit un isomorphisme entre un supplémentaire de $\operatorname{Ker} u$ et $\operatorname{Im} u$. Par conséquent,

$$\dim E = \operatorname{rg} u + \dim \operatorname{Ker} u.$$

Remarque

La formule du rang ne signifie pas que $\operatorname{Im} u$ et $\operatorname{Ker} u$ sont supplémentaires (d'ailleurs en général, ces sous-espaces ne sont pas dans le même espace).

Démonstration

La première assertion est déjà établie : $\operatorname{Im} u$ est donc isomorphe à un supplémentaire de $\operatorname{Ker} u$ donc de même dimension. Ainsi,

$$\operatorname{rg} u = \dim \operatorname{Im} u = \dim E - \dim \operatorname{Ker} u.$$

■

Corollaire 15.32.

Soit E et F deux espaces vectoriels de même dimension finie et $u \in \mathscr{L}(E,F)$. Il y a équivalence entre

- u est injective ;
- u est surjective ;
- u est bijective.

Démonstration

Il suffit d'appliquer la formule du rang pour trouver selon les cas que $\operatorname{Ker} u = \{0_E\}$ (car $\dim \operatorname{Ker} u = 0$) ou $\operatorname{Im} u = F$ (car $\operatorname{Im} u$ est un sous-espace de F et $\dim \operatorname{Im} u = \dim E = \dim F$). ■

Exemple

Soit $a_1, \dots, a_{n+1}$ des réels deux à deux distincts. Considérons l'application linéaire

$$\Phi : \left\{ \begin{array}{ccc} \mathbb{R}_n[X] & \to & \mathbb{R}^{n+1} \\ P & \mapsto & (P(a_1), \dots, P(a_{n+1})) \end{array} \right.$$

Alors, Φ est un isomorphisme ; en effet $\dim R_n[X] = n+1 = \dim \mathbb{R}^{n+1}$ et Φ est injective (tout polynôme du noyau est de degré au plus n et admet au moins $n+1$ racines donc est nul).

Ceci traduit une nouvelle fois la propriété d'interpolation de Lagrange (et les réels $P(a_k)$ pour tout $k \in [\![1, n+1]\!]$ sont les coordonnées dans la base des polynômes interpolateurs de Lagrange associés aux $a_1, \ldots, a_{n+1}$).

Exemple

Soit A un polynôme de degré $n \geq 1$. Considérons l'application linéaire

$$\Phi : \left\{ \begin{array}{ccc} \mathbb{R}_p[X] \times \mathbb{R}_{n-1}[X] & \to & \mathbb{R}_{n+p}[X] \\ (Q, R) & \mapsto & A.Q + R \end{array} \right.$$

Alors, Φ est un isomorphisme.

▹ D'une part, il y a égalité des dimensions des deux espaces considérés :

$$\begin{aligned} \dim \mathbb{R}_p[X] \times \mathbb{R}_{n-1}[X] &= \dim \mathbb{R}_p[X] + \dim \mathbb{R}_{n-1}[X] \\ &= p + 1 + (n-1) + 1 = n + p + 1 \\ &= \dim \mathbb{R}_{n+p}[X]. \end{aligned}$$

▹ D'autre part, l'application linéaire Φ est injective :

$$\begin{aligned} (Q, R) \in \operatorname{Ker} \Phi &\Leftrightarrow A.Q + R = 0 \\ &\Leftrightarrow A.Q = -R \\ &\Leftrightarrow \deg A + \deg Q = \deg R \text{ et } A.Q = -R \\ &\Leftrightarrow \deg Q = 0 \text{ et } A.Q = -R \\ &\Leftrightarrow Q = R = 0. \end{aligned}$$

D'après la proposition précédente, Φ est surjective : cela traduit la propriété de division euclidienne dans $\mathbb{R}_{n+p}[X]$.

Terminons cette section avec quelques propriétés du rang.

Proposition 15.33.

Soit $u \in \mathscr{L}(E, F)$ et $v \in \mathscr{L}(F, G)$ où E, F sont de dimension finie. Alors,
▹ $\operatorname{rg}(v \circ u) \leq \min(\operatorname{rg} u, \operatorname{rg} v)$;
▹ si u est surjective, alors $\operatorname{rg}(v \circ u) = \operatorname{rg} v$;
▹ si v est injective, alors $\operatorname{rg}(v \circ u) = \operatorname{rg} u$.

Démonstration

▹ Comme $\operatorname{Im}(v \circ u) \subset \operatorname{Im} v$, $\operatorname{rg}(v \circ u) \leq \operatorname{rg} v$. Comme $\operatorname{Ker}(v \circ u) \supset \operatorname{Ker} u$, la formule du rang donne

$$\operatorname{rg}(v \circ u) = \dim E - \dim \operatorname{Ker}(v \circ u) \leq \dim E - \dim \operatorname{Ker} u = \operatorname{rg}(u).$$

▹ Si u est surjective, alors

$$\operatorname{Im}(v \circ u) = v \circ u(E) = v(F) = \operatorname{Im} v.$$

▹ Si v est injective, alors la restriction de v entre $\operatorname{Im} u$ et $v(\operatorname{Im} u)$ est un isomorphisme d'où

$$\dim \operatorname{Im} u = \dim v(\operatorname{Im} u) = \dim \operatorname{Im} v \circ u.$$

■

On en déduit facilement le corollaire suivant.

Corollaire 15.34.

Soit u une application linéaire entre deux sous-espaces E et F de dimension finie, f un isomorphisme entre E_1 et E et g un isomorphisme entre F et F_1. Alors,

$$\mathrm{rg}(g \circ u \circ f) = \mathrm{rg}(u).$$

Remarque

Introduisons la relation d'équivalence sur les endomorphismes d'un espace vectoriel E de dimension finie. Les endomorphismes u, v sont équivalents s'il existe des automorphismes f et g de E tels que

$$v = g \circ u \circ f.$$

La proposition précédente indique que le rang est constant sur les classes d'équivalence d'endomorphismes. En fait, on peut montrer que si deux endomorphismes sont de même rang, alors ils sont équivalents.

Il est important de comprendre que l'application rg n'est pas linéaire mais que l'on peut toutefois la manipuler assez facilement.

Proposition 15.35.

Soit u, v deux applications linéaires d'un espace de dimension finie E dans un espace F et λ un scalaire non nul.
▹ $\mathrm{rg}(\lambda u) = \mathrm{rg}(u)$.
▹ $\mathrm{rg}(u+v) \leq \mathrm{rg}(u) + \mathrm{rg}(v)$.

Démonstration

▹$\mathrm{Im}(\lambda u) = \mathrm{Im}\, u$ car pour tout $x \in E$, $\lambda u(x) = u(\lambda x)$.
▹ $\mathrm{Im}(u+v) \subset \mathrm{Im}(u) + \mathrm{Im}(v)$ car pour tout $x \in E$, $u(x) + v(x) \in \mathrm{Im}(u) + \mathrm{Im}(v)$. ■

3.2. Noyaux itérés

Cette partie (hors-programme) développe un exemple fréquemment utilisé dans les exercices.

Proposition 15.36.

Soit u un endomorphisme de E. Alors, pour tout $k \in \mathbb{N}$, $\mathrm{Ker}\, u^k \subset \mathrm{Ker}\, u^{k+1}$ et $\mathrm{Im}\, u^{k+1} \subset \mathrm{Im}\, u^k$.

En d'autres termes, la suite des noyaux itérés est croissante (pour l'inclusion) alors que la suite des images itérées est décroissante.

Démonstration

▹ Soit $x \in \mathrm{Ker}\, u^k$. Alors, $u^{k+1}(x) = u(u^k(x)) = u(0_E) = 0_E$ d'où $x \in \mathrm{Ker}\, u^{k+1}$.
▹ Soit $y \in \mathrm{Im}\, u^{k+1}$ et $x \in E$ tel que $y = u^{k+1}(x)$. Alors, $y = u^k(u(x)) \in \mathrm{Im}\, u^k$. ■

Proposition 15.37.

Soit u un endomorphisme de E et $k \in \mathbb{N}$ tel que $\mathrm{Ker}\, u^k = \mathrm{Ker}\, u^{k+1}$. Alors, pour tout $p \in \mathbb{N}$, $\mathrm{Ker}\, u^{k+p} = \mathrm{Ker}\, u^k$.

Remarque

▹ La suite des noyaux itérés est donc strictement croissante puis stationnaire.

▹ De même, la suite des images itérées est strictement décroissantes puis stationnaire.

Démonstration

Montrons le résultat par récurrence sur p.

▹ Le résultat est évident pour $p=0$ et, par hypothèse, pour $p=1$.

▹ Soit $p \in \setminus\{0\}$ tel que $\operatorname{Ker} u^{k+p} = \operatorname{Ker} u^k$. On sait d'après la proposition précédente que $\operatorname{Ker} u^{k+p} \subset \operatorname{Ker} u^{k+p+1}$. Montrons l'autre inclusion : soit $x \in \operatorname{Ker} u^{k+p+1}$; alors, $u(x) \in \operatorname{Ker} u^{k+p} = \operatorname{Ker} u^{k+p-1}$ (hypothèse de récurrence) donc $u^{k+p}(x) = u^{k+p-1}(u(x)) = 0_E$ soit $x \in \operatorname{Ker} u^{k+p}$. ■

Remarque

Si l'espace E est de dimension finie, la suite croissante $(\dim \operatorname{Ker} u^k)_k$ est majorée donc converge ; comme il s'agit d'une suite d'entiers, elle est stationnaire ce qui assure l'existence de l'entier k de la proposition précédente.

En revanche, en dimension infinie, la suite des noyaux peut être strictement croissante (par exemple avec l'endomorphisme de dérivation dans $\mathbb{R}[X]$).

Proposition 15.38.

Soit u un endomorphisme de E supposé de dimension finie. La suite $(\dim \operatorname{Ker} u^{k+1} - \dim \operatorname{Ker} u^k)_k$ des sauts de dimensions des noyaux itérés est décroissante.

Démonstration

Considérons l'application linéaire

$$v : \left\{ \begin{array}{ccc} \operatorname{Ker} u^{k+1} & \longrightarrow & \operatorname{Ker} u \\ x & \longmapsto & u^k(x) \end{array} \right.$$

et appliquons-lui la formule du rang :

$$\begin{aligned} \dim \operatorname{Ker} u^{k+1} &= \dim \operatorname{Im} v + \dim \operatorname{Ker} v \\ &= \dim \operatorname{Im} u^k \cap \operatorname{Ker} u + \dim \operatorname{Ker} u^k \cap \operatorname{Ker} u^{k+1} \\ &= \dim \operatorname{Im} u^k \cap \operatorname{Ker} u + \dim \operatorname{Ker} u^k. \end{aligned}$$

Ainsi,

$$\dim \operatorname{Ker} u^{k+1} - \dim \operatorname{Ker} u^k = \dim \operatorname{Im} u^k \cap \operatorname{Ker} u.$$

Or, d'après la décroissance de la suite des images itérées, la suite $(\dim \operatorname{Im} u^k \cap \operatorname{Ker} u)$ est décroissante. ■

Exemple

Soit u un endomorphisme nilpotent d'indice n (tel que $u^n = 0_{\mathscr{L}(E)}$ et $u^{n-1} \neq 0_{\mathscr{L}(E)}$) dans un espace vectoriel E de dimension n. Alors, les différents noyaux $\operatorname{Ker} u^k$ pour $k \in [\![0, n]\!]$ sont distincts est de dimensions entre 0 et n. Ainsi, pour tout $k \in [\![0, n]\!]$, $\dim \operatorname{Ker} u^k = k$.

Exemple

Soit u un endomorphisme nilpotent d'indice $n-1$ (tel que $u^{n-1} = 0_{\mathscr{L}(E)}$ et $u^{n-2} \neq 0_{\mathscr{L}(E)}$) dans un espace vectoriel E de dimension n. Alors, $\dim \operatorname{Ker} u^1 - \dim \operatorname{Ker} u^0 = \dim \operatorname{Ker} u$ est supérieure ou égale à $\dim \operatorname{Ker} u^{k+1} - \dim \operatorname{Ker} u^k$ pour tout $k \in [\![1, n-2]\!]$. Or, la somme de ces $n-1$ termes vaut

(après télescopage) $\dim \operatorname{Ker} u^{n-1} = \dim E = n$. Par conséquent, $\dim \operatorname{Ker} u = 2$ (et tous les autres termes valent 1). En conclusion, pour tout $k \in [\![1, n-1]\!]$, $\dim \operatorname{Ker} u^k = k+1$.

3.3. Formes linéaires

Définition 15.39.

Soit E un espace vectoriel de dimension n et $(e_1, \dots, e_n)$ une base de E. La base duale de $(e_1, \dots, e_n)$ est la famille $(e_1^*, \dots, e_n^*)$ des formes linéaires coordonnées, c'est-à-dire la famille des formes linéaires vérifiant

$$\forall (i,j) \in [\![1,n]\!]^2, \qquad e_i^*(e_j) = \delta_{i,j} = \begin{cases} 1 & \text{si } i = j, \\ 0 & \text{sinon.} \end{cases}$$

Exemple

▹ La base duale de $(1, i)$ (de $\mathbb{C}$ comme $\mathbb{R}$-espace vectoriel) est $(\mathfrak{Re}, \mathfrak{Im})$.
▹ La base duale (de $\mathbb{R}_n[X]$) de $((X-a)^k)_{k\in[0,n]}$ est $(P \mapsto \frac{1}{k!}P^{(k)}(a))_{k\in[0,n]}$.
▹ La base duale (de $\mathbb{R}_n[X]$) de la base des polynômes de Lagrange associés aux réels $a_0 < \ldots < a_n$ est $(P \mapsto P(a_k))_{k\in[0,n]}$.

Proposition 15.40.

Soit E un espace vectoriel de dimension n et $(e_1, \dots, e_n)$ une base de E. La base duale de $(e_1, \dots, e_n)$ est une base de l'espace dual E^* (autrement dit, une base duale est une base du dual).
Par conséquent E^* est de dimension finie et $\dim E^* = \dim E$.

Démonstration

Remarquons tout d'abord que la famille $(e_1^*, \dots, e_n^*)$ est libre. En effet, si $\sum_{i=1}^n \lambda_i e_i^* = 0$, alors en évaluant en chaque e_i, on obtient que $\lambda_i = 0$. Montrons maintenant que cette famille est génératrice de E^*. Par définition des formes linéaires coordonnées, tout vecteur $x \in E$ s'écrit $x = \sum_{i=1}^n e_i^*(x).e_i$. Ainsi, pour toute forme linéaire $\varphi \in E^*$,

$$\forall x \in E, \qquad \varphi(x) = \sum_{i=1}^n e_i^*(x)\varphi(e_i).$$

Donc $\varphi = \sum_{i=1}^n \varphi(e_i).e_i^*$. En conclusion, $(e_1^*, \dots, e_n^*)$ est une base de E^*. ■

Exemple

Soit $a_0 < \ldots < a_n$. Les formes linéaires $(\varphi_k : P \mapsto P(a_k))_{k\in[0,n]}$ constituent une base de l'espace dual de $\mathbb{C}_n[X]$ (car c'est la base duale de la base des polynômes interpolateurs de Lagrange aux points $a_0 < \ldots < a_n$).
Par conséquent, la forme linéaire $\psi : P \mapsto \int_0^1 P(t)\,\mathrm{d}t$ s'écrit sur cette base : il existe des scalaires $c_0, \dots, c_n$ tels que

$$\psi = \sum_{k=0}^n c_k \varphi_k,$$

donc, il existe des scalaires $c_0, \dots, c_n$ tels que

$$\forall P \in \mathbb{C}_n[X], \qquad \int_0^1 P(t)\,\mathrm{d}t = \sum_{k=0}^n c_k P(a_k).$$

3.4. Formes linéaires et hyperplans

Définition 15.41.

Un hyperplan est le noyau d'une forme linéaire non nulle.

Exemple Équation d'un hyperplan de $\mathbb{K}^n$

Soit $\varphi \in (\mathbb{K}^n)^*$ non nulle, $H = \operatorname{Ker}\varphi$ un hyperplan de $\mathbb{K}^n$, $(e_1^*, \dots, e_n^*)$ la base duale de la base canonique. Comme φ est non nulle, il existe des scalaires $a_1, \dots, a_n$ non tous nuls tels que

$$\varphi = \sum_{k=1}^{n} a_k e_k^*.$$

Ainsi, un vecteur $x = (x_1, \dots, x_n)$ appartient à H si, et seulement s'il vérifie l'équation

$$\sum_{k=1}^{n} a_k x_k = 0.$$

On reconnaît la forme obtenue pour les droites vectorielles dans $\mathbb{R}^2$ et pour les plans dans $\mathbb{R}^3$.

Proposition 15.42.

Soit H un hyperplan de E et $x \notin H$. L'hyperplan H et la droite $\operatorname{Vect}(x)$ sont supplémentaires dans E.

Démonstration

Soit $\varphi \in E^*$ non nulle telle que $H = \operatorname{Ker}\varphi$. Montrons le caractère supplémentaire par analyse-synthèse.

▹ Soit $y \in E$ tel que $y = z + \lambda x$ avec $z \in H$ et $\lambda \in \mathbb{K}$. Alors, $\varphi(y) = \lambda\varphi(x)$ et comme $\varphi(x) \neq 0$ (car $x \notin H = \operatorname{Ker}\varphi$),

$$\lambda = \frac{\varphi(y)}{\varphi(x)}, \quad z = y - \lambda x.$$

▹ Soit $y \in E$; posons $\lambda = \frac{\varphi(y)}{\varphi(x)}$ et $z = y - \lambda x$. Alors, $y = z + \lambda x$ avec $z \in \operatorname{Ker}\varphi$ et $\lambda x \in \operatorname{Vect}(x)$.

En conclusion, H et $\operatorname{Vect}(x)$ sont supplémentaires dans E. ■

Proposition 15.43.

Deux formes linéaires non nulles sur E un espace vectoriel de dimension $n \geq 1$ définissent le même hyperplan si, et seulement si, elles sont proportionnelles.

Démonstration

▹ Le sens retour est évident.

▹ Soit φ et ψ deux formes linéaires non nulles telles que $H = \operatorname{Ker}\varphi = \operatorname{Ker}\psi$ et $x \notin H$. Le sous-espace $\operatorname{Vect}(x)$ est un supplémentaire de H. Posons $\lambda = \frac{\varphi(x)}{\psi(x)}$ et montrons que $\varphi = \lambda\psi$.

Pour tout $y \in E$, il existe $z \in H$ et $\mu \in \mathbb{K}$ tel que $y = z + \mu x$. Alors,

$$\varphi(y) = \varphi(z) + \mu\varphi(x) = \mu\varphi(x),$$
$$\frac{\varphi(x)}{\psi(x)}\psi(y) = \frac{\varphi(x)}{\psi(x)}\psi(z) + \mu\frac{\varphi(x)}{\psi(x)}\psi(x) = \mu\varphi(x),$$

d'où $\varphi = \lambda\psi$. ■

Proposition 15.44.

Une forme linéaire non nulle est de rang 1.

Démonstration

L'image d'une forme linéaire est un sous-espace du corps de base qui est de dimension 1 : c'est donc soit $\{0_K\}$, ce qui est exclu par hypothèse, soit $\mathbb{K}$. ■

Proposition 15.45.

Soit E un espace vectoriel de dimension $n \geq 1$. Les hyperplans de E sont les sous-espaces de dimension $n-1$.

Démonstration

▹ D'après la formule du rang, la dimension d'un hyperplan est $n-1$.

▹ Réciproquement, si H est de dimension $n-1$, alors on considère une base $(e_1,\dots,e_{n-1})$ de H que l'on complète en une base de E avec un vecteur e_n. Alors, H est le noyau de la forme linéaire non nulle e_n^*.

■

Remarque

Un hyperplan affine est le translaté d'un hyperplan, c'est-à-dire un sous-espace affine de dimension $n-1$.

Généralisons le résultat précédent.

Proposition 15.46.

Soit E un espace vectoriel de dimension n et F un sous-espace vectoriel de E. F est de dimension $n-p$ (on dit alors de codimension p) si, et seulement s'il existe p formes linéaires indépendantes $\varphi_1,\dots,\varphi_p$ telles que

$$F = \bigcap_{k=1}^{p} \operatorname{Ker} \varphi_k.$$

Démonstration

($\Rightarrow$) Complétons une base $(e_1,\dots,e_{n-p})$ de F (donc une famille libre de E) en une base $(e_1,\dots,e_n)$ de E. Alors,

$$F = \bigcap_{k=n-p+1}^{n} \operatorname{Ker} e_k^* = \bigcap_{k=1}^{p} \operatorname{Ker} e_{n-p+k}^*.$$

($\Leftarrow$) Réciproquement, complétons la famille libre $(\varphi_1,\dots,\varphi_p)$ en une base $(\varphi_1,\dots,\varphi_n)$ de E^* et considérons $(e_1,\dots,e_n)$ la base dont $(\varphi_1,\dots,\varphi_n)$ est la base duale (qui est l'image réciproque de la base canonique dans l'application $E \to \mathbb{R}^n$ qui à x associe $(\varphi_1(x),\dots,\varphi_n(x))$). Alors,

$$\bigcap_{k=1}^{p} \operatorname{Ker} \varphi_k = \operatorname{Vect}(e_{n-p+1},\dots,e_n).$$

■

Remarque

Plus généralement, pour toutes formes linéaires $\varphi_1,\dots,\varphi_p \in E^*$,

$$\dim \bigcap_{k=1}^{p} \operatorname{Ker} \varphi_k = n - \operatorname{rg}(\varphi_1,\dots,\varphi_p).$$

FICHE

SYNTHÈSE

Dans ce chapitre, on définit la dimension, on apprend à la calculer puis on exploite des identités entre les dimensions.

Obtention de base

Soit E un $\mathbb{K}$-espace vectoriel de dimension finie différent de $\{0_E\}$.
▹ On peut extraire une base de E de toute famille génératrice finie de E.
▹ On peut compléter toute famille libre de E en une base de E.
▹ Toutes les bases de E ont le même cardinal, appelé dimension de E.

Voici quelques formules de calcul de la dimension (ici énoncées pour deux espaces) ; pour pouvoir les utiliser, il est bon de connaître la dimension des espaces les plus courants.

Calculs de dimension

▹ Soit E et E' deux espaces vectoriels de dimension finie. Alors,

$$\dim E \times E' = \dim E + \dim E'.$$

▹ Soit E et E' deux espaces vectoriels de dimension finie. Alors,

$$\dim \mathscr{L}(E, E') = \dim E' \times \dim E.$$

▹ Soit F un sous-espace vectoriel d'un espace E de dimension finie. Alors, $\dim F \leq \dim E$ avec égalité si, et seulement si, $F = E$.
▹ Soit F et F' deux sous-espaces vectoriels d'un espace de dimension finie. Alors,

$$\dim F + F' = \dim F + \dim F' - \dim F \cap F'.$$

Lorsque F et F' sont en somme directe, $\dim F + F' = \dim F + \dim F'$.

Sans exhiber une base, on peut calculer la dimension directement en considérant un isomorphisme avec un espace mieux connu puisque les isomorphismes préservent la dimension.
Lorsque l'on retire la propriété de bijection, l'évolution de la dimension est décrite par la formule du rang.

Formule du rang

Soit u une application linéaire entre deux sous-espaces E et F avec E de dimension finie. Alors,

$$\dim E = \operatorname{rg} u + \dim \operatorname{Ker} u.$$

On utilise souvent la dimension pour remplacer une propriété (existence ou unicité, générateur ou libre, surjectif ou injectif).

Propriétés de dimension finie

▹ Soit F, G deux sous-espaces d'un espace E de dimension finie. Deux des trois propriétés entraînent la troisième :

- $E = F + G$;
- F et G sont en somme directe ;
- $\dim E = \dim F + \dim G$.

▹ Soit $(x_1, \ldots, x_p)$ une famille d'un espace E de dimension finie. Deux des trois propriétés entraînent la troisième :

- $(x_1, \ldots, x_p)$ est génératrice ;
- $(x_1, \ldots, x_p)$ est libre ;
- $\dim E = p$.

▹ Soit u une application linéaire entre E et F deux espaces de dimension finie. Deux des trois propriétés entraînent la troisième :

- u est surjective ;
- u est injective ;
- $\dim E = \dim F$.

EXERCICES

Vrai ou faux ?

		Vrai	Faux
a)	Une famille de vecteurs deux à deux non colinéaires est libre.	☐	☐
b)	Si $(e_1,\dots,e_n)$ est une famille libre de E et $x \in E$, alors $(e_1+x,\dots,e_n+x)$ est une famille libre.	☐	☐
c)	L'espace des fonctions de $\mathscr{C}^0(\mathbb{R},\mathbb{R})$ nulles en 0 est de dimension finie.	☐	☐
d)	L'espace des fonctions de $\mathscr{C}^0(\mathbb{R},\mathbb{R})$ périodiques et nulles en 0 est de dimension finie.	☐	☐
e)	Soit $N>0$. L'espace des suites de $\mathbb{R}^{\mathbb{N}}$ N-périodiques est de dimension finie.	☐	☐
f)	De toute famille génératrice d'un espace de dimension finie, on peut extraire une base.	☐	☐
g)	Si $(f_1,\dots,f_n)$ est une base de F et $(g_1,\dots,g_p)$ est une base de G, alors $\{f_1,\dots,f_n\}\cap\{g_1,\dots,g_p\}$ est une base de $F\cap G$.	☐	☐
h)	Si $(f_1,\dots,f_n)$ est une base de F, $(g_1,\dots,g_p)$ est une base de G et $(f_1,\dots,f_n,g_1,\dots,g_p)$ est une base de E, alors $E=F\oplus G$.	☐	☐
i)	Tout sous-espace de dimension $n-1$ d'un espace vectoriel de dimension $n\geq 1$ est le noyau d'une forme linéaire non nulle.	☐	☐
j)	L'intersection de trois hyperplans deux à deux distincts d'un espace vectoriel de dimension $n\geq 3$ est de dimension $n-3$.	☐	☐
k)	Soit F un sous-espace vectoriel d'un espace vectoriel de dimension n et H un hyperplan. Alors, $\dim F\cap H=\dim F-1$.	☐	☐
l)	Soit H un hyperplan et $x\notin H$. $\mathrm{Vect}(x)$ est un supplémentaire de H.	☐	☐
m)	Soit $u\in\mathscr{L}(E)$ avec E de dimension finie et F un sous-espace vectoriel de E tel que $F\subset u(F)$. Alors, $F=u(F)$.	☐	☐
n)	Si $u\in\mathscr{L}(E,F)$ avec E, F de dimension finie et u injective, alors $\dim E\leq\dim F$.	☐	☐
o)	Si $u\in\mathscr{L}(E,F)$ avec E, F de dimension finie et $\dim E\geq\dim F$, alors u est surjective.	☐	☐

Exercices d'application du chapitre 15

●○○ **Exercice 1**

Soit $n \geq 1$ et $P \in \mathbb{R}_n[X]$ non nul.

1. Montrer que l'ensemble F_P des polynômes de $\mathbb{R}_n[X]$ multiples de P est un sous-espace vectoriel de $\mathbb{R}_n[X]$.
2. En déduire la dimension de F_P en fonction du degré de P.
3. Soit $Q \in \mathbb{R}_n[X]$ un polynôme premier avec P et tel que $\deg P + \deg Q = n+1$.
   Montrer que $\mathbb{R}_n[X] = F_P \oplus F_Q$.

●○○ **Exercice 2**

Soit E un espace vectoriel de dimension finie n des sous-espaces $E_1, \ldots, E_p$, $F_1, \ldots, F_p$ tels que

- pour tout $i \in [\![1, p]\!]$, $F_i \subset E_i$,
- $\bigoplus_{i=1}^{p} E_i = \bigoplus_{i=1}^{p} F_i$.

Montrer que, pour tout $i \in [\![1, p]\!]$, $F_i = E_i$.

●●○ **Exercice 3**

Soit $a_0 < a_1 < \ldots < a_N$.

1. Montrer que l'espace E des fonctions continues sur $[a_0, a_N]$, affines par morceaux pour la subdivision $a_0 < a_1 < \ldots < a_N$ est de dimension $N+1$.
2. Déterminer une base de E.

●●● **Exercice 4**

Montrer qu'un sous-espace vectoriel de $\mathbb{R}^{\mathbb{R}}$ ne contenant que des applications de signe constant est de dimension au plus 1.

●●○ **Exercice 5**

Soit $(a, b, c) \in (\mathbb{R}^*)^3$. Définissons $\mathscr{S}$ l'ensemble des suites réelles satisfaisant la relation de récurrence, pour tout $n \in \mathbb{N}$,

$$u_{n+3} - (a+b+c)u_{n+2} + (ab+bc+ac)u_{n+1} - abc\,u_n = 0.$$

1. Montrer que $\mathscr{S}$ est un espace vectoriel et déterminer les suites géométriques de $\mathscr{S}$.
2. Montrer que l'application
   $$\varphi : \left\{ \begin{array}{ccc} \mathscr{S} & \to & \mathbb{R}^3 \\ (u_n)_{n\in\mathbb{N}} & \mapsto & (u_0, u_1, u_2) \end{array} \right.$$
   est un isomorphisme.
3. Déterminer une base de $\mathscr{S}$.
   On pourra commencer par les cas a, b et c deux à deux distincts.

●●○ **Exercice 6**

Un polynôme trigonométrique de degré au plus n est une fonction

$$T : \left\{ \begin{array}{ccc} \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & \frac{a_0}{2} + \sum_{k=1}^{n} (a_k \cos(kx) + b_k \sin(kx)) \end{array} \right.$$

avec $(a_0, \ldots, a_n, b_1, \ldots, b_n) \in \mathbb{R}^{2n+1}$. Définissons $\mathscr{T}_n$ l'ensemble des polynômes trigonométriques de degré au plus n.

1. Montrer que $\mathscr{T}_n$ est un espace vectoriel.
2. Soit $T \in \mathscr{T}_n$. Calculer, pour tout entier k, les intégrales
$$\int_{-\pi}^{\pi} \sin(kx)T(x)\,\mathrm{d}x \text{ et } \int_{-\pi}^{\pi} \cos(kx)T(x)\,\mathrm{d}x\,.$$
3. Montrer que la famille composée des fonctions $x \mapsto \cos(kx)$ pour $k \in [\![0, n]\!]$ et des fonctions $x \mapsto \sin(jx)$ pour $j \in [\![1, n]\!]$ est une base de $\mathscr{T}_n$. En déduire la dimension de $\mathscr{T}_n$.

●●○ **Exercice 7**

Soit E de dimension finie et $u,\ v \in \mathscr{L}(E)$ vérifiant $u^2 + u \circ v = \mathrm{Id}$. Montrer que u et v commutent.

●●● **Exercice 8**

Soit E un espace vectoriel de dimension finie, $u \in \mathscr{L}(E)$ nilpotent d'indice p (tel que $u^p = 0_{\mathscr{L}(E)}$ et $u^{p-1} \neq 0_{\mathscr{L}(E)}$) et $\Phi : v \in \mathscr{L}(E) \mapsto u \circ v - v \circ u$.

1. Montrer que, pour tout $n \in \mathbb{N}$ et tout $v \in \mathscr{L}(E)$,
$$\Phi^n(v) = \sum_{k=0}^{n} (-1)^k \binom{n}{k} u^{n-k} \circ v \circ u^k.$$
2. Montrer que Φ est nilpotente et majorer son indice de nilpotence.
3. Soit $a \in \mathscr{L}(E)$. Montrer qu'il existe $b \in \mathscr{L}(E)$ tel que $a \circ b \circ a = a$.
4. En déduire l'indice de nilpotence de Φ.

●●○ **Exercice 9**

Soit E, F deux espaces vectoriels de dimension finie et $u \in \mathscr{L}(E)$. Déterminer la dimension du sous-espace
$$\{v \in \mathscr{L}(E, F),\ v \circ u = 0_{\mathscr{L}(E,F)}\}.$$

●○○ **Exercice 10**

Soit u et v deux endomorphismes d'un espace vectoriel de dimension finie. Montrer que $\mathrm{rg}(u+v) \leq \mathrm{rg}(u) + \mathrm{rg}(v)$ puis que $|\mathrm{rg}(u) - \mathrm{rg}(v)| \leq \mathrm{rg}(u+v)$.

●○○ **Exercice 11**

Soit u et v deux endomorphismes d'un espace vectoriel E de dimension finie.

1. Préciser le noyau et l'image de l'application linéaire induite par v entre les espaces $\mathrm{Im}(u)$ et E.
2. Montrer que $\mathrm{rg}(u) = \mathrm{rg}(v \circ u) + \dim \mathrm{Ker}(v) \cap \mathrm{Im}(u)$.
3. En déduire que $\dim \mathrm{Ker}(v \circ u) \leq \dim \mathrm{Ker}(v) + \dim \mathrm{Ker}(u)$ et en particulier que, pour tout $k \in \mathbb{N}$,
$$\dim \mathrm{Ker}(u^k) \leq k.\dim \mathrm{Ker}(u), \mathrm{rg}(u^k) \geq k.\mathrm{rg}(u) - n(k-1).$$

●●● **Exercice 12**

Soit E de dimension finie et $u,\ v \in \mathscr{L}(E)$ tel que $u + v$ est bijectif.

1. Montrer que si $v \circ u = 0_{\mathscr{L}(E)}$, alors $\mathrm{rg}\, u + \mathrm{rg}\, v = \dim E$.
2. Montrer que si $\mathrm{rg}\, u + \mathrm{rg}\, v = \dim E$, alors il existe un projecteur p tel que
$$u = p \circ (u+v), \qquad v = (\mathrm{Id} - p) \circ (u+v).$$

EXERCICES

●●○ **Exercice 13**

Soit E un espace vectoriel de dimension n et $u \in \mathscr{L}(E)$ nilpotent. On rappelle que la suite $(\operatorname{Ker} u^k)_k$ est strictement croissante puis stationnaire et que la suite $(\dim \operatorname{Ker} u^{k+1} - \dim \operatorname{Ker} u^k)_k$ est décroissante.

1. Montrer que l'indice de nilpotence de u est n si, et seulement si, $\dim \operatorname{Ker} u = 1$.
2. Montrer que, si l'indice de nilpotence de u est $n-1$, alors $\dim \operatorname{Ker} u \cap \operatorname{Im} u = 1$.

●○○ **Exercice 14**

Montrer qu'un sous-espace F est stable par un endomorphisme u de rang 1 si, et seulement si, $\operatorname{Im} u \subset F$ ou $F \subset \operatorname{Ker} u$.

●●○ **Exercice 15**

Soit $u \in \mathscr{L}(\mathbb{R}^3)$ nilpotent d'ordre 2. Montrer les sous-espaces de dimension 2 stables par u sont ceux qui contiennent $\operatorname{Im} u$.

●●● **Exercice 16**

Soit E un espace vectoriel de dimension finie, $u \in \mathscr{L}(E)$ tel que $u^n = 0_{\mathscr{L}(E)}$ pour un $n \in \mathbb{N} \setminus \{0\}$ et F un sous-espace vectoriel de E stable par u vérifiant $E = F + \operatorname{Im} u$. Montrer que $F = E$.

●●○ **Exercice 17**

Définissons $P_0 = 1$ et, pour tout $k \in \mathbb{N} \setminus \{0\}$, $P_k(X) = \frac{1}{k!} X(X-k)^{k-1}$.

1. Montrer que, pour tout $k \in \mathbb{N} \setminus \{0\}$, $P_k'(X) = P_{k-1}(X-1)$.
2. Montrer que la famille $(P_0, \dots, P_n)$ est une base de $\mathbb{R}_n[X]$ puis que sa base duale est $(\varphi_0, \dots, \varphi_n)$ où
$$\varphi_k : P \mapsto P^{(k)}(k).$$

●●● **Exercice 18**

Soit $Q \in \mathbb{R}_n[X]$ de degré n ; posons, pour tout $i \in [\![0, n]\!]$, $Q_i(X) = Q(X+i)$.

1. Montrer que la famille $(Q, Q', Q'', \dots, Q^{(n)})$ est une base de $\mathbb{R}_n[X]$.
2. Montrer que, pour toute forme linéaire φ sur $\mathbb{R}_n[X]$, il existe des scalaires $\alpha_0, \dots, \alpha_n$ tels que, pour tout $P \in \mathbb{R}_n[X]$,
$$\varphi(P) = \alpha_0 P(0) + \alpha_1 P'(0) + \cdots + \alpha_n P^{(n)}(0).$$
3. Soit une forme linéaire φ sur $\mathbb{R}_n[X]$ telle que $\varphi(Q_0) = \dots = \varphi(Q_n) = 0$. Montrer que φ est nulle. On pourra considérer le polynôme $\alpha_0 Q + \cdots + \alpha_n Q^{(n)}$ avec $\alpha_0, \dots, \alpha_n$ les scalaires associés à φ dans la question précédente.
4. Conclure que la famille $(Q_0, \dots, Q_n)$ est une base de $\mathbb{R}_n[X]$.

●●● **Exercice 19**

Soit E un espace vectoriel de dimension n. Un endomorphisme f de E est cyclique s'il existe $x_0 \in E$ tel que $(x_0, f(x_0), \dots, f^{n-1}(x_0))$ soit une base de E.

1. Montrer que si $u^n = 0$ et $u^{n-1} \neq 0$ alors u est cyclique.
2. Soit $u \in \mathscr{L}(E)$ cyclique. Montrer qu'il existe des scalaires $(a_k)_{k \in [0, n-1]}$ tels que
$$u^n + \sum_{k=0}^{n-1} a_k u^k = 0.$$

Exercices d'approfondissement du chapitre 15

●●● **Exercice A**

Soit E un espace vectoriel de dimension finie et $(x_1, \dots, x_N)$ une famille génératrice de E. Notons

$$I = \{i \in [\![1, N]\!],\ x_i \notin \mathrm{Vect}(x_1, \dots, x_{i-1})\}.$$

Montrer que $(x_i)_{i \in I}$ est une base de E.

●○○ **Exercice B**

Soit $x_1, \dots, x_n$ des vecteurs d'un espace vectoriel E. Montrer que, pour tout $(j, k) \in [\![1, n]\!]^2$ tel que $j \leq k$,

$$\mathrm{rg}(x_1, \dots, x_j) - j \geq \mathrm{rg}(x_1, \dots, x_k) - k.$$

●●○ **Exercice C**

Soit $(f_1, \dots, f_n)$ une famille libre de fonctions dérivables de $\mathbb{R}$ dans $\mathbb{R}$. Montrer que la famille $(f_1', \dots, f_n')$ est de rang au moins $n-1$.

●●● **Exercice D**

Soit $V_1, \dots, V_p$ des sous-espaces d'un espace vectoriel E de dimension finie tels que

$$\sum_{k=1}^{p} \dim V_k > (p-1) \dim E.$$

Montrer que $\bigcap_{k=1}^{p} V_k \neq \{0_E\}$.

●○○ **Exercice E**

Déterminer les endomorphismes de $\mathbb{K}^n$ laissant stables chacun des axes de coordonnées et la droite engendrée par le vecteur $(1, 1, \dots, 1)$.

●●● **Exercice F**

Soit E un $\mathbb{K}$-espace vectoriel et $\varphi, \varphi_1, \dots, \varphi_p$ des formes linéaires sur E.

1. Supposons que $(\varphi_1, \dots, \varphi_p)$ est une famille libre de E^*. Montrer que $\varphi \in \mathrm{Vect}(\varphi_1, \dots, \varphi_p)$ si, et seulement si,
$$\bigcap_{i=1}^{p} \mathrm{Ker}\, \varphi_i \subset \mathrm{Ker}\, \varphi.$$
2. On suppose que
$$\bigcap_{i=1}^{p} \mathrm{Ker}\, \varphi_i = \{0_E\}.$$
Montrer que la famille $(\varphi_1, \dots, \varphi_p)$ est génératrice de E^*.

EXERCICES

Problèmes du chapitre 15

●○○ **Problème 1 Lemme de Hochschild**

Considérons l'espace vectoriel $E = \mathscr{C}^\infty(\mathbb{R}, \mathbb{C})$.

1. Soit $F \neq \{0_E\}$ un sous-espace vectoriel de E de dimension n. Pour tout $x \in \mathbb{R}$, notons φ_x la forme linéaire sur F définie par $\varphi_x : f \mapsto f(x)$.
   a. Justifier qu'il existe $x \in \mathbb{R}$ tel que φ_x soit non nulle.
   b. Soit k le cardinal maximal d'une famille de réels $(x_1, \dots, x_k)$ telle que $(\varphi_{x_1}, \dots, \varphi_{x_k})$ soit libre.
      i. Montrer que $k = n$.
      Indication : on pourra considérer la famille liée $(\varphi_x, \varphi_{x_1}, \dots, \varphi_{x_k})$.
      ii. En déduire qu'il existe des réels $x_1, \dots, x_n$ tels que $(\varphi_{x_1}, \dots, \varphi_{x_n})$ soit une base de F^*.
   c. En déduire qu'il existe une base $(v_1, \dots, v_n)$ de F telle que
   $$\forall i, j \in [\![1, n]\!], \qquad v_i(x_j) = \delta_{i,j}.$$
   Déterminer les coordonnées de $f \in F$ dans la base $(v_1, \dots, v_n)$.
2. Pour tout $a \in \mathbb{R}$, définissons l'endomorphisme τ_a de E défini par
   $$\forall f \in E,\ \forall x \in \mathbb{R}, \qquad \tau_a(f)(x) = f(x + a).$$
   Soit $f \in E$ et F le sous-espace de E engendré par les fonctions $\tau_a(f)$ pour tout $a \in \mathbb{R}$.
   Supposons que F est de dimension finie $n > 0$.
   a. Justifier qu'il existe des réels $(x_1, \dots, x_n)$ et une base $(v_1, \dots, v_n)$ de F tels que
   $$\forall g \in F,\ \forall a, x \in \mathbb{R}, \qquad g(a + x) = \sum_{k=1}^{n} g(a + x_k) v_k(x).$$
   b. En déduire que, pour tout $g \in F$, $g' \in F$.
   c. Montrer qu'il existe une équation différentielle linéaire à coefficients constants dont tous les éléments de F sont solutions.

●●● **Problème 2**

Définition 15.47.

Soit E et F deux $\mathbb{R}$-espaces vectoriels de dimension finie. Deux endomorphismes $u \in \mathscr{L}(E)$ et $v \in \mathscr{L}(F)$ sont croisés s'il existe $a \in \mathscr{L}(E, F)$ et $b \in \mathscr{L}(F, E)$ tels que
$$u = b \circ a, \qquad v = a \circ b.$$

Partie A – Décomposition de Fitting

Soit f un endomorphisme d'un espace vectoriel E de dimension finie. Notons
$$\mathscr{N}_f = \bigcup_{k \in \mathbb{N}} \operatorname{Ker} f^k, \qquad \mathscr{I}_f = \bigcap_{k \in \mathbb{N}} \operatorname{Im} f^k.$$

1. Montrer que la suite $(\operatorname{Ker} f^k)_k$ est croissante pour l'inclusion et que la suite $(\operatorname{Im} f^k)_k$ est décroissante pour l'inclusion.
2. a. Montrer qu'il existe un rang r à partir duquel la suite $(\operatorname{Ker} f^k)_k$ est stationnaire alors qu'elle est strictement croissante avant.
   b. Montrer que la suite $(\operatorname{Im} f^k)_k$ est également stationnaire à partir de ce rang r.
   c. Déterminer ce rang r pour un projecteur.
3. En déduire que $\mathscr{N}_f$ et $\mathscr{I}_f$ sont des sous-espaces vectoriels de E, stables par f.

4. Montrer que $E = \mathscr{N}_f \oplus \mathscr{I}_f$.
5. Montrer que l'endomorphisme induit par f sur $\mathscr{N}_f$ est nilpotent.
   Cet endomorphisme est noté $f_{\mathscr{N}}$.
6. Que dire de l'endomorphisme induit par f sur $\mathscr{I}_f$?
   Cet endomorphisme est noté $f_{\mathscr{I}}$.

Partie B – Cas général

Soit E et F deux $\mathbb{R}$-espaces vectoriels de dimension finie, $u \in \mathscr{L}(E)$ et $v \in \mathscr{L}(F)$. On considère $u_{\mathscr{N}}$, $u_{\mathscr{I}}$, $v_{\mathscr{N}}$ et $v_{\mathscr{I}}$ les endomorphismes associés d'après la partie précédente.

1. a. Supposons $E = E_1 \oplus E_2$ et considérons $f_1 \in \mathscr{L}(E_1, F)$ et $f_2 \in \mathscr{L}(E_2, F)$. Montrer qu'il existe une unique application linéaire $f \in \mathscr{L}(E, F)$ telle que $f(x) = f_1(x)$ si $x \in E_1$ et $f(x) = f_2(x)$ si $x \in E_2$.
   b. Montrer que si $u_{\mathscr{N}}$ et $v_{\mathscr{N}}$ sont croisés et si $u_{\mathscr{I}}$ et $v_{\mathscr{I}}$ sont croisés, alors u et v sont croisés.
2. Supposons qu'il existe $a \in \mathscr{L}(E, F)$ et $b \in \mathscr{L}(F, E)$ tels que
$$u = b \circ a, \qquad v = a \circ b.$$
   a. Montrer que, pour tout $k \in \mathbb{N}$, $a(\operatorname{Ker}(b \circ a)^k) \subset \operatorname{Ker}(a \circ b)^k$; en déduire que $a(\mathscr{N}_u) \subset \mathscr{N}_v$.
   b. Montrer que $a(\mathscr{I}_u) \subset \mathscr{I}_v$.
   c. Énoncer les résultats analogues concernant l'application linéaire b.
   d. Montrer que $u_{\mathscr{N}}$ et $v_{\mathscr{N}}$ sont croisés et que $u_{\mathscr{I}}$ et $v_{\mathscr{I}}$ sont croisés.

À la suite de ces deux parties, on est ramené à étudier le cas des isomorphismes et des applications linéaires nilpotentes : c'est l'objet des deux prochaines parties.

Partie C – Cas des isomorphismes

1. Soit E, F deux $\mathbb{R}$-espaces vectoriels de même dimension finie, deux isomorphismes $u \in \mathscr{L}(E)$ et $v \in \mathscr{L}(F)$ tels qu'il existe un isomorphisme $f \in \mathscr{L}(E, F)$ vérifiant $u = f^{-1} \circ v \circ f$.
   Montrer que u et v sont croisés.
2. Soit E, F deux $\mathbb{R}$-espaces vectoriels de dimension finie, deux isomorphismes croisés $u \in \mathscr{L}(E)$ et $v \in \mathscr{L}(F)$ et $a \in \mathscr{L}(E, F)$ et $b \in \mathscr{L}(F, E)$ tels que
$$u = b \circ a, \qquad v = a \circ b.$$
   a. Montrer que a est un isomorphisme.
   b. En déduire qu'il existe un isomorphisme $f \in \mathscr{L}(E, F)$ vérifiant
$$u = f^{-1} \circ v \circ f.$$

Partie D – Cas des endomorphismes nilpotents

1. Soit f un endomorphisme d'un espace vectoriel E de dimension finie.
   a. Montrer que, pour tout $k \in \mathbb{N}$,
$$\dim \operatorname{Ker} f^{k+1} = \dim \operatorname{Ker} f^k + \dim \operatorname{Ker} f \cap \operatorname{Im} f^k.$$
   Indication : on pourra considérer l'application linéaire
$$\left\{ \begin{array}{ccc} \operatorname{Ker} f^{k+1} & \to & E \\ x & \mapsto & f^k(x) \end{array} \right.$$
   b. En déduire que la suite $(\dim \operatorname{Ker} f^{k+1} - \dim \operatorname{Ker} f^k)_k$ est décroissante.
2. Soit E, F deux $\mathbb{R}$-espaces vectoriels de dimension finie, deux endomorphismes nilpotents croisés $u \in \mathscr{L}(E)$ et $v \in \mathscr{L}(F)$ et $a \in \mathscr{L}(E, F)$ et $b \in \mathscr{L}(F, E)$ tels que
$$u = b \circ a, \qquad v = a \circ b.$$
   Fixons $k \in \mathbb{N}$ et considérons G un supplémentaire de $\operatorname{Ker} u^{k+1}$ dans $\operatorname{Ker} u^{k+2}$, $H_1 = a(G) \cap \operatorname{Ker} v^{k+1}$ et H_2 un supplémentaire de H_1 dans $a(G)$.

PROBLÈMES

a. Montrer que $\dim a(G) = \dim G$ puis que $\dim v(H_2) = \dim H_2$.
b. Montrer que $H_1 \oplus v(H_2) \oplus \operatorname{Ker} v^k \subset \operatorname{Ker} v^{k+1}$.
c. En déduire que

$$\dim \operatorname{Ker} v^{k+1} - \dim \operatorname{Ker} v^k \geq \dim G = \dim \operatorname{Ker} u^{k+2} - \dim \operatorname{Ker} u^{k+1}.$$

On a montré que si deux endomorphismes nilpotents u et v sont croisés, alors, pour tout $k \in \mathbb{N}$,

$$\dim \operatorname{Ker} v^{k+1} - \dim \operatorname{Ker} v^k \geq \dim \operatorname{Ker} u^{k+2} - \dim \operatorname{Ker} u^{k+1},$$
$$\dim \operatorname{Ker} u^{k+1} - \dim \operatorname{Ker} u^k \geq \dim \operatorname{Ker} v^{k+2} - \dim \operatorname{Ker} v^{k+1}.$$

La réciproque est vraie.

CORRIGÉS

Corrigés des Vrai/Faux

a) Faux, par exemple, avec $(1,0)$, $(0,1)$ et $(1,1)$.

b) Faux, par exemple, pour $x = -e_1$.

c) Faux.

d) Faux.

e) Vrai, l'application linéaire $u \mapsto (u_0, \dots, u_{N-1})$ est injective de cet espace dans $\mathbb{R}^N$.

f) Vrai.

g) Faux, les deux familles peuvent être disjointes alors que l'intersection des sous-espaces peut être grande.

h) Vrai.

i) Vrai.

j) Faux.

k) Faux, si $F \subset H$.

l) Vrai.

m) Vrai, car $\dim u(F) \leq \dim F$.

n) Vrai.

o) Faux, u peut très bien être l'application nulle.

Corrigés des exercices d'application du chapitre 15

Exercice 1

1. La partie $F_P \subset \mathbb{R}_n[X]$ est non vide car elle contient le polynôme nul. Par ailleurs, une combinaison linéaire de polynômes multiples de P est encore un multiple de P et le degré d'une combinaison linéaire de polynômes de degré au plus n est de degré au plus n. Par conséquent, d'après la caractérisation, F_P est un sous-espace vectoriel de $\mathbb{R}_n[X]$.
2. D'après la question précédente, un polynôme de F_P est de la forme

$$P(X) \sum_{k=0}^{n-\deg P} a_k X^k,$$

donc une famille génératrice de F_P est $(P, XP, \dots, X^{n-\deg P}P)$. De plus, cette famille est échelonnée en degrés donc libre : c'est une base de F_P et $\dim F_P = n - \deg P + 1$.

3. ▹ Montrons déjà que F_P et F_Q sont en somme directe ; en effet si $A \in F_P \cap F_Q$, alors A est un multiple de P et Q qui sont premiers entre eux donc de PQ. Or, $\deg PQ > n$ et $\deg A \leq n$ donc $A = 0$. L'intersection est réduite à 0 donc les sous-espaces sont en somme directe.
▹ Comme $\dim F_P + \dim F_Q = n - \deg P + 1 + n - \deg Q + 1 = n + 1 = \dim \mathbb{R}_n[X]$, le fait d'être en somme directe entraîne que ces deux sous-espaces sont supplémentaires.

Exercice 2

Raisonnons par l'absurde et supposons qu'il existe un j tel que $F_j \neq E_j$. Ainsi, pour tout i, $\dim F_i \leq \dim E_i$ et $\dim F_j < \dim E_j$. Alors,

$$\dim \bigoplus_{i=1}^{p} E_i = \sum_{i=1}^{n} \dim E_i < \sum_{i=1}^{n} \dim F_i = \dim \bigoplus_{i=1}^{p} F_i,$$

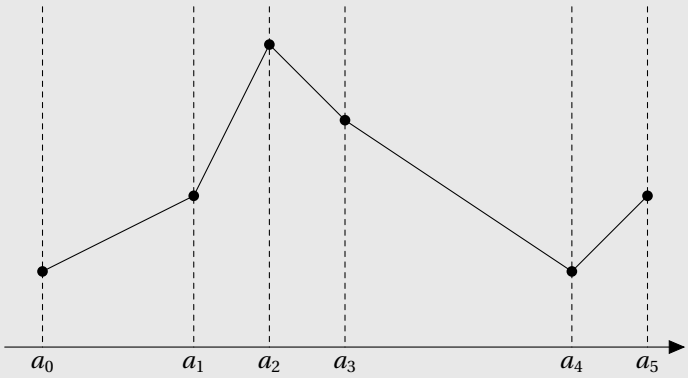
ce qui contredit l'égalité en hypothèse.

Exercice 3

1. L'application suivante, qui associe à une fonction de E ses valeurs aux points de subdivision,

$$u : \left\{ \begin{array}{ccc} E & \to & \mathbb{R}^{N+1} \\ f & \mapsto & (f(a_0), \ldots, f(a_N)) \end{array} \right.$$

est un isomorphisme : la linéarité est évidente et l'unique antécédent de $(b_0, \ldots, b_N) \in \mathbb{R}^{N+1}$ est la fonction dont la courbe est la ligne brisée joignant les points de coordonnées (a_k, b_k).



Par conséquent, $\dim E = \dim \mathbb{R}^{N+1} = N + 1$.
2. Considérons la famille des fonctions $f_k \in E$ pour $k \in [\![0, N]\!]$ définies par

$$f_k : \left\{ \begin{array}{ccc} [a_0, a_N] & \to & \mathbb{R} \\ x & \mapsto & |x - a_k| \end{array} \right. .$$

Pour montrer qu'il s'agit d'une base, il suffit de montrer la liberté puisque son cardinal est $N + 1 = \dim E$. Soit $\lambda_0, \ldots, \lambda_N \in \mathbb{K}$ tels que $\sum_{j=0}^{N} \lambda_j f_j = 0$. Alors, pour tout $k \in [\![1, N-1]\!]$,

$$\lambda_k f_k = -\sum_{\substack{j=0 \\ j \neq k}}^{N} \lambda_j f_j$$

est dérivable en a_k (avec l'expression de droite) donc $\lambda_k = 0$. L'égalité se réécrit alors $\lambda_0 f_0 = -\lambda_N f_N$ ce qui entraîne immédiatement $\lambda_0 = \lambda_N = 0$.

Exercice 4

Soit F un tel sous-espace. Raisonnons par l'absurde et supposons que $\dim F \geq 2$: il existe donc deux fonctions f et $g \in F$ non proportionnelles et positives (quitte à les remplacer par leurs opposés). Soit $a \in \mathbb{R}$ tel que $g(a) > 0$. Comme f est différente de $\frac{f(a)}{g(a)}g$, il existe $b \in \mathbb{R}$ tel que $f(b) \neq \frac{f(a)}{g(a)}g(b)$. Considérons alors λ strictement entre $\frac{f(a)}{g(a)}$ et $\frac{f(b)}{g(b)}$ (si $g(b) = 0$, on remplace cette dernière borne par $+\infty$) et posons $\varphi = f - \lambda g \in F$.
Alors, φ est de signe constant par hypothèse alors que $\varphi(a)$ et $\varphi(b)$ sont de signes différents (par définition de λ) : contradiction.

Exercice 5

1. ▹ La partie $\mathscr{S} \subset \mathbb{R}^n$ est non vide car elle contient la suite nulle. Soit u, $v \in \mathscr{S}$, $\lambda \in \mathbb{R}$. Alors, pour tout $n \in \mathbb{N}$,

$$(\lambda u_{n+3} + v_{n+3}) - (a+b+c)(\lambda u_{n+2} + v_{n+2}) + (ab+bc+ac)(\lambda u_{n+1} + v_{n+1}) - abc(\lambda u_n + v_n) = 0,$$

en regroupant les termes en u et les termes en v : $\lambda u + v \in \mathscr{S}$.
Ainsi, $\mathscr{S}$ est non vide, stable par combinaison linéaire donc est un sous-espace vectoriel de $\mathbb{R}^{\mathbb{N}}$.
▹ Une suite géométrique non nulle de raison r appartient à $\mathscr{S}$ si, et seulement si, r est solution de l'équation $r^3 - (a+b+c)r^2 + (ab+bc+ac)r - abc = 0$ qui, d'après les relations entre coefficients et racines, se réécrit $(r-a)(r-b)(r-c) = 0$. En conclusion, les suites géométriques non nulles de $\mathscr{S}$ sont les suites géométriques de raison dans $\{a, b, c\}$.
2. L'application φ est clairement linéaire. Elle est une bijection car une fois fixés les trois premiers termes, une suite de $\mathscr{S}$ s'obtient de manière unique à partir de la relation de récurrence.
3. Les vérifications des résultats suivants sont immédiates comme dans le cas d'une relation de récurrence linéaire d'ordre 2 à coefficients constants.
▹ Si a, b et c sont distincts les trois suites géométriques $(a^n)_n$, $(b^n)_n$ et $(c^n)_n$ forment une famille libre et $\dim \mathscr{S} = 3$ donc elles donnent une base de $\mathscr{S}$.
▹ Si $a = b \neq c$, alors les trois suites $(a^n)_n$, $(na^n)_n$ et $(c^n)_n$ forment une famille libre de $\mathscr{S}$ et $\dim \mathscr{S} = 3$ donc elles donnent une base de $\mathscr{S}$.
▹ Si $a = b = c$, alors les trois suites $(a^n)_n$, $(na^n)_n$ et $(n^2a^n)_n$ forment une famille libre de $\mathscr{S}$ et $\dim \mathscr{S} = 3$ donc elles donnent une base de $\mathscr{S}$.

Exercice 6

1. Par construction, il est évident que $\mathscr{T}_n$ est l'espace engendré par les fonctions $x \mapsto 1$, $x \mapsto \cos(kx)$ et $x \mapsto \sin(kx)$ avec $k \in [\![1, n]\!]$, c'est donc un espace vectoriel.
2. On calcule ces intégrales par linéarité puis en linéarisant les produits de fonctions trigonométriques.

Or,

$$\begin{aligned}
\int_{-\pi}^{\pi} \sin(kx)\sin(jx)\,\mathrm{d}x &= \frac{1}{2}\int_{-\pi}^{\pi} \cos((k-j)x) - \cos((k+j)x)\,\mathrm{d}x \\
&= \begin{cases} 0 & \text{si } k \neq j, \\ 0 & \text{si } k = j = 0, \\ \pi & \text{sinon.} \end{cases} \\
\int_{-\pi}^{\pi} \sin(kx)\cos(jx)\,\mathrm{d}x &= \frac{1}{2}\int_{-\pi}^{\pi} \sin((k+j)x) + \sin((k-j)x)\,\mathrm{d}x \\
&= \begin{cases} 0 & \text{si } k \neq j, \\ 0 & \text{si } k = j = 0, \\ \pi & \text{sinon.} \end{cases} \\
\int_{-\pi}^{\pi} \cos(kx)\cos(jx)\,\mathrm{d}x &= \frac{1}{2}\int_{-\pi}^{\pi} \cos((k+j)x) + \cos((k-j)x)\,\mathrm{d}x \\
&= \begin{cases} 0 & \text{si } k \neq j, \\ 2\pi & \text{si } k = j = 0, \\ \pi & \text{sinon.} \end{cases}
\end{aligned}$$

Ainsi, en posant $T(x) = \frac{a_0}{2} + \sum_{k=1}^{n}(a_k\cos(kx) + b_k\sin(kx))$, on obtient

$$\int_{-\pi}^{\pi} \sin(kx)T(x)\,\mathrm{d}x = b_k \qquad \text{et} \qquad \int_{-\pi}^{\pi} \cos(kx)T(x)\,\mathrm{d}x = a_k.$$

3. On vient de voir qu'on retrouve les coefficients d'une combinaison linéaire T de ces fonctions en calculant des intégrales ; par conséquent, si $T = 0$, tous les coefficients sont nuls. La dimension de $\mathscr{T}_n$ est donc $2n+1$.
La deuxième question nous permet de dire que les formes linéaires coordonnées sont

$$x \mapsto \int_{-\pi}^{\pi} \cos(kx)T(x)\,\mathrm{d}x, \qquad x \mapsto \int_{-\pi}^{\pi} \sin(kx)T(x)\,\mathrm{d}x\,.$$

Exercice 7

L'hypothèse entraîne que $u \circ (u+v) = \mathrm{Id}$. Comme u, $u+v$ sont des endomorphismes d'un espace vectoriel de dimension finie, ceci entraîne que u est inversible et que $u^{-1} = u + v$ donc u commute avec $u + v$ et avec $(u+v) - u = v$.

Exercice 8

1. Montrons le résultat par récurrence sur n.
▹ Le résultat est évident pour $n = 0$ (et $n = 1$).
▹ Soit $n \in \mathbb{N}$. Supposons que, pour tout $v \in \mathscr{L}(E)$,

$$\Phi^n(v) = \sum_{k=0}^{n} (-1)^k \binom{n}{k} u^{n-k} \circ v \circ u^k.$$

Alors, pour tout $v \in \mathscr{L}(E)$,

$$\begin{aligned}
\Phi^{n+1}(v) &= \Phi(\Phi^n(v)) \\
&= \sum_{k=0}^{n}(-1)^k\binom{n}{k}u^{n-k+1}\circ v\circ u^k - \sum_{k=0}^{n}(-1)^k\binom{n}{k}u^{n-k}\circ v\circ u^{k+1} \\
&= \sum_{k=0}^{n}(-1)^k\binom{n}{k}u^{n-k+1}\circ v\circ u^k - \sum_{k=1}^{n+1}(-1)^{k-1}\binom{n}{k-1}u^{n-k+1}\circ v\circ u^k \\
&= u^{n+1}\circ v + \sum_{k=1}^{n}(-1)^k\left(\binom{n}{k}+\binom{n}{k-1}\right)u^{n-k+1}\circ v\circ u^k - (-1)^n v\circ u^{n+1} \\
&= \sum_{k=0}^{n+1}(-1)^k\binom{n+1}{k}u^{n+1-k}\circ v\circ u^k,
\end{aligned}$$

avec la formule de Pascal.

2. Avec $n = 2p-1$, $\max\{k, n-k\} \geq p$ pour tout $0 \leq k \leq n$. Par conséquent, $\Phi^n(v) = 0$ pour tout $v \in \mathscr{L}(E)$. L'indice de nilpotence de Φ est inférieur ou égal à $2p-1$.

3. Considérons $(e_1, \ldots, e_n)$ une base de E telle que $(e_{r+1}, \ldots, e_n)$ soit une base du noyau de a. Alors, $(a(e_1), \ldots, a(e_r))$ est une base de l'image de a que l'on complète en une base $(a(e_1), \ldots, a(e_r), e'_{r+1}, \ldots, e'_n)$ de E. Définissons un endomorphisme b par l'image de cette base :

- pour tout $k \in \{1, \ldots, r\}$, $b(a(e_k)) = e_k$;
- pour tout $k \in \{r+1, \ldots, n\}$, $b(e'_k) = 0_E$.

Pour tout $x \in E$, il existe des scalaires $\lambda_1, \ldots, \lambda_n$ tels que $a(x) = \sum_{k=1}^{r}\lambda_k a(e_k)$.

Alors,

$$\begin{aligned}
a\circ b\circ a(x) &= a\circ b\left(\sum_{k=1}^{r}\lambda_k a(e_k)\right) \\
&= a\left(\sum_{k=1}^{r}\lambda_k b\circ a(e_k)\right) \\
&= a\left(\sum_{k=1}^{r}\lambda_k e_k\right) = a(x).
\end{aligned}$$

En conclusion, $a\circ b\circ a = a$.

4. Nous avons déjà établi que $\Phi^{2p-1} = 0$; montrons que $\Phi^{2p-2} \neq 0$. Pour tout $v \in \mathscr{L}(E)$,

$$\begin{aligned}
\Phi^{2p-2}(v) &= \sum_{k=0}^{2p-2}(-1)^k\binom{2p-2}{k}u^{2p-2-k}\circ v\circ u^k \\
&= \sum_{k=0}^{p-2}(-1)^k\binom{2p-2}{k}u^{2p-2-k}\circ v\circ u^k + (-1)^{p-1}\binom{2p-2}{p-1}u^{p-1}\circ v\circ u^{p-1} \\
&\quad + \sum_{k=p}^{2p-2}(-1)^k\binom{2p-2}{k}u^{2p-2-k}\circ v\circ u^k \\
&= (-1)^{p-1}\binom{2p-2}{p-1}u^{p-1}\circ v\circ u^{p-1}.
\end{aligned}$$

En considérant v l'application associée à u^{p-1} d'après la question précédente, on obtient

$$\Phi^{2p-2}(v) = (-1)^{p-1}\binom{2p-2}{p-1}u^{p-1} \neq 0$$

CORRIGÉS

En conclusion, Φ est nilpotente d'indice $2p-1$.

Exercice 9

Soit r le rang de u et $(e_1,\dots,e_r)$ une base de $\operatorname{Im} u$. Complétons cette famille libre en une base $(e_1,\dots,e_n)$ de E (avec $n=\dim E$).
Commençons par remarquer que, pour tout $v\in\mathscr{L}(E,F)$, $v\circ u=0_{\mathscr{L}(E,F)}$ si, et seulement si, $\operatorname{Im} u\subset\operatorname{Ker} v$, c'est-à-dire si, pour tout $i\in[\![1,r]\!]$, $v(e_r)=0_F$. Ainsi, une telle application v est uniquement déterminée par l'image des vecteurs e_{r+1}, …, e_n. Autrement dit, l'application $v\mapsto(v(e_{r+1}),\dots,v(e_n))$ est un isomorphisme entre l'espace dont on cherche la dimension et F^{n-r}. La dimension recherchée est donc

$$\dim F^{n-r}=\dim F(\dim E-\operatorname{rg} u).$$

Exercice 10

▹ Comme $\operatorname{Im}(u+v)\subset\operatorname{Im} u+\operatorname{Im} v$, on obtient en passant à la dimension $\operatorname{rg}(u+v)\leq\operatorname{rg}(u)+\operatorname{rg}(v)$.
▹ Utilisons la première inégalité à deux reprises

$$\operatorname{rg}(u)=\operatorname{rg}(u+v-v)\leq\operatorname{rg}(u+v)+\operatorname{rg}(-v)=\operatorname{rg}(u+v)+\operatorname{rg}(v)$$
$$\operatorname{rg}(v)=\operatorname{rg}(u+v-u)\leq\operatorname{rg}(u+v)+\operatorname{rg}(-u)=\operatorname{rg}(u+v)+\operatorname{rg}(u)$$

Ainsi, $|\operatorname{rg}(u)-\operatorname{rg}(v)|\leq\operatorname{rg}(u+v)$.

Exercice 11

1. Soit $\tilde{v}$ l'application induite par v entre $\operatorname{Im}(u)$ et E. Par construction, $\operatorname{Ker}\tilde{v}=\operatorname{Ker} v\cap\operatorname{Im} u$ et $\operatorname{Im}\tilde{v}=v(\operatorname{Im} u)=\operatorname{Im} v\circ u$.
2. En appliquant la formule du rang à $\tilde{v}$, on obtient exactement la formule recherchée.
3. ▹ On écrit encore la formule du rang : $\operatorname{rg} u=\dim E-\dim\operatorname{Ker} u$ et $\operatorname{rg} v\circ u=\dim E-\dim\operatorname{Ker} v\circ u$; l'inégalité précédente se réécrit

$$\dim\operatorname{Ker}(v\circ u)\leq\dim\operatorname{Ker} u+\dim\operatorname{Ker}(v)\cap\operatorname{Im}(u)\leq\dim\operatorname{Ker} u+\dim\operatorname{Ker}(v).$$

▹ en posant $v=u^k$, on obtient $\dim\operatorname{Ker} u^{k+1}\leq\dim\operatorname{Ker} u^k+\dim\operatorname{Ker} u$ d'où le premier résultat par récurrence. La seconde inégalité s'obtient alors avec la formule du rang pour u et u^k.

Exercice 12

1. La condition $v\circ u=0_{\mathscr{L}(E)}$ équivaut $\operatorname{Im} u\subset\operatorname{Ker} v$ ce qui entraîne en passant à la dimension

$$\operatorname{rg} u\leq\dim E-\operatorname{rg} v.$$

Par ailleurs, $u+v$ est inversible donc

$$n=\operatorname{rg}(u+v)\leq\operatorname{rg} u+\operatorname{rg} v.$$

En combinant ces deux inégalités, on obtient $\operatorname{rg} u+\operatorname{rg} v=\dim E$.
2. Si un tel projecteur existe, on a, d'après la première égalité, $\operatorname{Im} p=\operatorname{Im} u$ (car $u+v$ inversible) et, d'après la deuxième égalité, $\operatorname{Ker} p=\operatorname{Im} v$.
▹ Remarquons que $\operatorname{Im} u$ et $\operatorname{Im} v$ sont en somme directe car un vecteur x dans leur intersection vérifie $(u+v)(x)=0_E+0_E=0_E$ donc $x=0_E$ car $u+v$ est inversible.
▹ Comme $\operatorname{rg} u+\operatorname{rg} v=\dim E$, la condition de somme directe entraîne le caractère supplémentaire de ces deux sous-espaces. Posons p le projecteur sur $\operatorname{Im} u$ parallèlement à $\operatorname{Im} v$. Par construction, $p\circ u=u$ et $p\circ v=0_{\mathscr{L}(E)}$ donc en développant

$$u=p\circ(u+v),\qquad v=(\operatorname{Id}-p)\circ(u+v).$$

Exercice 13

1. ($\Rightarrow$) La suite $(\operatorname{Ker} u^k)_k$ est strictement croissante pour $k \in [\![0, n]\!]$ et $\dim \operatorname{Ker} u^n = \dim E = n$. Par conséquent, pour tout $k \in [\![0, n]\!]$, on a $\dim \operatorname{Ker} u^k = k$.
($\Leftarrow$) La suite $(\dim \operatorname{Ker} u^{k+1} - \dim \operatorname{Ker} u^k)_k$ est décroissante de premier terme 1. Ainsi, si un noyau itéré $\operatorname{Ker} u^k$ est de dimension n, alors $k \geq n$. Comme un indice de nilpotence est inférieur ou égal à n, l'indice de nilpotence de u est n.
2. ▹ Remarquons tout d'abord que $\dim \operatorname{Ker} u \geq 2$ d'après la première question donc $\dim \operatorname{Im} u \leq n-2$.
▹ L'endomorphisme induit $u_{\operatorname{Im} u}$ est nilpotent d'indice $n-2$, car u est d'indice de nilpotence $n-1$. Ainsi, $\dim \operatorname{Im} u = n-2$ et d'après la question précédente, $\dim \operatorname{Ker} u_{\operatorname{Im} u} = 1$. Or, $\operatorname{Ker} u_{\operatorname{Im} u} = \operatorname{Ker} u \cap \operatorname{Im} u$ d'où $\dim \operatorname{Ker} u \cap \operatorname{Im} u = 1$.

Exercice 14

Le sens retour est évident : si $\operatorname{Im} u \subset F$, alors $u(F) \subset \operatorname{Im} u \subset F$; si $F \subset \operatorname{Ker} u$, alors $u(F) = \{0_E\} \subset F$.
Pour le sens direct, considérons F stable par u tel que $F \not\subset \operatorname{Ker} u$. Il existe donc $x \in F$ tel que $u(x) \neq 0_E$. Mais comme u est de rang 1, $\operatorname{Im} u = \operatorname{Vect}(u(x)) \subset u(F) \subset F$.

Exercice 15

▹ Si un sous-espace F contient $\operatorname{Im} u$, alors il contient $u(F)$: par conséquent, F est stable.
▹ Soit F un plan de $\mathbb{R}^3$ stable par u. Remarquons que la condition $u^2 = 0_{\mathscr{L}(\mathbb{R}^3)}$ entraîne $\operatorname{Im} u \subset \operatorname{Ker} u$ et donc, grâce à la formule du rang $\dim \operatorname{Im} u = 1$ et $\dim \operatorname{Ker} u = 2$.
- Si $F = \operatorname{Ker} u$, alors $\operatorname{Im} u \subset \operatorname{Ker} u = F$.
- Sinon, il existe $x \in F$ tel que $u(x) \neq 0_{\mathbb{R}^3}$. Comme $\operatorname{Im} u$ est une droite, $\operatorname{Im} u = \operatorname{Vect} u(x) \subset u(F) \subset F$.

Dans les deux cas, F contient l'image de u.

Exercice 16

Montrons par récurrence sur $k \in \mathbb{N} \setminus \{0\}$ que $E = F + \operatorname{Im} u^k$.
- L'initialisation pour $k = 1$ est l'hypothèse de l'énoncé.
- Soit $k \in \mathbb{N} \setminus \{0\}$ tel que la propriété soit vraie au rang k. Soit $x \in E$. D'après l'hypothèse de récurrence, il existe $y \in F$ et $z \in E$ tels que $x = y + u^k(z)$. Par ailleurs, il existe $a \in F$ et $b \in E$ tels que $z = a + u(b)$. En combinant ces deux résultats, on a $x = y + u^k(a) + u^{k+1}(b) \in F + \operatorname{Im} u^{k+1}$. En conclusion, $E = F + \operatorname{Im} u^{k+1}$.

La propriété au rang n est $E = F + \operatorname{Ker} u^n = F$.

Exercice 17

1. Soit $k \in \mathbb{N} \setminus \{0\}$. Alors,

$$P_k'(X) = \frac{1}{k!}(X-k)^{k-1} + \frac{k-1}{k!}X(X-k)^{k-2} = \frac{1}{k!}(X-k)^{k-2}(X-k+(k-1)X) = P_{k-1}(X-1).$$

2. La famille considérée est échelonnée en degrés donc libre et compte $n+1 = \dim \mathbb{R}_n[X]$ éléments donc est une base de $\mathbb{R}_n[X]$.
Pour montrer que la famille de formes linéaires $(\varphi_0, \ldots, \varphi_n)$ est la base duale de cette base, il suffit de montrer que pour tous j, k, $\varphi_k(P_j) = \delta_{j,k}$. Or, d'après la première question, $P_k^{(j)}(X) = P_{k-j}(X-j)$ pour $k \geq j$ et $P_k^{(j)}(X) = 0$ sinon. Ainsi,

$$\varphi_k(P_j) = \begin{cases} 0 & \text{si } k > j \\ P_0(0) = 1 & \text{si } k = j \\ P_{j-k}(j-k) = 0 & \text{si } k < j \end{cases}$$

Le résultat est ainsi établi.

Exercice 18

1. La famille est échelonnée en degrés (pour tout $k \in [\![0,n]\!]$, $\deg Q^{(k)} = n-k$) et de cardinal $n+1 = \dim \mathbb{R}_n[X]$ donc est une base de $\mathbb{R}_n[X]$.
2. D'après la formule de Taylor pour les polynômes, la famille des formes linéaires $P \mapsto P^{(k)}(0)$ avec $k \in [\![0,n]\!]$ est la base duale de la base des $\frac{1}{k!}X^k$ avec $k \in [\![0,n]\!]$, c'est donc une base de l'espace dual de $\mathbb{R}_n[X]$. Par conséquent, pour toute forme linéaire φ sur $\mathbb{R}_n[X]$, il existe des scalaires $\alpha_0,\dots,\alpha_n$ tels que, pour tout $P \in \mathbb{R}_n[X]$,
$$\varphi(P) = \alpha_0 P(0) + \alpha_1 P'(0) + \cdots + \alpha_n P^{(n)}(0).$$
3. Soit $R = \alpha_0 Q + \cdots + \alpha_n Q^{(n)}$ avec $\alpha_0,\dots,\alpha_n$ les scalaires associés à φ dans la question précédente. La condition $\varphi(Q_k) = 0$ signifie que $R(k) = 0$. Par conséquent, R est un polynôme de degré au plus n qui admet $n+1$ racines il est donc nul. Comme la famille $(Q, Q', Q'', \dots, Q^{(n)})$ est libre, tous les coefficients α_k pour $k \in [\![0,n]\!]$ sont nuls et donc φ est la forme linéaire nulle.
4. Si la famille $(Q_0,\dots,Q_n)$ n'est pas une base de $\mathbb{R}_n[X]$, elle engendre un sous-espace vectoriel $F \neq \mathbb{R}_n[X]$. On peut alors définir une forme linéaire nulle sur F et non nulle sur un supplémentaire de F, ce qui contredit la question précédente. Par conséquent, $(Q_0,\dots,Q_n)$ est une base de $\mathbb{R}_n[X]$.

Exercice 19

1. Soit $x_0 \in E$ tel que $u^{n-1}(x_0) \neq 0_E$.
▹ Montrons par l'absurde que la famille $(x_0, u(x_0),\dots,u^{n-1}(x_0))$ est libre. Soit $(\lambda_0,\dots,\lambda_{n-1}) \in \mathbb{K}^n$ non nul tel que
$$0_{\mathscr{L}(E)} = \sum_{k=0}^{n-1} \lambda_k u^k(x_0).$$
Notons p le plus petit entier k tel que $\lambda_k \neq 0$ et composons l'égalité précédente par u^{n-1-p}
$$0_{\mathscr{L}(E)} = u^{n-1-p} \circ \sum_{k=p}^{n-1} \lambda_k u^k(x_0) = \sum_{k=p}^{n-1} \lambda_k u^{n-1-p+k}(x_0) = \lambda_p u^{n-1}(x_0)$$
donc $\lambda_p = 0$ car $u^{n-1}(x_0) \neq 0_E$: contradiction avec la définition de p.
▹ La famille $(x_0, u(x_0),\dots,u^{n-1}(x_0))$ est libre et de cardinal $n = \dim E$ donc est une base de E : u est cyclique.
2. Soit $x_0 \in E$ tel que $(x_0, u(x_0),\dots,u^{n-1}(x_0))$ est une base de E. Écrivons le vecteur $u^n(x_0)$ dans cette base : il existe des scalaires $(a_k)_{k\in[0,n-1]}$ tels que
$$u^n(x_0) = -\sum_{k=0}^{n-1} a_k u^k(x_0).$$
Pour tout $j \in [\![0,n-1]\!]$,
$$\begin{aligned} u^n(u^j(x_0)) = u^{n+j}(x_0) = u^j(u^n(x_0)) &= -u^j \circ \sum_{k=0}^{n-1} a_k u^k(x_0) \\ &= -\sum_{k=0}^{n-1} a_k u^{k+j}(x_0) = -\sum_{k=0}^{n-1} a_k u^k(u^j(x_0)). \end{aligned}$$
Alors, les endomorphismes u^n et $-\sum_{k=0}^{n-1} a_k u^k$ coïncident sur une base donc sont égaux
$$u^n = -\sum_{k=0}^{n-1} a_k u^k.$$

Corrigés des exercices d'approfondissement du chapitre 15

Exercice A

▹ Montrons, par récurrence sur $k \leq N$, que $\mathrm{Vect}(x_i)_{i\leq k} = \mathrm{Vect}(x_i)_{i\in I, i\leq k}$.

- L'égalité est vraie au rang 1.
- Soit $k < N$ tel que l'égalité est vraie au rang k.

Si $k+1 \in I$, alors

$$\mathrm{Vect}(x_i)_{i\leq k+1} = \mathrm{Vect}(x_i)_{i\leq k} + \mathrm{Vect}(x_{k+1}) = \mathrm{Vect}(x_i)_{i\in I, i\leq k} + \mathrm{Vect}(x_{k+1}) = \mathrm{Vect}(x_i)_{i\in I, i\leq k+1}.$$

Sinon, $\mathrm{Vect}(x_i)_{i\leq k+1} = \mathrm{Vect}(x_i)_{i\leq k}$ car $x_{k+1} \notin \mathrm{Vect}(x_1, \ldots, x_k)$ et $\mathrm{Vect}(x_i)_{i\in I, i\leq k+1} = \mathrm{Vect}(x_i)_{i\in I, i\leq k}$ car $k+1 \notin I$ donc

$$\mathrm{Vect}(x_i)_{i\leq k+1} = \mathrm{Vect}(x_i)_{i\in I, i\leq k+1}.$$

Par conséquent, en appliquant cette égalité au rang N

$$E = \mathrm{Vect}(x_i)_{i\leq N} = \mathrm{Vect}(x_i)_{i\in I}.$$

▹ Supposons par l'absurde que la famille $(x_i)_{i\in I}$ est liée : il existe une famille $(\lambda_i)_{i\in I}$ de scalaires non tous nuls tels que

$$\sum_{i\in I} \lambda_i x_i = 0_E.$$

Soit k le plus grand élément de I tel que $\lambda_k \neq 0$. Alors,

$$x_k = -\sum_{\substack{i\in I \\ i<k}} \frac{\lambda_i}{\lambda_k} x_i \in \mathrm{Vect}(x_i)_{i<k}$$

ce qui contredit l'appartenance de k à I.

Exercice B

Par définition de l'espace vectoriel engendré, pour $j \leq k$,

$$\mathrm{Vect}(x_1, \ldots, x_k) = \mathrm{Vect}(x_1, \ldots, x_j) + \mathrm{Vect}(x_{j+1}, \ldots, x_k)$$

donc en passant à la dimension

$$\mathrm{rg}(x_1, \ldots, x_k) \leq \mathrm{rg}(x_1, \ldots, x_j) + \mathrm{rg}(x_{j+1}, \ldots, x_k) \leq \mathrm{rg}(x_1, \ldots, x_j) + k - j,$$

ce qui est bien le résultat annoncé.

Exercice C

Supposons que $\mathrm{rg}(f'_1, \ldots, f'_n) \leq n-2$ et, sans perte de généralité que $(f'_1, \ldots, f'_{n-2})$ est génératrice de

$$\mathrm{Vect}(f'_1, \ldots, f'_n)$$

. Alors, il existe $\lambda_1, \ldots, \lambda_{n-2}, \mu_1, \ldots, \mu_{n-2} \in \mathbb{R}$ tels que

$$f'_{n-1} = \sum_{k=1}^{n-2} \lambda_k f'_k, \qquad f'_n = \sum_{k=1}^{n-2} \mu_k f'_k.$$

Il existe donc des constantes A, $B \in \mathbb{R}$ telles que

$$f_{n-1} - \sum_{k=1}^{n-2} \lambda_k f_k = A, \qquad f_n - \sum_{k=1}^{n-2} \mu_k f_k = B.$$

Comme la famille $(f_1,\dots,f_n)$ est libre, les constantes A et B ne sont pas nulles (sinon on aurait une combinaison linéaire nulle à coefficients non tous nuls) ; alors

$$B\Big(f_{n-1}-\sum_{k=1}^{n-2}\lambda_k f_k\Big)-A\Big(f_n-\sum_{k=1}^{n-2}\mu_k f_k\Big)=0$$

fournit une combinaison linéaire nulle à coefficients non tous nuls : contradiction.

Exercice D

Considérons l'application linéaire

$$u:\left\{\begin{array}{ccc} V_1\times\ldots\times V_p & \to & E^{p-1}\\ (x_1,\dots,x_p) & \mapsto & (x_1-x_2,\dots,x_{p-1}-x_p)\end{array}\right.$$

D'après l'hypothèse, $\dim V_1\times\ldots\times V_p>\dim E^{p-1}$ donc u n'est pas injective. Il existe donc un $(x_1,\dots,x_p)\in$ $\operatorname{Ker} u$ non nul. Par définition de u, $x_1=x_2=\ldots=x_p$ donc x_1 appartient à $\bigcap_{k=1}^{p} V_k$ et est non nul.

Exercice E

Notons e_1, e_2, …, e_n la base canonique de $\mathbb{K}^n$ et u un endomorphisme de $\mathbb{K}^n$ qui vérifie les conditions de l'énoncé. Par définition, il existe des scalaires λ_1, …, λ_n, $\lambda\in\mathbb{K}$ tels que $u(e_k)=\lambda_k e_k$ pour tout $k\in[\![1,n]\!]$ et $u(e_1+\ldots+e_n)=\lambda(e_1+\ldots+e_n)$. Par linéarité, on a

$$\lambda_1 e_1+\ldots+\lambda_n e_n=\lambda(e_1+\ldots+e_n)$$

et donc $\lambda_1=\ldots=\lambda_n=\lambda$. En conclusion, u est une homothétie. Réciproquement, il est évident que les homothéties vérifient bien les conditions de l'énoncé.

Exercice F

1. ▹ Le sens direct est évident.
▹ Pour le sens retour, considérons l'application linéaire

$$\Phi:\left\{\begin{array}{ccc} E & \to & \mathbb{K}^{p+1}\\ x & \mapsto & (\varphi(x),\varphi_1(x),\dots,\varphi_p(x)).\end{array}\right.$$

Le vecteur $(1,0,\dots,0)$ n'appartient pas à l'image de Φ : en effet, s'il existe $x\in E$ tel que $\Phi(x)=(1,0,\dots,0)$ alors $x\in\cap_{i=1}^{p}\operatorname{Ker}\varphi_i$ et $x\notin\operatorname{Ker}\varphi$ ce qui entre en conflit avec notre hypothèse sur les noyaux. Ainsi, $\operatorname{Im}\Phi$ est un sous-espace strict de $\mathbb{K}^{p+1}$ donc inclus dans un hyperplan H de $\mathbb{K}^{p+1}$; par ailleurs, il existe des scalaires $a_0,\dots,a_p\in\mathbb{K}$ non tous nuls tels que $(x_0,x_1,\dots,x_p)\in H$ si, et seulement si,

$$\sum_{k=0}^{p}a_k x_k=0.$$

Comme $\operatorname{Im}\Phi\subset H$, on a, pour tout $x\in E$,

$$a_0\varphi(x)+\sum_{k=1}^{p}a_k\varphi_k(x)=0.$$

Par conséquent, $a_0\varphi+\sum_{k=1}^{p}a_k\varphi_k$ est la forme linéaire nulle : $(\varphi_1,\dots,\varphi_p)$ est libre et $(\varphi,\varphi_1,\dots,\varphi_p)$ est liée. En conclusion, $\varphi\in\operatorname{Vect}(\varphi_1,\dots,\varphi_p)$.
2. Pour tout $\varphi\in E^*$,

$$\bigcap_{i=1}^{p}\operatorname{Ker}\varphi_i=\{0_E\}\subset\operatorname{Ker}\varphi.$$

D'après la question précédente, $\varphi\in\operatorname{Vect}(\varphi_1,\dots,\varphi_p)$ donc $\operatorname{Vect}(\varphi_1,\dots,\varphi_p)=E^*$.

Corrigés des problèmes du chapitre 15

Problème 1

1. a. Comme $F \neq \{0_E\}$, il existe une fonction non nulle $f \in F$. Puisque f est non nulle, il existe $x \in \mathbb{R}$ tel que $f(x) \neq 0$, donc $\varphi_x(f) \neq 0$ et φ_x est non nulle.
b. i. Remarquons tout d'abord que $k \leq n$ car le cardinal d'une famille libre est inférieur à

$$\dim F^* = \dim F = n.$$

Pour tout $x \in \mathbb{R}$, la famille $(\varphi_x, \varphi_{x_1}, \ldots, \varphi_{x_k})$ est liée et la famille $(\varphi_{x_1}, \ldots, \varphi_{x_k})$ est libre donc

$$\varphi_x \in \mathrm{Vect}(\varphi_{x_1}, \ldots, \varphi_{x_k}).$$

Ainsi, il existe des scalaires $c_1(x), \ldots, c_k(x)$ tels que

$$\varphi_x = \sum_{j=1}^{k} c_j(x)\varphi_{x_j}.$$

Ainsi, pour tout $f \in F$,

$$f(x) = \sum_{j=1}^{k} c_j(x) f(x_j)$$

soit $F \in \mathrm{Vect}(c_1, \ldots, c_n)$ est de dimension au plus k : $n \leq k$.
ii. Soit $x_1, \ldots, x_n$ la famille de la question précédente. La famille $(\varphi_{x_1}, \ldots, \varphi_{x_n})$ est libre et son cardinal est $n = \dim F^*$ donc est une base de F^*.
c. L'application $\left\{\begin{array}{ccc} F & \to & \mathbb{C}^n \\ f & \mapsto & (\varphi_{x_1}(f), \ldots, \varphi_{x_n}(f)) \end{array}\right.$ est clairement linéaire, injective (avec les notations précédentes $f(x) = \sum_{j=1}^{k} c_j(x) f(x_j)$) donc un isomorphisme. Considérons la famille des fonctions $v_1, \ldots, v_n$ de F dont l'image par cet isomorphisme est la base canonique de $\mathbb{C}^n$. C'est une base comme image d'une base par un isomorphisme (c'est la base antéduale de $(\varphi_{x_1}, \ldots, \varphi_{x_n})$).
Pour tout $f \in F$, $f = \sum_{j=1}^{n} f(x_j) v_j$.
2. a. Considérons les réels $(x_1, \ldots, x_n)$ et la base $(v_1, \ldots, v_n)$ de F définis à la question 1. Ainsi

$$\forall g \in F,\ \forall a, x \in \mathbb{R}, \qquad \tau_a(g)(x) = \sum_{k=1}^{n} \tau_a(g)(x_k) v_k(x).$$

b. Pour tout $g \in F$ et tous $a, x \in \mathbb{R}$,

$$\frac{\tau_a(g)(x) - g(x)}{a} = \sum_{k=1}^{n} \frac{\tau_a(g)(x_k) - g(x_k)}{a} v_k(x).$$

En passant à la limite $a \to 0$,

$$g'(x) = \sum_{k=1}^{n} g'(x_k) v_k(x).$$

En conclusion, $g' \in F$.
c. D'après la question précédente, toutes les dérivées d'une fonction de F appartiennent à F. La famille $(f, f', \ldots, f^{(n)})$ est de cardinal $n+1$ dans l'espace F de dimension n donc est liée. Ainsi, il existe des constantes $a_0, \ldots a_n$ telles que $\sum_{k=0}^{n} a_k f^{(k)} = 0$.
Toutes les translatées de f vérifient cette équation donc toutes les fonctions de F sont solutions de cette équation.

Problème 2

Partie A – Décomposition de Fitting

1. Soit $k \in \mathbb{N}$.

$$\begin{aligned} x \in \operatorname{Ker} f^k &\Leftrightarrow f^k(x) = 0_E \\ &\Rightarrow f^{k+1}(x) = 0_E \\ &\Leftrightarrow x \in \operatorname{Ker} f^{k+1}. \end{aligned}$$

D'où $\operatorname{Ker} f^k \subset \operatorname{Ker} f^{k+1}$.

$$\begin{aligned} x \in \operatorname{Im} f^{k+1} &\Leftrightarrow \exists t \in E,\ x = f^{k+1}(t) \\ &\Rightarrow \exists y \in E,\ x = f(y) \\ &\Leftrightarrow x \in \operatorname{Im} f^k. \end{aligned}$$

D'où $\operatorname{Im} f^{k+1} \subset \operatorname{Im} f^k$.
En conclusion, la suite $(\operatorname{Ker} f^k)_k$ est croissante pour l'inclusion et que la suite $(\operatorname{Im} f^k)_k$ est décroissante pour l'inclusion.
2. a. La suite $(\dim \operatorname{Ker} f^k)_k$ est croissante majorée par $\dim E$ donc est stationnaire. Soit

$$r = \min\{k \in \mathbb{N},\ \dim \operatorname{Ker} f^k = \dim \operatorname{Ker} f^{k+1}\}$$

et $p > r$. Alors

$$\begin{aligned} x \in \operatorname{Ker} f^p &\Leftrightarrow f^p(x) = 0_E \\ &\Leftrightarrow f^{r+1}(f^{p-r-1}(x)) = 0_E \\ &\Leftrightarrow f^r(f^{p-r-1}(x)) = 0_E \\ &\Leftrightarrow f^{p-1}(x) = 0_E. \end{aligned}$$

Ainsi, la suite est strictement croissante jusqu'au rang r puis constante après ce rang.
b. Comme on a déjà une inclusion entre les images, on obtient

$$\begin{aligned} \operatorname{Im} f^k = \operatorname{Im} f^{k+1} &\Leftrightarrow \dim \operatorname{Im} f^k = \dim \operatorname{Im} f^{k+1} \\ &\Leftrightarrow \dim \operatorname{Ker} f^k = \dim \operatorname{Ker} f^{k+1} && \text{par le théorème du rang} \\ &\Leftrightarrow \operatorname{Ker} f^k = \operatorname{Ker} f^{k+1} && \text{d'après l'inclusion de la question 1.} \end{aligned}$$

Ainsi, la suite $(\operatorname{Im} f^k)_k$ est stationnaire à partir du même rang que $(\operatorname{Ker} f^k)_k$.
c. Soit p un projecteur. Par définition, $p^2 = p$ donc $\operatorname{Ker} p = \operatorname{Ker} p^2$ et ainsi l'indice recherché r est inférieur ou égal à 1. Si $r = 0$, on a $\{0_E\} = \operatorname{Ker} p^0 = \operatorname{Ker} p$ donc p est inversible et $p = \mathrm{Id}_E$. En conclusion, si p est l'identité, alors $r = 0$; sinon, $r = 1$.
3. D'après ce qui précède et en conservant la notation r, $\mathscr{N}_f = \operatorname{Ker} f^r$ et $\mathscr{I}_f = \operatorname{Im} f^r$. Ces deux ensembles sont bien des sous-espaces (comme noyau et image) stables par f (car f et f^r commutent).
4. ▹ D'après la formule du rang appliquée à f^r, $\dim \mathscr{N}_f + \dim \mathscr{I}_f = \dim E$.
▹ Soit $x \in \mathscr{N}_f \cap \mathscr{I}_f$; alors $f^r(x) = 0_E$ et il existe $t \in E$ tel que $x = f^r(t)$. Ainsi, $f^{2r}(t) = 0_E$, c'est-à-dire $t \in \operatorname{Ker} f^{2r}$. Par définition de r, $t \in \operatorname{Ker} f^r$ donc $x = 0_E$.
Ainsi, $\mathscr{N}_f$ et $\mathscr{I}_f$ sont en somme directe et les dimensions concordent donc $E = \mathscr{N}_f \oplus \mathscr{I}_f$.
5. Pour tout $x \in \mathscr{N}_f = \operatorname{Im} f^r$, $(f_{\mathscr{N}})^r(x) = f^r(x) = 0_E$. D'où $f_{\mathscr{N}}$ est nilpotent d'indice au plus r (en fait, c'est exactement r).
6. L'application $f_{\mathscr{I}}$ est inversible car $\operatorname{Ker} f_{\mathscr{I}} = \operatorname{Ker} f \cap \mathscr{I}_f \subset \mathscr{N}_f \cap \mathscr{I}_f = \{0_E\}$.

Partie B – Cas général

1. a. ▹ Supposons qu'il existe une telle application linéaire f. Alors, pour tout $x \in E$ d'écriture $x_1 + x_2$ avec $x_1 \in E_1$ et $x_2 \in E_2$,

$$f(x) = f(x_1) + f(x_2) = f_1(x_1) + f_2(x_2).$$

Ainsi, l'application f, si elle existe, est unique.
▹ Réciproquement, notons p_1 et p_2 les projections sur E_1 parallèlement à E_2 et sur E_2 parallèlement à E_1. L'application $f = f_1 \circ p_1 + f_2 \circ p_2$ convient.
b. Par hypothèse, il existe des applications linéaires $a_1 \in \mathscr{L}(\mathscr{N}_u, \mathscr{N}_v)$, $b_1 \in \mathscr{L}(\mathscr{N}_v, \mathscr{N}_u)$, $a_2 \in \mathscr{L}(\mathscr{I}_u, \mathscr{I}_v)$ et $b_2 \in \mathscr{L}(\mathscr{I}_v, \mathscr{I}_u)$ telles que

$$u_{\mathscr{N}} = b_1 \circ a_1,\ u_{\mathscr{I}} = b_2 \circ a_2,\ v_{\mathscr{N}} = a_1 \circ b_1,\ v_{\mathscr{I}} = a_2 \circ b_2.$$

Soit $a \in \mathscr{L}(E, F)$ et $b \in \mathscr{L}(F, E)$ les applications linéaires associées par la question 1 à a_1 et a_2 et b_1 et b_2 respectivement.
Alors, pour tout $x \in E = \mathscr{N}_u \oplus \mathscr{I}_u$, il existe $x_1 \in \mathscr{N}_u$ et $x_2 \in \mathscr{I}_u$ tel que $x = x_1 + x_2$. Ainsi,

$$\begin{aligned} b \circ a(x) &= b(a_1(x_1) + a_2(x_2)) = b_1(a_1(x_1)) + b_2(a_2(x_2)) \\ &= u_{\mathscr{N}}(x_1) + u_{\mathscr{I}}(x_2) = u(x). \end{aligned}$$

Par conséquent, $u = b \circ a$ et on montre de même que $v = a \circ b$.
2. a. Soit $k \in_N$ et $x \in \operatorname{Ker}(b \circ a)^k$. Montrons que $a(x) \in \operatorname{Ker}(a \circ b)^k$.

$$(a \circ b)^k(a(x)) = (a \circ b)^k \circ a(x) = a \circ (b \circ a)^k(x) = a(0_E) = 0_E.$$

D'où $a(\operatorname{Ker}(b \circ a)^k) \subset \operatorname{Ker}(a \circ b)^k$. Ce résultat se réécrit encore $a(\operatorname{Ker} u^k) \subset \operatorname{Ker} v^k \subset \mathscr{N}_v$. Comme ceci est valable pour tout $k \in \mathbb{N}$, $a(\mathscr{N}_u) \subset \mathscr{N}_v$.
b. Procédons de la même manière. Soit $k \in_N$ et $y \in \operatorname{Ker}(b \circ a)^k$; il existe $x \in E$ tel que $y = (b \circ a)^k(x)$. Alors, $a(y) = (a \circ b)^k(a(x)) \in \operatorname{Im}(a \circ b)^k$. Ainsi, $a(\operatorname{Im} u^k) \subset \operatorname{Im} v^k$ et donc $a(\mathscr{I}_u) \subset \mathscr{I}_v$.
c. $b(\mathscr{N}_v) \subset \mathscr{N}_u$ et $b(\mathscr{I}_v) \subset \mathscr{I}_u$.
d. Définissons $a_1 : \left\{ \begin{array}{ccc} \mathscr{N}_u & \to & \mathscr{N}_v \\ x & \mapsto & a(x) \end{array} \right.$ et $b_1 : \left\{ \begin{array}{ccc} \mathscr{N}_v & \to & \mathscr{N}_u \\ x & \mapsto & b(x). \end{array} \right.$ Alors, on vérifie immédiatement que $u_{\mathscr{N}} = b_1 \circ a_1$ et $v_{\mathscr{N}} = a_1 \circ b_1$. On procède de même avec $u_{\mathscr{I}}$ et $v_{\mathscr{I}}$.

Partie C – Cas des isomorphismes

1. Il suffit de prendre $a = f$ et $b = f^{-1} \circ v$ et on obtient le résultat.
2. a. a est injective car $b \circ a = u$ est injective ; a est surjective car $a \circ b = v$ est surjective. Par conséquent, a est bijective donc un isomorphisme.
b. Cette fois-ci (et conformément à la première question de cette partie), il suffit de poser $f = a$.

Partie D – Cas des endomorphismes nilpotents

1. a. Soit $u : \left\{ \begin{array}{ccc} \operatorname{Ker} f^{k+1} & \to & E \\ x & \mapsto & f^k(x) \end{array} \right.$. Alors,

- $\operatorname{Ker} u = \{x \in \operatorname{Ker} f^{k+1},\ f^k(x) = 0_E\} = \operatorname{Ker} f^k \cap \operatorname{Ker} f^{k+1} = \operatorname{Ker} f^k$.
- $\operatorname{Im} u = \operatorname{Im} f^k \cap f^k(\operatorname{Ker} f^{k+1})$. Comme $f^k(\operatorname{Ker} f^{k+1}) \subset \operatorname{Ker} f$, $\operatorname{Im} u \subset \operatorname{Im} f^k \cap \operatorname{Ker} f$.

Montrons l'inclusion réciproque. Soit $y \in \operatorname{Im} f^k \cap \operatorname{Ker} f$. Par définition, $f(y) = 0_E$ et il existe $x \in E$ tel que $y = f^k(x)$; par conséquent, $f^{k+1}(x) = 0_E$ et donc $y \in f^k(\operatorname{Ker} f^{k+1})$. En conclusion, $\operatorname{Im} u = \operatorname{Im} f^k \cap \operatorname{Ker} f$.
Le théorème du rang appliqué à u donne alors

$$\dim \operatorname{Ker} f^{k+1} = \dim \operatorname{Ker} f^k + \dim \operatorname{Ker} f \cap \operatorname{Im} f^k.$$

b. On a déjà établi la décroissance de la suite $(\operatorname{Im} f^k)_k$. Par conséquent, pour tout $k \in \mathbb{N}$

$$\operatorname{Ker} f \cap \operatorname{Im} f^{k+1} \subset \operatorname{Ker} f \cap \operatorname{Im} f^k.$$

En passant à la dimension et en considérant la question précédente, on obtient

$$\dim \operatorname{Ker} f^{k+2} - \dim \operatorname{Ker} f^{k+1} \leq \dim \operatorname{Ker} f^{k+1} - \dim \operatorname{Ker} f^{k}.$$

Donc la suite $(\dim \operatorname{Ker} f^{k+1} - \dim \operatorname{Ker} f^{k})_k$ est décroissante.

2. a. L'application linéaire $\left\{\begin{array}{ccc} G & \longrightarrow & a(G) \\ x & \longmapsto & a(x) \end{array}\right.$ est bijective car elle est surjective par construction et

$$\begin{aligned} G \cap \operatorname{Ker} a &\subset G \cap \operatorname{Ker} u \text{ car } u = b \circ a \\ &\subset G \cap \operatorname{Ker} u^{k+1} = \{0_E\}. \end{aligned}$$

Ainsi, $\dim G = \dim a(G)$.

L'application linéaire $\left\{\begin{array}{ccc} H_2 & \longrightarrow & v(H_2) \\ x & \longmapsto & v(x) \end{array}\right.$ est bijective car elle est surjective par construction et

$$H_2 \cap \operatorname{Ker} v \subset H_2 \cap \operatorname{Ker} v^{k+1} = H_2 \cap H_1 = \{0_E\}.$$

Ainsi, $\dim H_2 = \dim v(H_2)$.

b. ▹ Remarquons tout d'abord les inclusions $H_1 \subset \operatorname{Ker} v^{k+1}$ (par définition), $\operatorname{Ker} v^k \subset \operatorname{Ker} v^{k+1}$ (croissance de la suite des noyaux itérés) et

$$v(H_2) \subset v(a(G)) \subset v(a(\operatorname{Ker} u^{k+1})) \subset v(\operatorname{Ker} v^{k+1}) \subset \operatorname{Ker} v^{k+1}.$$

▹ Passons maintenant à l'aspect somme directe en considérant $y_1 \in H_1$, $y_2 \in H_2$ tels que $y_1 + v(y_2) \in \operatorname{Ker} v^k$. Comme $H_1 \cup H_2 \subset a(G)$, introduisons x_1 et $x_2 \in G$ tels que $y_1 = a(x_1)$ et $y_2 = a(x_2)$. Alors,

$$\begin{aligned} 0_E &= v^k(y_1 + v(y_2)) = v^k(a(x_1)) + v^{k+1}(a(x_2)) \\ &= a \circ (b \circ a)^k(x_1) + a \circ (b \circ a)^{k+1}(x_2) \\ &= a \circ u^k(x_1) + a \circ u^{k+1}(x_2). \end{aligned}$$

Ainsi, en composant par b, $u^{k+1}(x_1) + u^{k+2}(x_2) = 0_E$. Or, $x_2 \in G \subset \operatorname{Ker} u^{k+2}$ d'où $u^{k+1}(x_1) = 0_E$ mais alors $x_1 \in \operatorname{Ker} u^{k+1} \cap G = \{0_E\}$ et $y_1 = 0_E$. Par conséquent, $v(y_2) \in \operatorname{Ker} v^k$ soit encore $y_2 \in \operatorname{Ker} v^{k+1}$. Comme $\operatorname{Ker} v^{k+1} \cap H_2 = \{0_E\}$, on en déduit $y_2 = 0_E$.

Nous avons donc établi que la seule décomposition nulle sur la somme $H_1 + v(H_2) + \operatorname{Ker} v^k$ est celle dont les trois composantes sont nulles : la somme est directe.

c. D'après la question précédente,

$$\begin{aligned} \dim \operatorname{Ker} v^{k+1} &\geq \dim \operatorname{Ker} v^k + \dim H_1 + \dim v(H_2) \\ &\geq \dim \operatorname{Ker} v^k + \dim H_1 + \dim H_2 \\ &\geq \dim \operatorname{Ker} v^k + \dim a(G) \\ &\geq \dim \operatorname{Ker} v^k + \dim G \\ &\geq \dim \operatorname{Ker} v^k + \dim \operatorname{Ker} u^{k+2} - \dim \operatorname{Ker} u^{k+1}. \end{aligned}$$

En conclusion,

$$\dim \operatorname{Ker} v^{k+1} - \dim \operatorname{Ker} v^k \geq \dim G = \dim \operatorname{Ker} u^{k+2} - \dim \operatorname{Ker} u^{k+1}.$$

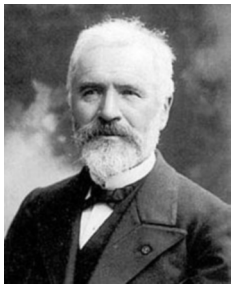
Chapitre 16

Calcul matriciel

1. Matrices — **2.** Produit matriciel — **3.** Calcul de puissances

Objectifs et compétences du programme

- Connaître les structures algébriques des ensembles matriciels.
- Calculer des produits de matrices.
- Calculer des puissances de matrices carrées.
- Manipuler trace et transposition de matrices.



Camille Jordan, mathématicien français (1838-1922)
Il a révolutionné plus d'un domaine des mathématiques et on trouve sa trace dans de nombreuses appelations :

- avec le théorème de Jordan-Hölder, il fournit un outil puissant d'étude des groupes finis ;
- le théorème de Jordan-Schur se penche sur le cas particulier des groupes finis des groupes linéaires complexes ;
- une courbe de Jordan est une courbe plane simple et fermée et le théorème associé (qui est loin d'être élémentaire contrairement aux apparences) est que son complémentaire dans le plan comporte deux composantes connexes dont l'une est bornée ;
- la réduction de Jordan permet de donner une forme standardisée pour les endomorphismes.

En revanche, la méthode du pivot de Gauss ou méthode d'élimination de Gauss-Jordan est due à un homonyme contemporain : Wilhelm Jordan (1842 – 1899), un géodésien allemand.

COURS

Dans ce chapitre, on introduit les matrices comme un objet mathématique qui permet de traduire les propriétés des applications linéaires entre espaces de dimension finie. L'addition et la composition des applications linéaires deviennent des opérations sur les matrices dont il convient de connaître aussi bien les propriétés théoriques que le maniement pratique.

1. Matrices

1.1. Définitions

Définition 16.1.

Soit $n, p \in \mathbb{N} \setminus \{0\}$. Une matrice à n lignes et p colonnes (ou de taille n, p) à coefficients dans le corps $\mathbb{K}$ est une famille $(a_{i,j})_{i\in[1,n], j\in[1,p]}$ d'éléments de $\mathbb{K}$. On la note

$$\begin{bmatrix} a_{1,1} & \dots & a_{1,p} \\ \vdots & & \vdots \\ a_{n,1} & \dots & a_{n,p} \end{bmatrix} \text{ ou } \begin{pmatrix} a_{1,1} & \dots & a_{1,p} \\ \vdots & & \vdots \\ a_{n,1} & \dots & a_{n,p} \end{pmatrix}.$$

L'ensemble des matrices de taille n, p à coefficients dans $\mathbb{K}$ est noté $\mathscr{M}_{n,p}(\mathbb{K})$ ou plus simplement $\mathscr{M}_n(\mathbb{K})$ lorsque $n = p$.

On note quelquefois $[A]_{i,j}$ le coefficient en position (i, j) de A.

Exemple

La matrice $A = (a_{i,j}) \in \mathscr{M}_{2,3}(\mathbb{R})$ dont les coefficients sont

$$a_{i,j} = \begin{cases} 1 & \text{si } i + j = 0[2] \\ 0 & \text{sinon,} \end{cases}$$

est $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$.

Définition 16.2.

Soit $A = (a_{i,j}) \in \mathscr{M}_{n,p}(\mathbb{K})$.

▹ Pour tout $k \in [\![1, p]\!]$, la $k^{\text{ième}}$ colonne de A est la matrice

$$C_k(A) = \begin{bmatrix} a_{1,k} \\ \vdots \\ a_{n,k} \end{bmatrix} \in \mathscr{M}_{n,1}(\mathbb{K}).$$

▹ Pour tout $k \in [\![1, n]\!]$, la $k^{\text{ième}}$ ligne de A est la matrice

$$L_k(A) = \begin{bmatrix} a_{k,1} & \dots & a_{k,p} \end{bmatrix} \in \mathscr{M}_{1,p}(\mathbb{K}).$$

Définition 16.3.

▹ Soit $A \in \mathscr{M}_{n,p}(\mathbb{K})$ et $B \in \mathscr{M}_{n,q}(\mathbb{K})$ (les deux matrices ont le même nombre de lignes). La matrice $\begin{bmatrix} A & B \end{bmatrix}$ est la matrice de $\mathscr{M}_{n,p+q}(\mathbb{K})$ obtenue en concaténant les colonnes de A puis les colonnes de B.
▹ Soit $A \in \mathscr{M}_{m,p}(\mathbb{K})$ et $B \in \mathscr{M}_{n,p}(\mathbb{K})$ (les deux matrices ont le même nombre de colonnes). La matrice $\begin{bmatrix} A \\ B \end{bmatrix}$ est la matrice de $\mathscr{M}_{m+n,p}(\mathbb{K})$ obtenue en concaténant les lignes de A puis les lignes de B.

Exemple

Si $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ et $B = \begin{bmatrix} 5 & 6 & 7 \\ 8 & 9 & 10 \end{bmatrix}$, alors $\begin{bmatrix} A & B \end{bmatrix} = \begin{bmatrix} 1 & 2 & 5 & 6 & 7 \\ 3 & 4 & 8 & 9 & 10 \end{bmatrix}$.

Remarque

On peut ainsi définir une matrice par blocs $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$ pour des matrices vérifiant

- A et B (respectivement C et D) ont le même nombre de lignes ;
- A et C (respectivement B et D) ont le même nombre de colonnes.

1.2. Matrice d'une application linéaire

Définition 16.4.

Soit E un espace vectoriel de dimension p et $\mathscr{B} = (e_1, \dots, e_p)$ une base de E. La matrice d'un vecteur $x = \sum_{i=1}^{p} x_i e_i \in E$ dans la base $\mathscr{B}$ est la matrice colonne

$$\mathrm{mat}_{\mathscr{B}}(x) = \begin{bmatrix} x_1 \\ \vdots \\ x_p \end{bmatrix} \in \mathscr{M}_{p,1}(\mathbb{K}).$$

Exemple

Dans la base de $\mathbb{R}_n[X]$ formée des polynômes interpolateurs de Lagrange $(L_0, \dots, L_n)$ aux points $a_0 < \dots < a_n$, le polynôme $P \in \mathbb{R}_n[X]$ admet pour matrice

$$\begin{bmatrix} P(a_0) \\ \vdots \\ P(a_n) \end{bmatrix},$$

puisque $P = \sum_{k=0}^{n} P(a_k) L_k$.

Définition 16.5.

Soit E, F deux espaces vectoriels de dimension p et n respectivement, $\mathscr{B}=(e_1,\dots,e_p)$ une base de E, $\mathscr{C}$ une base de F et $u\in\mathscr{L}(E,F)$. La matrice de u relativement aux bases $\mathscr{B}$ et $\mathscr{C}$ est la matrice de $\mathscr{M}_{n,p}(\mathbb{K})$ dont les colonnes sont les matrices colonnes des images des vecteurs de $\mathscr{B}$ dans la base $\mathscr{C}$

$$\mathrm{mat}_{\mathscr{B},\mathscr{C}}(u)=\begin{bmatrix} \mathrm{mat}_{\mathscr{C}}(u(e_1)) & \dots & \mathrm{mat}_{\mathscr{C}}(u(e_p)) \end{bmatrix}\in\mathscr{M}_{n,p}(\mathbb{K}).$$

En notant $\mathscr{B}=(e_1,\dots,e_p)$ une base de E, $\mathscr{C}=(f_1,\dots,f_n)$ une base de F, la matrice $A=(a_{i,j})$ de $u\in\mathscr{L}(E,F)$ dans les bases $\mathscr{B}$ et $\mathscr{C}$ est donc

$$\begin{array}{cc} \begin{array}{cccc} u(e_1) & u(e_2) & \cdots & u(e_p) \\ \downarrow & \downarrow & & \downarrow \end{array} & \\ \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,p} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,p} \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,p} \end{bmatrix} & \begin{array}{c} \to f_1 \\ \to f_2 \\ \vdots \\ \vdots \\ \to f_n \end{array} \end{array}$$

Remarque

Si les bases $\mathscr{B}$ et $\mathscr{C}$ sont fixées, alors la matrice $\mathrm{mat}_{\mathscr{B},\mathscr{C}}(u)$ suffit à caractériser l'application linéaire puisque les colonnes déterminent les images d'une base par u (en l'occurrence $\mathscr{B}$).

Exemple

La matrice de l'endomorphisme

$$\varphi:\left\{\begin{array}{ccc} \mathbb{R}_n[X] & \to & \mathbb{R}_n[X] \\ P(X) & \mapsto & P(X+1) \end{array}\right.$$

dans la base canonique de $\mathbb{R}_n[X]$ (au départ comme à l'arrivée) est la matrice de $\mathscr{M}_{n+1}(\mathbb{R})$ définie par

$$\begin{bmatrix} 1 & 1 & 1 & \dots & \dots & \binom{n}{0} \\ 0 & 1 & 2 & & & \binom{n}{1} \\ 0 & 0 & 1 & & & \vdots \\ \vdots & \vdots & \vdots & & & \vdots \\ \vdots & \vdots & \vdots & & & \binom{n}{n-1} \\ 0 & 0 & 0 & \dots & \dots & \binom{n}{n} \end{bmatrix},$$

car $\varphi(X^k)=\sum_{i=0}^{k}\binom{k}{i}X^i$ pour tout $k\in[\![0,n]\!]$ d'après la formule du binôme. Le coefficient en position (i,j) de la matrice est donc $\binom{j-1}{i-1}$ si $i<j$, 0 sinon.

Définition 16.6.

L'application linéaire canoniquement associée à $A\in\mathscr{M}_{n,p}(\mathbb{K})$ est l'application linéaire de $\mathbb{K}^p$ dans $\mathbb{K}^n$ dont la matrice relativement aux bases canoniques est A.

Exemple

L'endomorphisme canoniquement associé à la matrice $\begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix}$ est

$$u : \left\{ \begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ (x,y) & \longmapsto & (x+2y, 3x) \end{array} \right.$$

car

$$\begin{aligned} u(x,y) &= x\,u(1,0) + y\,u(0,1) \\ &= x(1(1,0)+3(0,1)) + y(2(1,0)+0(0,1)) \\ &= (x+2y, 3x). \end{aligned}$$

1.3. Structure d'espace vectoriel

Définition 16.7.

Soit $A = (a_{i,j}) \in \mathscr{M}_{n,p}(\mathbb{K})$, $B = (b_{i,j}) \in \mathscr{M}_{n,p}(\mathbb{K})$ et $\lambda \in \mathbb{K}$.
▹ La matrice $A+B$ est la matrice de coefficients $a_{i,j} + b_{i,j}$ pour tout i, j.
▹ La matrice λA est la matrice de coefficients $\lambda . a_{i,j}$ pour tout i, j.
Muni de ces deux lois (interne et externe), $\mathscr{M}_{n,p}(\mathbb{K})$ est un espace vectoriel.

Proposition 16.8.

Soit E, F deux espaces vectoriels de dimension p et n respectivement, $\mathscr{B}$ une base de E et $\mathscr{C}$ une base de F. L'application

$$\Phi : \left\{ \begin{array}{ccc} \mathscr{L}(E,F) & \longrightarrow & \mathscr{M}_{n,p}(\mathbb{K}) \\ u & \longmapsto & \mathrm{mat}_{\mathscr{B},\mathscr{C}}(u) \end{array} \right.$$

est un isomorphisme. Par conséquent, $\dim_{\mathbb{K}} \mathscr{M}_{n,p}(\mathbb{K}) = np$.

Démonstration
La propriété de linéarité provient essentiellement de la linéarité des applications coordonnées. La propriété de bijection découle de l'unique détermination d'une application linéaire par l'image d'une base. ■

Remarque
Intuitivement, la propriété $\dim \mathscr{M}_{n,p}(\mathbb{K}) = np$ est relativement facile à comprendre : pour déterminer une matrice de $\mathscr{M}_{n,p}(\mathbb{K})$ il faut fixer $n \times p$ coefficients.

1.4. Matrices remarquables

Définition 16.9.

La matrice nulle de taille n, p est la matrice de $\mathscr{M}_{n,p}(\mathbb{K})$, notée $0_{n,p}$ ou 0_n si $n = p$, dont tous les coefficients sont nuls. C'est l'élément neutre de l'addition dans $\mathscr{M}_{n,p}(\mathbb{K})$. La seule application linéaire associée à cette matrice est l'application nulle.

Définition 16.10.

La matrice identité de taille n est la matrice de $\mathscr{M}_n(\mathbb{K})$, notée I_n, dont les coefficients sont

$$a_{i,j} = \delta_{i,j} = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{sinon.} \end{cases}$$

L'endomorphisme canoniquement associé à cette matrice est l'application linéaire identité.

Définition 16.11.

Dans $\mathscr{M}_{n,p}(\mathbb{K})$, la matrice $E_{i,j}$ a tous les coefficients nuls sauf celui en position i, j qui vaut 1 ; en d'autres termes, le coefficient en position k, l est $\delta_{i,k}\delta_{j,l}$.

Exemple

Voici les matrices $E_{i,j}$ de $\mathscr{M}_{2,3}(\mathbb{K})$

$$E_{1,1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad E_{1,2} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad E_{1,3} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix},$$

$$E_{2,1} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad E_{2,2} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad E_{2,3} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Proposition 16.12.

Les matrices $E_{i,j}$ pour $i \in [\![1, n]\!]$ et $j \in [\![1, p]\!]$ forment une base de $\mathscr{M}_{n,p}(\mathbb{K})$ appelée base canonique. Plus précisément, la matrice $A = (a_{i,j}) \in \mathscr{M}_{n,p}(\mathbb{K})$ s'écrit

$$A = \sum_{i=1}^{n} \sum_{j=1}^{p} a_{i,j} E_{i,j}.$$

Définition 16.13.

Une matrice carrée $A = (a_{i,j}) \in \mathscr{M}_n(\mathbb{K})$ est diagonale si $a_{i,j} = 0$ pour tout $i \neq j$. On note

$$\mathrm{diag}(\alpha_1, \ldots, \alpha_n) = \begin{bmatrix} \alpha_1 & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & \alpha_n \end{bmatrix}.$$

Une matrice est scalaire si elle est diagonale à coefficients diagonaux tous égaux, c'est-à-dire si elle est proportionnelle à I_n.

Remarque

Pour des raisons d'interprétation en termes d'endomorphisme comme pour des facilités calculatoires, on ne considère jamais sérieusement l'autre « diagonale » du coin bas gauche vers le coin haut droit dans une matrice carrée.

Proposition 16.14.

L'ensemble des matrices diagonales de $\mathscr{M}_n(\mathbb{K})$ est un espace vectoriel de dimension n ; une base est constituée des matrices $E_{i,i}$ avec $i \in [\![1, n]\!]$.

Définition 16.15.

Une matrice carrée $A = (a_{i,j}) \in \mathscr{M}_n(\mathbb{K})$ est triangulaire supérieure (respectivement triangulaire strictement supérieure) si $a_{i,j} = 0$ pour tout $i > j$ (respectivement $i \geq j$).

On définit de même les matrices triangulaires inférieures et strictement inférieures.

Proposition 16.16.

L'ensemble des matrices triangulaires supérieures est un espace vectoriel de dimension $\frac{n(n+1)}{2}$ dont un supplémentaire est l'espace des matrices triangulaires strictement inférieures.

Démonstration

Si $A = (a_{i,j}) \in \mathscr{M}_n(\mathbb{K})$ est triangulaire supérieure, alors

$$A = \sum_{i,j} a_{i,j} E_{i,j} = \sum_{i \leq j} a_{i,j} E_{i,j}.$$

Ainsi, la famille $(E_{i,j})_{i \leq j}$ est génératrice de l'espace des matrices triangulaires supérieures ; cette famille est libre (car extraite de la base canonique de $\mathscr{M}_n(\mathbb{R})$) et de cardinal $\frac{n(n+1)}{2}$. En conclusion, l'espace des matrices triangulaires supérieures admet pour base $(E_{i,j})_{i \leq j}$ et sa dimension est $\frac{n(n+1)}{2}$ ■

1.5. Transposition

Définition 16.17.

La transposée d'une matrice $A = (a_{i,j}) \in \mathscr{M}_{n,p}(\mathbb{K})$ est la matrice

$$A^{\mathrm{T}} = (a_{j,i}) \in \mathscr{M}_{p,n}(\mathbb{K}).$$

Intuitivement, la matrice A^{T} est obtenue par symétrie de A par rapport à la première diagonale.

Exemple

La transposée de $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$ est $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$.

Remarque

Une matrice carrée A est triangulaire supérieure si, et seulement si, A^{T} est triangulaire inférieure.

Proposition 16.18.

L'application $\left\{\begin{array}{ccc} \mathscr{M}_{n,p}(\mathbb{K}) & \to & \mathscr{M}_{p,n}(\mathbb{K}) \\ A & \mapsto & A^{\mathrm{T}} \end{array}\right.$ est un isomorphisme.

Démonstration

L'application est clairement linéaire et la bijection réciproque (distincte de l'application de départ si les dimensions n et p sont différentes) est

$$\left\{\begin{array}{ccc} \mathscr{M}_{p,n}(\mathbb{K}) & \to & \mathscr{M}_{n,p}(\mathbb{K}) \\ A & \mapsto & A^{\mathrm{T}}. \end{array}\right.$$

■

Définition 16.19.

Soit $A \in \mathscr{M}_n(\mathbb{K})$.
▹ La matrice A est symétrique si $A^{\mathrm{T}} = A$.
▹ La matrice A est antisymétrique si $A^{\mathrm{T}} = -A$.
Notons $\mathscr{S}_n(\mathbb{K})$ et $\mathscr{A}_n(\mathbb{K})$ les ensembles des matrices symétriques et antisymétriques respectivement.

Remarque

La diagonale d'une matrice antisymétrique est composée de 0 puisque par définition, chaque élément diagonal est égal à son opposé.

Proposition 16.20.

Les ensembles $\mathscr{S}_n(\mathbb{K})$ et $\mathscr{A}_n(\mathbb{K})$ sont des sous-espaces vectoriels supplémentaires dans $\mathscr{M}_n(\mathbb{K})$.
De plus,
▹ $\mathscr{S}_n(\mathbb{K})$ est de dimension $\frac{n(n+1)}{2}$ et une base est donnée par les vecteurs $E_{i,i}$ pour $i \in [\![1,n]\!]$ et les vecteurs $E_{i,j} + E_{j,i}$ pour $1 \leq i < j \leq n$.
▹ $\mathscr{A}_n(\mathbb{K})$ est de dimension $\frac{n(n-1)}{2}$ et une base est donnée par les vecteurs $E_{i,j} - E_{j,i}$ pour $1 \leq i < j \leq n$.

Démonstration

▹ L'application linéaire

$$u : \left\{\begin{array}{ccc} \mathscr{M}_n(\mathbb{K}) & \to & \mathscr{M}_n(\mathbb{K}) \\ A & \mapsto & A^{\mathrm{T}} \end{array}\right.$$

est une symétrie (car la transposition est involutive sur $\mathscr{M}_n(\mathbb{K})$) de $\mathscr{M}_n(\mathbb{K})$ par rapport à $\mathscr{S}_n(\mathbb{K}) = \mathrm{Ker}(u - \mathrm{Id})$ et parallèlement à $\mathscr{A}_n(\mathbb{K}) = \mathrm{Ker}(u + \mathrm{Id})$ donc les deux sous-espaces sont supplémentaires.
▹ Soit $S = (s_{i,j}) \in \mathscr{S}_n(\mathbb{K})$.

$$\begin{aligned} S &= \sum_{i=1}^{n} \sum_{j=1}^{n} s_{i,j} E_{i,j} \\ &= \sum_{i=1}^{n} s_{i,i} E_{i,i} + \sum_{1 \leq i < j < n} s_{i,j}(E_{i,j} + E_{j,i}). \end{aligned}$$

La famille des vecteurs $E_{i,i}$ pour $i \in [\![1,n]\!]$ et des vecteurs $E_{i,j} + E_{j,i}$ pour $1 \leq i < j \leq n$ est donc génératrice de $\mathscr{S}_n(\mathbb{K})$ or elle est libre donc c'est une base de $\mathscr{S}_n(\mathbb{K})$.
On procède de même pour $\mathscr{A}_n(\mathbb{K})$. ■

Exemple

Considérons une matrice $A \in \mathscr{M}_3(\mathbb{K})$. Par définition, elle est de la forme

$$A = \begin{bmatrix} 0 & a_{1,2} & a_{1,3} \\ -a_{1,2} & 0 & a_{2,3} \\ -a_{1,3} & -a_{2,3} & 0 \end{bmatrix}.$$

Sa décomposition dans la base précédente est donc

$$A = a_{1,2} \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + a_{1,3} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} + a_{2,3} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}.$$

2. Produit matriciel

2.1. Définitions et exemples

Définition 16.21.

Soit $A = (a_{i,j}) \in \mathscr{M}_{n,p}(\mathbb{K})$ et $X = (x_i) \in \mathscr{M}_{p,1}(\mathbb{K})$. La matrice $Y = AX = (y_i) \in \mathscr{M}_{n,1}(\mathbb{K})$ est définie par les coefficients

$$\forall i \in [\![1, n]\!], \qquad y_i = \sum_{k=1}^{p} a_{i,k} x_k.$$

De manière équivalente, $Y = \sum_{k=1}^{p} x_k C_k(A)$.

Conseils méthodologiques

Remarquons que pour pouvoir calculer le produit AX, il faut que le nombre de colonnes de A soit le nombre de lignes de X ; dans ce cas, le nombre de lignes de AX est le nombre de lignes de A.

Exemple

Dans le calcul matriciel ci-dessous, le calcul du coefficient en deuxième position est obtenu par $0 \times 1 + 1 \times 1 + 2 \times 0 + 3 \times (-1) = -2$.

$$\begin{bmatrix} 1 & 2 & 3 & -1 \\ 0 & 1 & 2 & 3 \\ 1 & 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 4 \\ -2 \\ -1 \end{bmatrix}.$$

Proposition 16.22.

Soit E, F deux espaces vectoriels de dimension finie, $\mathscr{B}$ une base de E et $\mathscr{C}$ une base de F. Alors, pour tout $u \in \mathscr{L}(E, F)$ et tout $x \in E$,

$$\mathrm{mat}_{\mathscr{C}}(u(x)) = \mathrm{mat}_{\mathscr{B},\mathscr{C}}(u)\,\mathrm{mat}_{\mathscr{B}}(x).$$

Démonstration

Notons $\mathscr{B}=(e_1,\dots,e_p)$, $\mathscr{C}=(\epsilon_1,\dots,\epsilon_n)$, $\mathrm{mat}_{\mathscr{B},\mathscr{C}}(u)=(a_{i,j})_{i\in[\![1,n]\!],j\in[\![1,p]\!]}$, $\mathrm{mat}_{\mathscr{B}}(x)=(x_i)_{i\in[\![1,p]\!]}$ et $\mathrm{mat}_{\mathscr{C}}(u(x))=(y_i)_{i\in[\![1,n]\!]}$. Alors

$$\begin{aligned} u(x)&=\sum_{j=1}^{p} x_j\, u(e_j)=\sum_{j=1}^{p} x_j \sum_{i=1}^{n} a_{i,j}\epsilon_i \\ &=\sum_{i=1}^{n}\left(\sum_{j=1}^{p} a_{i,j}x_j\right)\epsilon_i. \end{aligned}$$

En conclusion, $y_i=\sum_{j=1}^{p} a_{i,j}x_j$. ■

Définition 16.23.

Soit $A=(a_{i,j})\in\mathscr{M}_{n,p}(\mathbb{K})$ et $B=(b_{i,j})\in\mathscr{M}_{p,q}(\mathbb{K})$. La matrice $C=AB=(c_{i,j})\in\mathscr{M}_{n,q}(\mathbb{K})$ est définie par les coefficients

$$\forall i\in[\![1,n]\!],\forall j\in[\![1,q]\!],\qquad c_{i,j}=\sum_{k=1}^{p} a_{i,k}b_{k,j}.$$

Remarque

Cette définition correspond à la relation suivante sur les colonnes $C_k(AB)=AC_k(B)$ pour tout $k\in[\![1,q]\!]$ ou à celle-ci sur les lignes $L_k(AB)=L_k(A)B$ pour tout $k\in[\![1,n]\!]$.

Conseils méthodologiques

Le produit matriciel n'est en général pas commutatif (si le nombre de lignes de A est différent du nombre de colonnes de B, le produit BA n'est même pas défini). De plus, le produit de deux matrices non nulles peut être la matrice nulle. Par exemple,

$$\begin{bmatrix}1&0\\0&0\end{bmatrix}\begin{bmatrix}0&0\\0&1\end{bmatrix}=\begin{bmatrix}0&0\\0&0\end{bmatrix},\qquad \begin{bmatrix}0&1\\0&0\end{bmatrix}\begin{bmatrix}0&1\\0&0\end{bmatrix}=\begin{bmatrix}0&0\\0&0\end{bmatrix}.$$

$$\begin{bmatrix}1&1\\-1&-1\end{bmatrix}\begin{bmatrix}1&1\\-1&-1\end{bmatrix}=\begin{bmatrix}0&0\\0&0\end{bmatrix}$$

Remarque

En utilisant l'associativité de l'addition dans la formule définissant les coefficients du produit, on remarque que les produits peuvent être effectués par blocs (sous réserve de tailles concordantes des blocs et de faire attention à l'ordre des produits entre blocs). Voici par exemple un produit par blocs

$$\begin{bmatrix}A_{1,1}&A_{1,2}\\A_{2,1}&A_{2,2}\end{bmatrix}\begin{bmatrix}B_{1,1}&B_{1,2}\\B_{2,1}&B_{2,2}\end{bmatrix}=\begin{bmatrix}A_{1,1}B_{1,1}+A_{1,2}B_{2,1}&A_{1,1}B_{1,2}+A_{1,2}B_{2,2}\\A_{2,1}B_{1,1}+A_{2,2}B_{2,1}&A_{2,1}B_{1,2}+A_{2,2}B_{2,2}\end{bmatrix}$$

Exemple

Une disposition pratique pour le calcul effectif du produit matriciel consiste à placer la première matrice à gauche et la seconde au-dessus du résultat. Ainsi, pour effectuer le calcul

On écrit

$$\begin{array}{cc} & \begin{bmatrix} 1 & 2 & 3 & 4 \\ -1 & 0 & 1 & -2 \end{bmatrix} \\ \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} & \begin{bmatrix} -1 & 2 & 5 & 0 \\ -1 & 6 & 13 & 4 \\ -1 & 10 & 21 & 8 \end{bmatrix} \end{array}.$$

Remarque

Soit $A=(a_{i,j})\in\mathscr{M}_{n,p}(\mathbb{K})$ et $B=(b_{i,j})\in\mathscr{M}_{p,q}(\mathbb{K})$.

▹ Si la $k^{\text{ième}}$ ligne de A est nulle alors la $k^{\text{ième}}$ ligne de AB est nulle car $L_k(AB = L_k(A)B$; on le visualise ainsi avec la disposition des calculs :

$$\begin{array}{cc} & \begin{bmatrix} \star & \star & \star \\ \star & \star & \star \\ \star & \star & \star \\ \star & \star & \star \end{bmatrix} \\ \begin{bmatrix} \star & \star & \star & \star \\ 0 & \dots & \dots & 0 \\ \star & \star & \star & \star \end{bmatrix} & \begin{bmatrix} \star & \star & \star \\ 0 & \dots & 0 \\ \star & \star & \star \end{bmatrix} \end{array}.$$

▹ Si la $k^{\text{ième}}$ colonne de B est nulle alors la $k^{\text{ième}}$ colonne de AB est nulle.

Proposition 16.24.

Soit E, F, G trois espaces vectoriels de dimension finie, $\mathscr{B}$ une base de E, $\mathscr{C}$ une base de F et $\mathscr{D}$ une base de G. Alors, pour tout $u\in\mathscr{L}(E,F)$ et tout $v\in\mathscr{L}(F,G)$,

$$\mathrm{mat}_{\mathscr{B},\mathscr{D}}(v\circ u)=\mathrm{mat}_{\mathscr{C},\mathscr{D}}(v)\mathrm{mat}_{\mathscr{B},\mathscr{C}}(u).$$

Démonstration

Notons $\mathscr{B}=(e_1,\dots,e_p)$. Pour tout $k\in[\![1,p]\!]$,

$$\begin{aligned} C_k(\mathrm{mat}_{\mathscr{B},\mathscr{D}}(v\circ u)) &= \mathrm{mat}_{\mathscr{D}}(v\circ u(e_k)) = \mathrm{mat}_{\mathscr{C},\mathscr{D}}(v)\mathrm{mat}_{\mathscr{C}}(u(e_k)) \\ &= \mathrm{mat}_{\mathscr{C},\mathscr{D}}(v)C_k(\mathrm{mat}_{\mathscr{B},\mathscr{C}}(u)). \end{aligned}$$

Ainsi, $\mathrm{mat}_{\mathscr{B},\mathscr{D}}(v\circ u)=\mathrm{mat}_{\mathscr{C},\mathscr{D}}(v)\mathrm{mat}_{\mathscr{B},\mathscr{C}}(u)$. ■

Si on utilise la proposition précédente, on déduit les propriétés suivantes du produit matriciel des propriétés correspondantes de la composition.

Proposition 16.25.

Soit A, B, C trois matrices et $\lambda \in \mathbb{K}$. Lorsque les produits considérés existent (c'est-à-dire quand les tailles des matrices concordent),

- $(A+B)C = AC + BC$
- $A(B+C) = AB + AC$
- $(AB)C = A(BC)$
- $(\lambda A)B = A(\lambda B) = \lambda(AB)$.

Démonstration

Montrons par exemple l'avant-dernière propriété.

Pour cela, posons $A = (a_{i,j}) \in \mathscr{M}_{n,p}(\mathbb{K})$, $B = (b_{i,j}) \in \mathscr{M}_{p,q}(\mathbb{K})$ et $C = (c_{i,j}) \in \mathscr{M}_{q,r}(\mathbb{K})$. Alors, pour tout $i \leq n$ et tout $j \leq r$,

$$\sum_{l=1}^{q}\left(\sum_{k=1}^{p} a_{i,k} b_{k,l}\right) c_{l,j} = \sum_{k=1}^{p} a_{i,k}\left(\sum_{l=1}^{q} b_{k,l} c_{l,j}\right).$$

Par conséquent, $(AB)C$ et $A(BC)$ ont le même coefficient en position (i,j) pour tout i et j donc ces matrices sont égales. ■

Exemple

Pour toutes matrices $A \in \mathscr{M}_{n,p}(\mathbb{K})$ et $B \in \mathscr{M}_{p,q}(\mathbb{K})$, $(AB)^{\mathrm{T}} = B^{\mathrm{T}} A^{\mathrm{T}}$ car le coefficient en position (i,j) de $(AB)^{\mathrm{T}}$ (c'est-à-dire le coefficient en position (j,i) de BA) est

$$\sum_{k=1}^{p} a_{j,k} b_{k,i} = \sum_{k=1}^{p} b_{k,i} a_{j,k}.$$

2.2. Cas des matrices carrées

Exemple

Dans $\mathscr{M}_n(\mathbb{K})$, le produit de deux matrices de la base canonique est $E_{i,j} E_{k,l} = \delta_{j,k} E_{i,l}$.

Proposition 16.26.

L'espace $\mathscr{M}_n(\mathbb{K})$ est un anneau d'éléments neutres 0_n et I_n.

Proposition 16.27.

L'ensemble des matrices triangulaires supérieures est un sous-anneau de $\mathscr{M}_n(\mathbb{K})$.

Démonstration

Il suffit de montrer que le produit de deux matrices triangulaires supérieures $A = (a_{i,j})$ et $B = (b_{i,j})$ est une matrice triangulaire supérieure $AB = (c_{i,j})$. Par définition, $a_{i,j} = b_{i,j} = 0$ pour $i > j$. Par conséquent,

pour $i > j$,

$$\begin{aligned} c_{i,j} &= \sum_{k=1}^{n} a_{i,k} b_{k,j} = \sum_{k=1}^{i-1} a_{i,k} b_{k,j} + \sum_{k=i}^{n} a_{i,k} b_{k,j} \\ &= \sum_{k=1}^{i-1} 0.b_{k,j} + \sum_{k=i}^{n} a_{i,k}.0 = 0. \end{aligned}$$

Ainsi, AB est triangulaire supérieure. ■

Remarque

Avec les notations de la démonstration précédente, pour tout $i \in [\![1, n]\!]$,

$$c_{i,i} = \sum_{k=1}^{i-1} a_{i,k} b_{k,i} + a_{i,i} b_{i,i} + \sum_{k=i+1}^{n} a_{i,k} b_{k,i} = a_{i,i} b_{i,i}.$$

Ainsi, les termes diagonaux du produit de deux matrices triangulaires supérieures est le produit des éléments diagonaux.

Définition 16.28.

Le groupe linéaire d'indice n sur $\mathbb{K}$ est le groupe, noté $\mathrm{GL}_n(\mathbb{K})$, des matrices inversibles de l'anneau $\mathscr{M}_n(\mathbb{K})$.

Proposition 16.29.

L'inverse d'une matrice triangulaire supérieure inversible est une matrice triangulaire supérieure.

Démonstration

Notons $\mathscr{T}_n(\mathbb{K})$ l'ensemble des matrices triangulaires supérieures. Considérons une matrice $A \in \mathscr{T}_n(\mathbb{K})$ inversible et posons

$$\varphi : \left\{ \begin{array}{ccc} \mathscr{T}_n(\mathbb{K}) & \to & \mathscr{T}_n(\mathbb{K}) \\ M & \mapsto & AM \end{array} \right.$$

L'endomorphisme φ est injectif car la condition $\varphi(M) = 0_n$ équivaut successivement à $AM = 0_n$ puis $A^{-1}AM = 0_n$ et enfin $M = 0_n$.

Par conséquent, φ est aussi surjectif et il existe $B \in \mathscr{T}_n(\mathbb{K})$ tel que $\varphi(B) = \mathrm{I}_n$, c'est-à-dire $AB = \mathrm{I}_n$ donc $B = A^{-1}$. ■

Remarque

Soit $A \in \mathrm{GL}_n(\mathbb{K})$. On obtient que la matrice A^{T} est inversible d'inverse $(A^{\mathrm{T}})^{-1} = (A^{-1})^{\mathrm{T}}$ en transposant la relation $AA^{-1} = A^{-1}A = \mathrm{I}_n$.

2.3. Trace

Définition 16.30.

La trace de la matrice carrée $A = (a_{i,j}) \in \mathscr{M}_n(\mathbb{K})$ est le scalaire, noté $\operatorname{tr} A$, égal à la somme des coefficients diagonaux $\sum_{i=1}^{n} a_{i,i}$.

Proposition 16.31.

L'application trace est une forme linéaire sur $\mathscr{M}_n(\mathbb{K})$.

Proposition 16.32.

Pour toutes matrices A, $B \in \mathscr{M}_n(\mathbb{K})$, $\operatorname{tr}(AB) = \operatorname{tr}(BA)$. En particulier, si $A \in \mathscr{M}_n(\mathbb{K})$ et $P \in \mathrm{GL}_n(\mathbb{K})$, alors $\operatorname{tr}(PAP^{-1}) = \operatorname{tr}(A)$.

Démonstration

▹ Notons $A = (a_{i,j})$ et $B = (b_{i,j})$. Alors,

$$\operatorname{tr}(AB) = \sum_{i=1}^{n}\sum_{k=1}^{n} a_{i,k}\, b_{k,i} = \sum_{k=1}^{n}\sum_{i=1}^{n} b_{k,i}\, a_{i,k} = \operatorname{tr}(BA).$$

▹ Pour $A \in \mathscr{M}_n(\mathbb{K})$ et $P \in \mathrm{GL}_n(\mathbb{K})$, $\operatorname{tr}(PAP^{-1}) = \operatorname{tr}(AP^{-1}P) = \operatorname{tr}(A)$. ■

Exemple

Il n'existe pas de matrices $A, B \in \mathscr{M}_n(\mathbb{R})$ telles que $AB - BA = \mathrm{I}_n$ puisque

$$\operatorname{tr}(AB - BA) = \operatorname{tr}(AB) - \operatorname{tr}(BA) = 0, \qquad \operatorname{tr}(\mathrm{I}_n) = n.$$

Exemple

Pour toute matrice $A = (a_{i,j}) \in \mathscr{M}_{n,p}(\mathbb{K})$,

$$\operatorname{tr}(A^{\mathrm{T}} A) = \sum_{i=1}^{n}\sum_{j=1}^{p} a_{i,j}^2;$$

en effet, le coefficient en position i, j de $A^{\mathrm{T}} A$ est $\sum_{k=1}^{n} a_{k,i} a_{k,j}$.

3. Calcul de puissances

Définition 16.33.

Soit $A \in \mathscr{M}_n(\mathbb{K})$.

▹ Les puissances de A sont définies récursivement par $A^0 = \mathrm{I}_n$ et

$$\forall k \in \mathbb{N}, \qquad A^{k+1} = A.A^k = A^k.A.$$

▹ Pour tout $P(X) = \sum_{k=0}^{p} \alpha_k X^k \in \mathbb{K}[X]$, la matrice $P(A)$ est définie par

$$P(A) = \sum_{k=0}^{p} \alpha_k A^k.$$

Exemple

Motivons l'étude des puissances d'une matrice carrée par un exemple issu des suites récurrentes linéaires. Considérons une suite $(u_n)_n$ vérifiant

$$\forall n \in \mathbb{N}, \qquad u_{n+4} = u_{n+3} + 2u_{n+2} + 3u_{n+1} + 4u_n.$$

Posons, pour tout $n \in \mathbb{N}$,

$$U_n = \begin{bmatrix} u_{n+3} \\ u_{n+2} \\ u_{n+1} \\ u_n \end{bmatrix}, \qquad A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

La relation de récurrence de la suite se réécrit alors $U_{n+1} = AU_n$ pour tout $n \in \mathbb{N}$.
Une rapide récurrence donne alors $U_n = A^n U_0$ pour tout $n \in \mathbb{N}$ d'où l'intérêt de savoir calculer A^n pour exprimer le terme général de la suite $(u_n)_n$ en fonction des premiers termes (et de n).

3.1. Matrices diagonalisables

Commençons par le calcul (élémentaire) des puissances de matrices diagonales.

Exemple

Soit $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ et $D = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$. Alors, pour tout $k \in \mathbb{N}$, $D^k = \operatorname{diag}(\lambda_1^k, \dots, \lambda_n^k)$ et, par conséquent, pour tout $P \in \mathbb{K}[X]$, $P(D) = \operatorname{diag}(P(\lambda_1), \dots, P(\lambda_n))$.

Définition 16.34.

Une matrice $A \in \mathscr{M}_n(\mathbb{K})$ est diagonalisable s'il existe $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ et $U \in \mathrm{GL}_n(\mathbb{K})$ tels que

$$A = U \operatorname{diag}(\lambda_1, \dots, \lambda_n) U^{-1}.$$

Toutes les matrices ne sont pas diagonalisables et trouver une telle décomposition, quand elle existe, est un travail délicat.

Proposition 16.35.

Soit $A = U \operatorname{diag}(\lambda_1, \dots, \lambda_n) U^{-1}$ avec $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ et $U \in \mathrm{GL}_n(\mathbb{K})$. Alors, pour tout $k \in \mathbb{N}$,

$$A^k = U \operatorname{diag}(\lambda_1^k, \dots, \lambda_n^k) U^{-1}$$

et, par conséquent, pour tout $P \in \mathbb{K}[X]$, $P(A) = U \operatorname{diag}(P(\lambda_1), \dots, P(\lambda_n)) U^{-1}$.

Démonstration

Il suffit d'écrire effectivement le produit et d'utiliser l'associativité du produit matriciel dans $\mathscr{M}_n(\mathbb{K})$

$$\begin{aligned} A^k &= (U \operatorname{diag}(\lambda_1, \dots, \lambda_n) U^{-1})(U \operatorname{diag}(\lambda_1, \dots, \lambda_n) U^{-1}) \dots (U \operatorname{diag}(\lambda_1, \dots, \lambda_n) U^{-1}) \\ &= U \operatorname{diag}(\lambda_1, \dots, \lambda_n)(U^{-1}U) \operatorname{diag}(\lambda_1, \dots, \lambda_n)(U^{-1}U) \dots (U^{-1}U) \operatorname{diag}(\lambda_1, \dots, \lambda_n) U^{-1} \\ &= U \operatorname{diag}(\lambda_1, \dots, \lambda_n)^k U^{-1} \qquad \text{car } U^{-1}U = \mathrm{I}_n \\ &= U \operatorname{diag}(\lambda_1^k, \dots, \lambda_n^k) U^{-1}. \end{aligned}$$

■

Exemple

Déterminons les puissances de la matrice $A \in \mathscr{M}_2(\mathbb{R})$ définie par

$$A = \begin{bmatrix} 0 & 1 \\ -2 & 3 \end{bmatrix}$$

en admettant la décomposition suivante

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}$$

où la première matrice et la troisième matrice sont inverses l'une de l'autre.
Alors, pour tout $k \in \mathbb{N}$,

$$A^k = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2^k \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 2-2^k & 2^k-2 \\ 2-2^{k+1} & 2^{k+1}-1 \end{bmatrix}.$$

3.2. Formule du binôme

Remarquons que comme $\mathscr{M}_n(\mathbb{K})$ est un anneau, on peut utiliser la formule du binôme de Newton.

Conseils méthodologiques

Pour calculer les puissances d'une matrice A, on l'écrit comme la somme de deux matrices B et C telles que

- B et C commutent, c'est-à-dire $BC = CB$,
- on sait calculer les puissances de B et de C.

On applique alors la formule du binôme à $A^k = (B+C)^k$.

Exemple

La matrice de $A \in \mathscr{M}_n(\mathbb{R})$ définie par

$$A = \begin{bmatrix} 3 & 1 & \dots & 1 \\ 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \dots & 1 & 3 \end{bmatrix}$$

s'écrit comme la somme de $2\mathrm{I}_n$ et de la matrice $J \in \mathscr{M}_n(\mathbb{R})$, dont tous les coefficients sont égaux à 1.
Remarquons la commutation $(2\mathrm{I}_n)J = 2J = J(2\mathrm{I}_n)$ donc nous pouvons appliquer la formule du binôme.
Par ailleurs, un rapide calcul donne $J^2 = nJ$; ainsi, par une récurrence immédiate, on obtient que $J^l = n^{l-1}J$ pour tout $l \in \mathbb{N} \setminus \{0\}$.

La formule du binôme donne alors, pour tout $k \in \mathbb{N} \setminus \{0\}$,

$$\begin{aligned} A^k &= \sum_{l=0}^{k} \binom{k}{l} 2^{k-l} J^l \\ &= 2^k \mathrm{I}_n + \sum_{l=1}^{k} \binom{k}{l} 2^{k-l} n^{l-1} J \\ &= 2^k \mathrm{I}_n + \frac{1}{n} \left(\sum_{l=1}^{k} \binom{k}{l} 2^{k-l} n^l \right) J \\ &= 2^k \mathrm{I}_n + \frac{1}{n}((n+2)^k - 2^k) J. \end{aligned}$$

Parmi les matrices dont les puissances sont faciles à calculer (et donc qui peuvent servir pour obtenir des décompositions simples auxquelles appliquer la formule du binôme), figurent les matrices nilpotentes.

Définition 16.36.

Une matrice $A \in \mathscr{M}_n(\mathbb{K})$ est nilpotente s'il existe $k \in \mathbb{N}$ tel que $A^k = 0_n$. Le plus petit entier $k \in \mathbb{N}$ tel que $A^k = 0_n$ est alors appelé indice de nilpotence de A.

Remarque

Si A est nilpotente d'indice k, alors, pour tout $l \geq k$, $A^l = A^k A^{l-k} = 0_n A^{l-k} = 0_n$.

Proposition 16.37.

Les matrices triangulaires supérieures strictes sont nilpotentes.

Démonstration

Si u est un endomorphisme associé à une matrice triangulaire supérieure stricte dans une base $(e_1, \ldots, e_n)$, alors $u(e_1) = 0_E$ et

$$\forall j \in [\![2, n]\!], \qquad u(e_j) \in \mathrm{Vect}(e_1, \ldots, e_{j-1}).$$

Par une récurrence immédiate, $u^k(e_j) \in \mathrm{Vect}(e_1, \ldots, e_{j-k})$ pour $j > k$ et $u^k(e_j) = 0_E$ sinon. Ainsi, u^n est l'application nulle. ■

Exemple

Voici une illustration de calcul des puissances pour une matrice triangulaire stricte :

$$A = \begin{bmatrix} 0 & 1 & 2 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{bmatrix}, \qquad A^2 = \begin{bmatrix} 0 & 0 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad A^3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Exemple

Calculons les puissances de la matrice $A \in \mathscr{M}_n(\mathbb{R})$ définie par

$$A = \begin{bmatrix} 2 & 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \dots & \dots & 0 & 2 \end{bmatrix}.$$

Notons $J = A - 2\mathrm{I}_n$ et calculons les premières puissances de J (car J est nilpotente en tant que matrice triangulaire stricte).

▹ Soit u l'endomorphisme associé à J dans la base $(e_1, \dots, e_n)$, alors $u(e_1) = 0_E$ et, pour tout $j \in [\![2, n]\!]$, $u(e_j) = e_{j-1}$. Alors, l'endomorphisme u^k associé à J^k dans cette base vérifie $u(e_j) = 0_E$ pour tout $j \leq k$ et $u(e_j) = e_{j-k}$ pour tout $j > k+1$. Ainsi, la matrice J^k est la matrice qui admet des 1 sur la $k^{\text{ième}}$ « surdiagonale » et des 0 ailleurs.

▹ La formule du binôme donne alors, pour tout $k \in \mathbb{N}$,

$$A^k = \sum_{l=0}^{k} \binom{k}{l} 2^{k-l} J^l.$$

La matrice A^k est donc triangulaire supérieure avec

- 2^k sur la diagonale,
- $\binom{k}{1} 2^{k-1}$ sur la surdiagonale,
- $\binom{k}{2} 2^{k-2}$ sur la deuxième surdiagonale,
- $\dots$

3.3. Utilisation d'un polynôme annulateur

Définition 16.38.

Soit $A \in \mathscr{M}_n(\mathbb{K})$. Un polynôme annulateur de A est un polynôme $P \in \mathbb{K}[X]$ tel que $P(A) = 0_n$.

Proposition 16.39.

Soit $A \in \mathscr{M}_n(\mathbb{K})$. Alors, il existe un polynôme annulateur de A non nul.

Démonstration

La famille $(\mathrm{I}_n, A, A^2, \dots, A^{n^2})$ comporte n^2+1 éléments dans l'espace $\mathscr{M}_n(\mathbb{K})$ qui est de dimension n^2 : elle est donc liée. Ainsi, il existe des scalaires $\alpha_0, \alpha_1, \dots, \alpha_{n^2}$ non tous nuls tels que

$$\sum_{k=0}^{n^2} \alpha_k A^k = 0_n,$$

ce qui indique que le polynôme $\sum_{k=0}^{n^2} \alpha_k X^k$ est annulateur de A et non nul. ■

Remarque

La démonstration ci-dessus indique qu'une matrice de taille n admet un polynôme annulateur non nul de degré au plus n^2. On peut montrer qu'il existe en fait un polynôme annulateur non nul de degré au plus n.

Attention, il n'y a pas de méthode élémentaire pour trouver un polynôme annulateur non nul de degré minimal.

Conseils méthodologiques

On peut utiliser un polynôme annulateur de A non nul pour calculer les puissances de A. Pour cela, on effectue la division de X^k par P : $X^k = Q_k P + R_k$ avec $\deg R_k < \deg P$. Alors,

$$A^k = Q_k(A)P(A) + R_k(A) = R_k(A).$$

Ainsi, la connaissance de R_k suffit pour calculer A^k (remarquons qu'il n'est pas nécessaire de connaître le quotient Q_k).

Remarque

Si A admet un polynôme annulateur de degré $p \in \mathbb{N}$, on obtient que toute puissance de A appartient au sous-espace $\mathrm{Vect}(\mathrm{I}_n, A, \ldots, A^{p-1})$.

Exemple

Considérons la matrice $A \in \mathscr{M}_3(R)$ définie par

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix}$$

Comme

$$A^2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & 4 \end{bmatrix}$$

on obtient la relation $A^2 - 3A + 2\mathrm{I}_3 = 0_3$ donc que $P(X) = X^2 - 3X + 2 = (X-1)(X-2)$ est annulateur. D'après la propriété de division euclidienne, pour tout $k \in \mathbb{N}$, il existe $Q_k \in \mathbb{R}[X]$, a_k, $b_k \in \mathbb{R}$ tels que

$$X^k = P(X)Q_k(X) + a_k X + b_k.$$

En évaluant en 1 et 2, on obtient le système

$$\left\{ \begin{array}{rcrcl} a_k & + & b_k & = & 1 \\ 2a_k & + & b_k & = & 2^k \end{array} \right.$$

D'où $a_k = 2^k - 1$ et $b_k = 2 - 2^k$ et par conséquent,

$$A^k = (2^k - 1)A + (2 - 2^k)\mathrm{I}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2^k - 1 \\ 0 & 0 & 2^{k+1} - 2^k \end{bmatrix}.$$

FICHE

SYNTHÈSE

Les matrices sont intuitivement des tableaux de nombres introduits pour « résumer » les vecteurs et les applications linéaires entre espaces de dimensions finies.

Espaces de matrices

L'espace $\mathscr{M}_{n,p}(\mathbb{K})$ est de dimension np. La base canonique est constituée des matrices $E_{i,j}$ dont le seul coefficient non nul est en positon (i,j) et vaut 1.

▹ Les matrices triangulaires supérieures forment un sous-espace de $\mathscr{M}_n(\mathbb{K})$ de dimension $\frac{n(n+1)}{2}$ dont un supplémentaire est l'espace des matrices triangulaires strictement inférieures.

▹ Les matrices symétriques forment un sous-espace de $\mathscr{M}_n(\mathbb{K})$ de dimension $\frac{n(n+1)}{2}$ dont un supplémentaire est l'espace des matrices antisymétriques.

▹ L'opération de transposition est une symétrie de l'espace $\mathscr{M}_n(\mathbb{K})$; la trace est une forme linéaire de celui-ci.

Il est important de savoir calculer la matrice d'une application linéaire dans des bases données ou de savoir retrouver l'endomorphisme canoniquement associé à une matrice.

Outre ces propriétés linéaires, on peut effectuer un produit de matrices (quand les tailles sont compatibles).

Produit matriciel

Soit $A=(a_{i,j})\in\mathscr{M}_{n,p}(\mathbb{K})$ et $B=(b_{i,j})\in\mathscr{M}_{p,q}(\mathbb{K})$. La matrice $C=AB=(c_{i,j})\in\mathscr{M}_{n,q}(\mathbb{K})$ est définie par les coefficients

$$\forall i\in[\![1,n]\!],\forall j\in[\![1,q]\!],\qquad c_{i,j}=\sum_{k=1}^{p}a_{i,k}b_{k,j}.$$

Muni de sa structure vectorielle et de ce produit, $\mathscr{M}_n(\mathbb{K})$ est un anneau. On note $\mathrm{GL}_n(\mathbb{K})$ le groupe des éléments inversibles de cet anneau.
Pour toutes matrices A, $B\in\mathscr{M}_n(\mathbb{K})$, $\mathrm{tr}(AB)=\mathrm{tr}(BA)$ et $(AB)^{\mathrm{T}}=B^{\mathrm{T}}A^{\mathrm{T}}$.

La structure d'anneau permet d'utiliser les formules de calcul habituelles (formule de Bernoulli, formule du binôme de Newton, sous réserve de vérifier l'hypothèse de commutation) en particulier pour calculer les puissances d'une matrice.

Le produit matriciel traduit exactement la composition des applications linéaires. Il hérite donc des propriétés de celle-ci.

EXERCICES

Vrai ou faux ?

		Vrai	Faux
a)	L'ensemble des matrices non-inversibles est un sous-espace vectoriel de $\mathscr{M}_n(\mathbb{K})$.	☐	☐
b)	La dimension de l'espace des matrices triangulaires supérieures est $\frac{n(n+1)}{2}$.	☐	☐
c)	L'ensemble des matrices triangulaires supérieures avec des 1 sur la diagonale est un sous-groupe de $\mathrm{GL}_n(\mathbb{K})$.	☐	☐
d)	$E_{i,j}E_{k,l} = 0_n$ si, et seulement si, $j \neq k$.	☐	☐
e)	Les matrices diagonales de $\mathscr{M}_n(\mathbb{K})$ commutent avec toutes les matrices de $\mathscr{M}_n(\mathbb{K})$.	☐	☐
f)	Les matrices qui commutent avec $\mathrm{diag}(1,2,\ldots,n)$ sont les matrices diagonales.	☐	☐
g)	Il existe une base de $\mathscr{M}_n(\mathbb{R})$ composée de matrices inversibles.	☐	☐
h)	La matrice de l'application identité d'un espace de dimension n est I_n quelque soit la base choisie.	☐	☐
i)	Toute matrice est somme de deux matrices inversibles.	☐	☐
j)	Pour $A \in \mathscr{M}_n(\mathbb{R})$, $\mathrm{Ker}\, A = \mathrm{Ker}\, A^{\mathrm{T}}$.	☐	☐
k)	Pour $A, B \in \mathscr{M}_n(\mathbb{R})$, $\mathrm{tr}(AB) = \mathrm{tr}(A)\,\mathrm{tr}(B)$.	☐	☐
l)	Pour $A, B \in \mathscr{M}_n(\mathbb{R})$, $\mathrm{tr}(AB) = \mathrm{tr}(BA)$.	☐	☐
m)	Pour $A, B, C \in \mathscr{M}_n(\mathbb{R})$, $\mathrm{tr}(ABC) = \mathrm{tr}(ACB)$.	☐	☐

Exercices d'application du chapitre 16

●○○ **Exercice 1**

Déterminer la dimension de l'espace des matrices de $\mathscr{M}_{3,2}(\mathbb{R})$ telles que la somme sur toute ligne et sur toute colonne est nulle.

●○○ **Exercice 2**

Déterminer la dimension de l'espace des matrices $M = (m_{i,j}) \in \mathscr{M}_3(\mathbb{R})$ telles que

$$a_{1,1} + a_{1,2} + a_{1,3} = a_{1,3} + a_{2,3} + a_{3,3} = a_{3,1} + a_{3,2} + a_{3,3} = a_{1,1} + a_{2,1} + a_{3,1}.$$

●●○ **Exercice 3**

Soit A, B et $C \in \mathscr{M}_2(\mathbb{R})$. Montrer qu'il existe λ, μ et $\nu \in \mathbb{R}$ non tous nuls tels que $\lambda A + \mu B + \nu C$ soit triangulaire supérieure à coefficients diagonaux égaux.

●○○ **Exercice 4**

Calculer JMJ où $M \in \mathscr{M}_n(\mathbb{R})$, et $J \in \mathscr{M}_n(\mathbb{R})$ est la matrice dont tous les coefficients sont égaux à 1.

●○○ **Exercice 5**

Soit $\mathscr{M}$ l'ensemble des matrices $A(t) \in \mathscr{M}_3(\mathbb{R})$, définies pour $t \in \mathbb{R}$ par

$$A(t) = \begin{bmatrix} 1-t & -t & 0 \\ -t & 1-t & 0 \\ -t & t & 1-2t \end{bmatrix}.$$

1. Montrer que $\mathscr{M}$ est stable par produit matriciel.
2. Pour quelles valeurs de t, la matrice $A(t)$ admet-elle un inverse dans $\mathscr{M}$?

●●● **Exercice 6**

Soit un entier $n \geq 2$. Déterminer les matrices de $\mathscr{M}_n(\mathbb{R})$ qui commutent avec toutes les matrices antisymétriques.

●●○ **Exercice 7**

Une matrice $A = (a_{i,j}) \in \mathscr{M}_n(\mathbb{R})$ est centrosymétrique si

$$\forall i, j \leq n, \qquad a_{n+1-i,n+1-j} = a_{i,j}.$$

Montrer que l'ensemble des matrices centrosymétriques est un sous-espace vectoriel de $\mathscr{M}_n(\mathbb{R})$, stable par produit. Déterminer sa dimension.

●○○ **Exercice 8**

Déterminer les réels λ tels qu'il existe une matrice $A \in \mathscr{M}_n(\mathbb{R})$ non nulle vérifiant $A^{\mathrm{T}} = \lambda A$.

●●○ **Exercice 9**

Soit φ une forme linéaire sur $\mathscr{M}_n(\mathbb{R})$ telle que

$$\forall A, B \in \mathscr{M}_n(\mathbb{R}), \qquad \varphi(AB) = \varphi(BA).$$

Montrer que φ est proportionnelle à la forme linéaire tr.
Indication : on pourra considérer l'image par φ de la base canonique de $\mathscr{M}_n(\mathbb{R})$.

●●○ **Exercice 10**

Calculer l'inverse de la matrice

$$A = \begin{bmatrix} 1 & \cdots & \cdots & 1 \\ 0 & 1 & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 \end{bmatrix}.$$

●○○ **Exercice 11**

Soit $\theta \in \mathbb{R}$. Calculer les puissances de la matrice

$$A = \begin{bmatrix} 1+\sin\theta & \cos\theta \\ \cos\theta & 1-\sin\theta \end{bmatrix}$$

●●○ **Exercice 12**

Calculer les puissances de la matrice $M \in \mathscr{M}_n(\mathbb{R})$ de coefficients $[M]_{i,j} = 1$ si $i = j - 1[n]$ et $[M]_{i,j} = 0$ sinon.

●○○ **Exercice 13**

Considérons les matrices $C \in \mathscr{M}_{n,1}(\mathbb{R})$ et $L \in \mathscr{M}_{1,n}(\mathbb{R})$. Calculer $(LC)^k$ et $(CL)^k$, pour tout $k \in \mathbb{N}$.

●●○ **Exercice 14**

Considérons la matrice $A \in \mathscr{M}_n(\mathbb{R})$ définie par

$$A = \begin{bmatrix} \alpha_1 & 1 & 0 & \dots & 0 \\ \alpha_2 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ \vdots & \vdots & & \ddots & 1 \\ \alpha_n & 0 & \dots & \dots & 0 \end{bmatrix}$$

et $\mathscr{C}_A$ l'espace vectoriel des matrices commutant avec A.

1. Montrer l'injectivité de l'application linéaire

$$\Psi : \left\{ \begin{array}{ccc} \mathscr{C}_A & \to & \mathscr{M}_{n,1}(\mathbb{R}) \\ (m_{i,j})_{i,j} & \mapsto & (m_{i,n})_i \end{array} \right.$$

2. En déduire que, pour tout $M \in \mathscr{C}_A$, il existe $P \in \mathbb{R}_{n-1}[X]$ tel que $M = P(A)$.

Exercices d'approfondissement du chapitre 16

●○○ **Exercice A**

Montrer que l'ensemble des matrices $M = (m_{i,j}) \in \mathscr{M}_n(\mathbb{R})$ telles que $m_{i,j} = 0$ pour tout (i, j) tel que $j = i + 1[2]$ est un sous-espace vectoriel de $\mathscr{M}_n(\mathbb{R})$, stable par produit. Préciser sa dimension.

●○○ **Exercice B**

1. Montrer que, pour toute matrice $M \in \mathscr{M}_2(\mathbb{R})$, $M^2 - (\operatorname{tr} M)M \in \operatorname{Vect}(\mathrm{I}_2)$.
2. Soit $A, B \in \mathscr{M}_2(\mathbb{R})$. Montrer que la matrice $(AB - BA)^2$ est proportionnelle à I_2

●○○ **Exercice C**

Soit A et $B \in \mathscr{M}_n(\mathbb{R})$ telles que AB soit nilpotente d'indice p.

1. Montrer que BA est nilpotente d'indice inférieur ou égal à $p + 1$.
2. Peut-il y avoir égalité dans l'inégalité précédente ?

●●○ **Exercice D**

Soit un entier $l \in \mathbb{N}$ impair, $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ deux à deux distincts et $D = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$.
Considérons l'équation (E) $M^{l+1} - M^l = D$ d'inconnue $M \in \mathscr{M}_n(\mathbb{R})$.

1. Montrer que si M est solution de (E), alors M commute avec D.
2. En déduire que M est diagonale.
3. Montrer que (E) admet des solutions si, et seulement si, $\lambda_k \geq -\frac{1}{l+1}\left(\frac{l}{l+1}\right)^l$ pour tout $k \in [\![1, n]\!]$.
4. Déterminer, dans ce cas, le nombre de solutions de (E).

EXERCICES

●●○ **Exercice E**

Soit A, $B \in \mathscr{M}_n(\mathbb{R})$. Résoudre l'équation $X + \operatorname{tr}(X)A = B$ d'inconnue $X \in \mathscr{M}_n(\mathbb{R})$.

●●○ **Exercice F**

1. Montrer que, pour toute matrice $A \in \mathrm{GL}_n(\mathbb{R})$, il existe un polynôme $P \in \mathbb{R}[X]$ tel que $A^{-1} = P(A)$.
2. Existe-t-il un polynôme $P \in \mathbb{R}[X]$ tel que, pour toute matrice $A \in \mathrm{GL}_n(\mathbb{R})$, $P(A) = A^{-1}$?

●●● **Exercice G**

Soit $A \in \mathscr{M}_n(\mathbb{R})$ de la forme $U \operatorname{diag}(\alpha_1, \dots, \alpha_n)U^{-1}$ avec $U \in \mathrm{GL}_n(\mathbb{R})$. Montrer qu'il existe $P \in \mathbb{K}[X]$ tel que $A = P(A^3)$.

●●● **Exercice H**

Soit $M \in \mathscr{M}_n(\mathbb{C})$ telle qu'il existe α, $\beta \in \mathbb{C}$ et A, $B \in \mathscr{M}_n(\mathbb{C})$ vérifiant

$$\forall k \in \{1,2,3\}, \qquad M^k = \alpha^k A + \beta^k B.$$

Montrer que, pour tout $k \in \mathbb{N} \setminus \{0\}$, $M^k = \alpha^k A + \beta^k B$.

Problèmes du chapitre 16

●●○ **Problème 1**

Définition 16.40.

Pour toute matrice $N \in \mathscr{M}_n(\mathbb{R})$ nilpotente d'indice p, l'exponentielle de N et le logarithme de $\mathrm{I}_n + N$ sont les matrices définies par

$$\exp N = \sum_{k=0}^{p-1} \frac{1}{k!} N^k, \qquad \ln(\mathrm{I}_n + N) = \sum_{k=1}^{p-1} \frac{(-1)^{k-1}}{k} N^k.$$

1. Justifier que les matrices $N = (n_{i,j})$ suivantes sont nilpotentes puis calculer $\exp N$.
   a. $n_{i,j} = \begin{cases} 1 & \text{si } i = 1 \text{ et } j = n, \\ 0 & \text{sinon.} \end{cases}$
   b. $n_{i,j} = \begin{cases} 1 & \text{si } j = i+1, \\ 0 & \text{sinon.} \end{cases}$
   c. $n_{i,j} = \begin{cases} i & \text{si } j = i+1, \\ 0 & \text{sinon.} \end{cases}$
2. Soit $N \in \mathscr{M}_n(\mathbb{R})$ nilpotente d'indice p.
   a. Montrer que $\ln(\mathrm{I}_n + N)$ et $\exp(N) - \mathrm{I}_n$ sont des matrices nilpotentes.
   b. Considérons les polynômes

   $$P(X) = \sum_{k=1}^{p} \frac{1}{k!} X^k, \qquad Q(X) = \sum_{k=1}^{p} \frac{(-1)^{k-1}}{k} X^k.$$

   Vérifier les développements limites suivants

   $$P(Q(x)) \underset{0}{=} x + o(x^p), \qquad Q(P(x)) \underset{0}{=} x + o(x^p).$$

   c. En déduire que les restes de la division euclidienne par X^p de $P \circ Q$ et de $Q \circ P$ sont égaux à X.

d. Calculer $\ln(\exp N)$ et $\exp(\ln(\mathrm{I}_n + N))$.

3. a. Soit N et $M \in \mathscr{M}_n(\mathbb{R})$ sont deux matrices nilpotentes qui commutent. Montrer que $M+N$ est encore une matrice nilpotente.

b. Soit N et $M \in \mathscr{M}_n(\mathbb{R})$ sont deux matrices nilpotentes qui commutent. Vérifier que

$$\exp(M+N) = \exp(M)\exp(N).$$

c. Montrer que, pour tout $N \in \mathscr{M}_n(\mathbb{R})$ nilpotente, $\exp N$ est inversible puis calculer son inverse.

●○○ **Problème 2**

Soit des entiers $n \geq 3$, δ entre 2 et $n-1$. Considérons une matrice $M = (m_{i,j}) \in \mathscr{M}_n(\mathbb{R})$ vérifiant les conditions suivantes

- M est symétrique ;
- les coefficients diagonaux de M sont nuls ;
- chaque ligne de M compte δ coefficients égaux à 1 et $n-\delta$ coefficients nuls ;
- pour tous indices i et j distincts, $m_{i,j}$ est nul si, et seulement s'il existe un indice k tel que $m_{i,k} = m_{j,k} = 1$; de plus, cet entier k est alors unique.

Notons $J \in \mathscr{M}_n(\mathbb{R})$ la matrice dont tous les coefficients sont égaux à 1.

1. Soit deux indices i et j. Montrer que, si $m_{i,j} = 1$, alors, pour tout k, $m_{i,k}m_{j,k} = 0$.
2. a. Calculer les coefficients de M^2.

   b. Montrer que $M^2 + M = J + \alpha \mathrm{I}_n$ avec $\alpha \in \mathbb{N}$.
3. Soit u et ϕ les endomorphismes canoniquement associés à M et J.

   a. Déterminer $\operatorname{Im}\phi$.

   b. Calculer $\operatorname{Ker}(u - \delta \mathrm{Id})$.

   c. En déduire que $n = \delta^2 + 1$.

PROBLÈMES

CORRIGÉS

Corrigés des Vrai/Faux

a) Faux.

b) Vrai.

c) Vrai.

d) Vrai.

e) Faux, ce sont les matrices scalaires.

f) Vrai.

g) Vrai.

h) Vrai.

i) Vrai.

j) Faux.

k) Faux, par exemple $A = B = \mathrm{I}_n$.

l) Vrai.

m) Faux, dans le cas général, on a juste $\mathrm{tr}(ABC) = \mathrm{tr}(BCA) = \mathrm{tr}(CAB)$.

Corrigés des exercices d'application du chapitre 16

Exercice 1

La matrice $M = (M_{i,j})$ est dans cet espace si, et seulement si,

$$\left\{\begin{array}{lcr} m_{1,2} & = & -m_{1,1} \\ m_{2,2} & = & -m_{2,1} \\ m_{3,1} & = & -m_{1,1} - m_{2,1} \\ m_{3,2} & = & m_{1,1} + m_{2,1} \end{array}\right.$$

Donc l'espace admet pour base $(E_{1,1} - E_{1,2} - E_{3,1} + E_{3,2}, E_{2,1} - E_{2,2} - E_{3,1} + E_{3,2})$ donc est de dimension 2.

Exercice 2

Soit E l'espace considéré. Une matrice A appartient à E si, et seulement si

$$\left\{\begin{array}{lcl} a_{3,3} & = & a_{1,1} + a_{1,2} - a_{2,3} \\ a_{3,1} & = & a_{1,3} + a_{2,3} - a_{3,2} \\ a_{2,1} & = & a_{3,2} + a_{3,3} - a_{1,1} \end{array}\right.$$

soit encore

$$\begin{cases} a_{3,3} &= a_{1,1}+a_{1,2}-a_{2,3} \\ a_{3,1} &= a_{1,3}+a_{2,3}-a_{3,2} \\ a_{2,1} &= a_{3,2}+a_{1,2}-a_{2,3} \end{cases}$$

Ceci est équivalent à dire que A est de la forme suivante

$$a_{2,2}E_{2,2}+a_{1,1}(E_{1,1}+E_{3,3})+a_{1,2}(E_{1,2}+E_{2,1}+E_{3,3})+a_{1,3}(E_{1,3}+E_{3,1})+a_{2,3}(E_{2,3}+E_{3,1}-E_{3,3}-E_{2,1})+a_{3,2}(E_{3,2}+E_{2,1}-E_{3,1})$$

La famille génératrice de E qu'on vient d'obtenir est bien libre : ainsi, $\dim E = 6$.

Exercice 3

Comme $\dim \mathscr{M}_2(\mathbb{R}) = 4$, la famille à cinq éléments $(A, B, C, \mathrm{I}_2, E_{1,2})$ est liée. Ainsi, il existe cinq scalaires non tous nuls λ, μ, ν, α et β tels que

$$\lambda A + \mu B + \nu C + \alpha \mathrm{I}_2 + \beta E_{1,2} = 0_2.$$

Supposons $\lambda = \mu = \nu = 0$, alors $\alpha \mathrm{I}_2 + \beta E_{1,2} = 0_2$ donc $\alpha = \beta = 0$ car les matrices I_2 et $E_{1,2}$ ne sont pas proportionnelles (ou, de manière équivalente, forment une famille libre) : contradiction.

Ainsi, les scalaires λ, μ et ν sont non tous nuls et $\lambda A + \mu B + \nu C \in \mathrm{Vect}(\mathrm{I}_2, E_{1,2})$ donc est triangulaire supérieure à coefficients diagonaux égaux.

Exercice 4

Soit $(i, j) \in [\![1, n]\!]^2$. Alors,

$$\begin{aligned} [JMJ]_{i,j} &= \sum_{k=1}^{n} [J]_{i,k}[MJ]_{k,j} \\ &= \sum_{k=1}^{n} [J]_{i,k} \sum_{l=1}^{n} [M]_{k,l}[J]_{l,j} \\ &= \sum_{k=1}^{n}\sum_{l=1}^{n} [M]_{k,l}. \end{aligned}$$

Ainsi, la matrice JMJ est la matrice dont chaque coefficient est égal à la somme de tous les coefficients de M.

Exercice 5

1. Un simple calcul montre que $A(t)A(s) = A(t+s-2st)$ pour tous s, $t \in \mathbb{R}$.
2. Soit $t \in \mathbb{R}$. La matrice $A(t)$ admet un inverse dans $\mathscr{M}$ si, et seulement s'il existe un paramètre $s \in \mathbb{R}$ tel que $A(t)A(s) = \mathrm{I}_3$, c'est-à-dire, avec la première question, $s + t - 2st = 0$. Cette équation en s admet une solution si, et seulement si, $t \neq \frac{1}{2}$.

Exercice 6

Soit $A \in \mathscr{M}_n(\mathbb{R})$ une telle matrice. Considérons pour $i \neq j$ l'égalité $A(E_{i,j} - E_{j,i}) = (E_{i,j} - E_{j,i})A$.

▹ En regardant la $j^{\text{ième}}$ colonne de ces matrices, on obtient à gauche la $i^{\text{ème}}$ colonne de A soit en effectuant le calcul soit en écrivant

$$C_j(A(E_{i,j} - E_{j,i})) = A(C_j(E_{i,j}) - C_j(E_{j,i})) = AC_j(E_{i,j}) = AC_i(\mathrm{I}_n) = C_i(A\mathrm{I}_n) = C_i(A).$$

À droite, les lignes d'indices autres que i et j sont nulles (car les lignes correspondantes dans la matrice $E_{i,j} - E_{j,i}$ sont nulles).

Ainsi, pour tout $k \notin \{i, j\}$, $[A]_{k,i} = 0$. En faisant varier j et i, on obtient que la matrice A est diagonale.

▹ Déterminons plus précisément le coefficient en position i, j dans cette égalité. On a déjà établi qu'à gauche, on trouve $[A]_{i,i}$. En calculant à droite, on trouve $[A]_{j,j}$ donc $[A]_{i,i} = [A]_{j,j}$. En faisant varier j et i, on obtient que la matrice A est scalaire (c'est-à-dire proportionnelle à l'identité).

Réciproquement, les matrices scalaires commutent avec toutes les matrices donc en particulier avec les matrices antisymétriques.

Exercice 7

L'ensemble des matrices centrosymétriques est non vide, stable par combinaison linéaire et contient la matrice nulle : c'est donc un sous-espace vectoriel de $\mathscr{M}_n(\mathbb{R})$.

Il suffit donc d'obtenir la stabilité par produit. Pour toutes matrices A et B centrosymétriques et tout $(i, j) \in \{1, 2, \dots, n\}$,

$$\begin{aligned}
[AB]_{n+1-i,n+1-j} &= \sum_{k=1}^{n} [A]_{n+1-i,k}[B]_{k,n+1-j} \\
&= \sum_{l=1}^{n} [A]_{n+1-i,n+1-l}[B]_{n+1-l,n+1-j} \\
&= \sum_{l=1}^{n} [A]_{i,l}[B]_{l,j} = [AB]_{i,j}.
\end{aligned}$$

Pour la dimension, on discute selon la parité de n :

- si n est pair, une base est $(E_{i,j} + E_{n+1-i,n+1-j})_{i \leq n, j \leq \frac{n}{2}}$ et la dimension est $\frac{n^2}{2}$;
- si n est impair, une base est constituée de $E_{\frac{n-1}{2},\frac{n-1}{2}}$ et des sommes distinctes de la forme $E_{i,j}+E_{n+1-i,n+1-j}$;

la dimension est donc $1 + \frac{n^2-1}{2} = \frac{n^2+1}{2}$.

Exercice 8

▹ Soit $\lambda \in \mathbb{R}$ et $A \in \mathscr{M}_n(\mathbb{R})$ non nulle vérifiant $A^{\mathrm{T}} = \lambda A$. En transposant, on obtient $A = \lambda A^{\mathrm{T}}$. En remplaçant alors A dans la relation de départ, on obtient $A^{\mathrm{T}} = \lambda^2 A^{\mathrm{T}}$.

La matrice A^{T} est non nulle donc $\lambda^2 = 1$ soit $\lambda \in \{-1, 1\}$.

▹ Réciproquement les valeurs $\lambda = 1$ et $\lambda = -1$ conviennent avec des matrices A non nulles symétriques ou antisymétriques respectivement.

Exercice 9

Commençons par rappeler que $E_{i,j}E_{k,l} = E_{i,l}$ si $j = k$ et 0_n sinon.

▹ Pour tous $i \neq j$, $\varphi(E_{i,j}) = \varphi(E_{i,i}E_{i,j}) = \varphi(E_{i,j}E_{i,i}) = \varphi(0_n) = 0$.

▹ Pour tous $i \neq j$, $\varphi(E_{i,i}) = \varphi(E_{i,j}E_{j,i}) = \varphi(E_{j,i}E_{i,j}) = \varphi(E_{j,j})$.

Fort de ces résultats et avec la linéarité de φ, on obtient pour toute matrice $A = (a_{i,j})$

$$\begin{aligned}
\varphi(A) &= \sum_{i=1}^{n}\sum_{j=1}^{n} a_{i,j}\varphi(E_{i,j}) \\
&= \sum_{i=1}^{n} a_{i,i}\varphi(E_{i,i}) && \text{d'après le premier point} \\
&= \sum_{i=1}^{n} a_{i,i}\varphi(E_{1,1}) && \text{d'après le second point} \\
&= \varphi(E_{1,1})\operatorname{tr}(A).
\end{aligned}$$

Ainsi, $\varphi = \varphi(E_{1,1})\operatorname{tr}$.

Remarque

On aurait pu conclure plus rapidement en indiquant que les formes linéaires φ et $\varphi(E_{1,1})\mathrm{tr}$ coïncident sur la base canonique donc sont égales.

Exercice 10

Proposons deux calculs de A^{-1}.

▹ Soit u l'endomorphisme de $\mathbb{R}_{n-1}[X]$ associé à cette matrice dans la base canonique. Alors, pour tout $k \leq n$, $u(X^{k-1}) = \sum_{j=0}^{k-1} X^j$. Ainsi,

$$u^{-1}(X^{k-1}) = u^{-1}(u(X^{k-1}) - u(X^{k-2})) = X^{k-1} - X^{k-2}.$$

Par conséquent,

$$[A^{-1}]_{i,j} = \begin{cases} 1 & \text{si } i = j, \\ -1 & \text{si } i+1 = j, \\ 0 & \text{sinon.} \end{cases}$$

▹ Considérons $J = (\delta_{j,i+1}) \in \mathscr{M}_n(\mathbb{R})$. Par construction,

$$A = \sum_{k=0}^{n-1} J^k.$$

Comme $J^n = 0_n$, $(\mathrm{I}_n - J)A = \mathrm{I}_n$ et donc $A^{-1} = \mathrm{I}_n - J$.

Exercice 11

▹ Notons $S_\theta = A - \mathrm{I}_2$ et remarquons que $S_\theta^2 = \mathrm{I}_2$. Ainsi, par une récurrence immédiate, pour tout $k \in \mathbb{N}$, $S_\theta^{2k} = \mathrm{I}_2$ et $S_\theta^{2k+1} = S_\theta$.

▹ Comme S_θ et I_2 commutent, on peut calculer les puissances de A par la formule du binôme. Pour tout $k \in \mathbb{N} \setminus \{0\}$,

$$\begin{aligned} A^k &= \sum_{l=0}^{k} \binom{k}{l} S_\theta^l \\ &= \sum_{\substack{l=0 \\ \text{impair}}}^{k} \binom{k}{l} S_\theta + \sum_{\substack{l=0 \\ \text{pair}}}^{k} \binom{k}{l} \mathrm{I}_2 \\ &= 2^{k-1} S_\theta + 2^{k-1} \mathrm{I}_2. \end{aligned}$$

Exercice 12

Soit $(e_1, \ldots, e_n)$ la base canonique de $\mathbb{R}^n$ et u l'endomorphisme canoniquement associé à M. Par définition de la matrice, $u(e_j) = e_{j-1[n]}$ pour tout $j \in [\![1, n]\!]$.

Par une récurrence immédiate sur $k \in \mathbb{N}$, $u^k(e_j) = e_{j-k[n]}$ pour tout $j \in [\![1, n]\!]$. Par conséquent, la matrice M^k admet pour coefficient en position (i, j)

$$[M^k]_{i,j} = \begin{cases} 1 & \text{si } i = j - k[n], \\ 0 & \text{sinon.} \end{cases}$$

Exercice 13

La matrice LC est de taille $1,1$ donc admet un unique coefficient $x = \sum_{k=1}^{n}[C]_k[L]_k$.
Par conséquent, pour tout $k \in \mathbb{N}$, $(LC)^k$ est la matrice $1,1$ dont l'unique coefficient est x^k.
Avec l'associativité du produit, $(CL)^k = C(LC)^{k-1}L = x^{k-1}CL$ pour tout $k \in \mathbb{N} \setminus \{0\}$.

Exercice 14

1. Soit $M \in \operatorname{Ker}\Psi$. On trouve rapidement $C_1(MA) = \sum_{k=1}^{n} \alpha_k C_k(M)$ et pour tout $k \geq 2$, $C_k(MA) = C_{k-1}(M)$.
Montrons par récurrence sur $k \in [\![0, n-1]\!]$ que $C_j(M) = 0_{n,1}$ pour $j \in [\![n-k, n]\!]$.
- L'initialisation vient de l'hypothèse $M \in \operatorname{Ker}\Psi$.
- Soit $k \in [\![0, n-2]\!]$ tel que $C_j(M) = 0_{n,1}$ pour $j \in [\![n-k, n]\!]$. Alors, pour tout $j \in [\![n-k, n]\!]$, $C_j(AM) = 0_{n,1}$ mais alors

$$C_{n-k-1}(M) = C_{n-k}(MA) = C_{n-k}(AM) = 0_{n,1}.$$

La propriété au rang n indique que $M = 0_n$. En conclusion, l'endomorphisme Ψ est injectif.
2. D'après la première question, $\dim \mathscr{C}_A \leq n$. Par ailleurs, $\{P(A),\ P \in \mathbb{R}_{n-1}[X]\} \subset \mathscr{C}_A$.
Or, $(E_n, AE_n, \ldots, A^{n-1}E_n)$ est une base de $\mathscr{M}_{n,1}(\mathbb{R})$ donc A n'admet pas de polynôme annulateur non nul de degré strictement inférieur à n : en conclusion, $\dim\{P(A),\ P \in \mathbb{R}_{n-1}[X]\} = n$ et $\{P(A),\ P \in \mathbb{R}_{n-1}[X]\} = \mathscr{C}_A$.

Corrigés des exercices d'approfondissement du chapitre 16

Exercice A

Soit F l'ensemble de ces matrices.
▹ L'ensemble F est non vide, stable par combinaison linéaire : si $M, N \in F$ et $\lambda \in \mathbb{R}$, alors, pour tout (i, j) tel que $j = i + 1[2]$,

$$[M + \lambda N]_{i,j} = [M]_{i,j} + \lambda[N]_{i,j} = 0,$$

donc $M + \lambda N \in F$: F est un sous-espace vectoriel de $\mathscr{M}_n(\mathbb{R})$.
▹ Montrons la stabilité par produit. Pour toutes matrices A et $B \in F$ et pour tout (i, j) tel que $j = i + 1[2]$,

$$[AB]_{i,j} = \sum_{k=1}^{n}[A]_{i,k}[B]_{k,j} = 0,$$

car soit i et k sont de parité différente (et alors $[A]_{i,k} = 0$), soit j et k sont de parité différente (et alors $[B]_{k,j} = 0$). Ainsi, $AB \in F$.
▹ Pour la dimension de F, on discute selon la parité de n en dénombrant la base de F composée des $E_{i,j}$ avec $i = j[2]$.
- Si n est pair, la dimension est donc $\frac{n^2}{2}$;
- si n est impair, la dimension est donc $\frac{n^2+1}{2}$.

Exercice B

1. Il suffit de faire le calcul. Avec $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, on obtient

$$M^2 - (\operatorname{tr} M)M = \begin{bmatrix} a^2 + bc & ab + bd \\ ac + dc & bc + d^2 \end{bmatrix} - (a+d)\begin{bmatrix} a & b \\ c & d \end{bmatrix} = (bc - ad)\mathrm{I}_2$$

2. Comme $\operatorname{tr}(AB-BA)=\operatorname{tr}(AB)-\operatorname{tr}(BA)=0$, la formule de la question précédente donne

$$(AB-BA)^2 \in \operatorname{Vect}(\mathrm{I}_2).$$

Exercice C

1. Calculons $(AB)^{p+1}$ en utilisant l'associativité du produit matriciel :

$$(AB)^{p+1}=(AB)(AB)\dots(AB)=A(BA)^pB=A0_nB=0_n.$$

2. La réponse est positive : on peut trouver des matrices A et B telles qu'il existe un entier p vérifiant $(AB)^p=0_n$ et $(BA)^p\neq 0_n$; par exemple $p=1$ et les matrices

$$A=\begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}, \qquad B=\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}.$$

Exercice D

1. Si M est solution de (E), alors D est un polynôme en M donc commute avec M.
2. Les coefficients de MD et DM sont respectivement $(\lambda_j m_{i,j})$ et $(\lambda_i m_{i,j})$: comme ces deux matrices sont égales et que les coefficients $(\lambda_i)_i$ sont deux à deux distincts, on en déduit que $m_{i,j}=0$ dès que $i\neq j$, c'est-à-dire que M est diagonale.
3. Si $M=\operatorname{diag}(a_1,\dots,a_n)$, alors $M^{l+1}-M^l=\operatorname{diag}(a_1^{l+1}-a_1^l,\dots,a_n^{l+1}-a_n^l)$. Par conséquent, (E) admet des solutions si, et seulement si, $\lambda_k\in f(\mathbb{R})$ où $f : x\mapsto x^{l+1}-x^l$. Une rapide étude de cette dernière fonction donne le tableau de variations suivant :

	$-\infty$		$\frac{l}{l+1}$		$+\infty$
f	$+\infty$	$\searrow$	$-\frac{1}{l+1}\left(\frac{l}{l+1}\right)^l$	$\nearrow$	$+\infty$

Donc $f(\mathbb{R})=[-\frac{1}{l+1}\left(\frac{l}{l+1}\right)^l,+\infty[$ d'où la condition annoncée.

4. D'après le tableau de variations, chaque équation $f(x)=\lambda_k$ admet deux solutions si $\lambda_k>-\frac{1}{l+1}\left(\frac{l}{l+1}\right)^l$, une si $\lambda_k=-\frac{1}{l+1}\left(\frac{l}{l+1}\right)^l$. Il y a 2^n solutions, 2^{n-1} si $-\frac{1}{l+1}\left(\frac{l}{l+1}\right)^l$ est un terme de la matrice diagonale D.

Exercice E

Commençons par passer cette équation à la trace : $\operatorname{tr}X(1+\operatorname{tr}A)=\operatorname{tr}B$.

- Si $\operatorname{tr}A\neq -1$, alors $\operatorname{tr}X=\frac{\operatorname{tr}B}{1+\operatorname{tr}A}$ donc la seule solution est (après vérification immédiate)

$$B-\frac{\operatorname{tr}B}{1+\operatorname{tr}A}A.$$

- Si $\operatorname{tr}A=-1$,
  - si $\operatorname{tr}B\neq 0$, il n'y a pas de solution.
  - sinon, on ne dispose d'aucune information sur $\operatorname{tr}X$. L'équation indique seulement que les solutions sont de la forme $B+\lambda A$ avec $\lambda\in\mathbb{R}$. Reportons dans l'équation

$$B+\lambda A+\operatorname{tr}(B+\lambda A)A=B+\lambda A+\lambda\operatorname{tr}AA=B.$$

En conclusion, les matrices $B+\lambda A$ avec $\lambda\in\mathbb{R}$ sont les solutions.

Exercice F

1. Soit $A \in \mathrm{GL}_n(\mathbb{R})$. Il existe un polynôme annulateur Q de A non nul (par exemple, en remarquant que la famille $(\mathrm{I}_n, A, \dots, A^{n^2})$ est liée). Écrivons alors Q de la forme $X^k R(X)$ avec k sa valuation et R un polynôme tel que $R(0) \neq 0$. Comme A est inversible, A^k l'est également et la condition $Q(A) = 0_n$ entraîne $R(A) = 0_n$. Notons $R = \sum_{k=0}^{r} \alpha_k X^k$ et isolons le terme d'indice 0 dans l'égalité matricielle précédente :

$$\alpha_0 \mathrm{I}_n = -\sum_{k=1}^{r} \alpha_k A^k$$

et par conséquent, en multipliant par A^{-1}, on obtient la forme recherchée

$$A^{-1} = -\sum_{k=1}^{r} \frac{\alpha_k}{\alpha_0} A^{k-1}.$$

2. La réponse est non. Si un tel polynôme P existait, on aurait, pour tout $\lambda \in \mathbb{R}^*$, $P(\lambda \mathrm{I}_n) = \lambda^{-1} \mathrm{I}_n$ donc $\lambda P(\lambda) = 1$. Comme cette identité est vraie pour une infinité de valeurs, on obtient l'égalité entre polynôme $XP(X) = 1$ qui donne une contradiction lorsque l'on évalue en 0.

Exercice G

Comme vu dans le cours, $A^3 = U \operatorname{diag}(\alpha_1^3, \dots, \alpha_n^3) U^{-1}$.

Considérons le polynôme P d'interpolation de Lagrange qui vaut, pour tout $k \in [\![1, n]\!]$, α_k en α_k^3 (ce qui est possible car $\alpha_k^3 = \alpha_j^3$ entraîne $\alpha_k = \alpha_j$). Alors, $\operatorname{diag}(\alpha_1, \dots, \alpha_n) = P(\operatorname{diag}(\alpha_1^3, \dots, \alpha_n^3))$ et donc

$$\begin{aligned} A &= U \operatorname{diag}(\alpha_1, \dots, \alpha_n) U^{-1} \\ &= U P(\operatorname{diag}(\alpha_1^3, \dots, \alpha_n^3)) U^{-1} \\ &= P(U \operatorname{diag}(\alpha_1^3, \dots, \alpha_n^3) U^{-1}) = P(A^3). \end{aligned}$$

Exercice H

▹ Soit $P(X) = X(X-\alpha)(X-\beta)$. Commençons par remarquer $P(M) = P(\alpha)A + P(\beta)B = 0_n$ c'est-à-dire que P est annulateur de M.

▹ Pour tout $k \in \mathbb{N} \setminus \{0\}$, il existe $Q_k \in \mathbb{C}[X]$, a_k, b_k et $c_k \in \mathbb{C}$ tels que $X^k = Q_k.P + a_k X^2 + b_k X + c_k$. En évaluant en 0, α et β, on trouve le système

$$\left\{ \begin{array}{rcl} c_k & = & 0 \\ a_k \alpha^2 + b_k \alpha & = & \alpha^k \\ a_k \beta^2 + b_k \beta & = & \beta^k \end{array} \right.$$

Par conséquent,

$$\begin{aligned} M^k &= a_k M^2 + b_k M + c_k \mathrm{I}_n \\ &= a_k(\alpha^2 A + \beta^2 B) + b_k(\alpha A + \beta B) \\ &= (a_k \alpha^2 + b_k \alpha) A + (a_k \beta^2 + b_k \beta) B \\ &= \alpha^k A + \beta^k B. \end{aligned}$$

Corrigés des problèmes du chapitre 16

Problème 1

1. a. Cette matrice N est la matrice $E_{1,n}$ de la base canonique de $\mathscr{M}_n(\mathbb{R})$: elle est triangulaire supérieure stricte donc nilpotente. Plus précisément, $N^2 = 0_n$ d'où $\exp N = \mathrm{I}_n + N$.
b. La matrice N est triangulaire supérieure stricte et d'indice de nilpotence n :

$$\begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 1 \\ 0 & \dots & \dots & \dots & 0 \end{bmatrix}.$$

On montre par récurrence que les coefficients non nuls de N^k sont égaux à 1 pour $j = i + k$ (c'est-à-dire sur la $k^{\text{ième}}$ surdiagonale).
Par conséquent, les coefficients de $\exp N$ sont $\frac{1}{(j-i)!}$ si $j \geq i$, 0 sinon.
c. La matrice N est triangulaire supérieure stricte et d'indice de nilpotence n :

$$\begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & & \ddots & n-1 \\ 0 & \dots & \dots & \dots & 0 \end{bmatrix}.$$

On montre par récurrence que les coefficients non nuls de N^k sont égaux à $\frac{(i+k-1)!}{(i-1)!}$ pour $j = i + k$.
Par conséquent, $\exp N$ est triangulaire supérieure et le coefficient en position (i, j) est $\binom{j-1}{i-1}$.
2. a. Par définition, les matrices $\ln(\mathrm{I}_n + N)$, $\exp N - \mathrm{I}_n$ s'écrivent toutes les deux sous la forme NA (en factorisant les sommes par N) avec A un polynôme en N. Comme A et N commutent, la puissance $p^{\text{ième}}$ vaut $A^p N^p = 0_n$.
b. Remarquons que $e^x = 1 + P(x) + o(x^p)$ et $\ln(1 + x) = Q(x) + o(x^p)$.
En composant les développements limités, on obtient

$$1 + x = e^{\ln(1+x)} \underset{0}{=} 1 + P(Q(x)) + o(x^p) \qquad x = \ln e^x \underset{0}{=} Q(P(x)) + o(x^p).$$

c. Comme $P \circ Q$ et $Q \circ P$ sont des polynômes, l'unicité de la partie régulière dans ces développements limités indique que la troncature de $1 + P(Q(X))$ à l'ordre p est $1 + X$ et la troncature de $Q(P(X))$ à l'ordre p est X.
En d'autres termes, les polynômes $P(Q(X)) - X$ et $Q(P(X)) - X$ sont divisibles par X^p.
d. Utilisons le résultat précédent : il existe $A \in \mathbb{R}[X]$ tel que $P(Q(X)) = X + X^p A(X)$ soit

$$P(Q(N)) = N + N^p A(N) = N$$

car $N^p = 0_n$. Ainsi,

$$\ln(\exp N) = Q(P(N)) = N.$$

De même,

$$\exp(\ln(\mathrm{I}_n + N)) = (1 + P) \circ Q(N) = \mathrm{I}_n + N.$$

CORRIGÉS

3. a. Si N et M sont deux matrices nilpotentes d'indice p et q, alors $M+N$ est une matrice nilpotente d'indice au plus $p+q-1$; en effet, la formule du binôme donne

$$\begin{aligned}(N+M)^{p+q-1} &= \sum_{j=0}^{p+q-1}\binom{p+q-1}{j}N^jM^{p+q-1-j}\\ &= \sum_{j=0}^{q-1}\binom{p+q-1}{j}N^j\underbrace{M^{p+q-1-j}}_{=0_n}+\sum_{j=q}^{p+q-1}\binom{p+q-1}{j}\underbrace{N^j}_{=0_n}M^{p+q-1-j}\\ &= 0_n\end{aligned}$$

car si $j\geq q$, alors $N^j=0_n$ et si $j<q$, alors $p+q-1-j\geq p$ et $M^{p+q-1-j}=0_n$.
Le formule du binôme donne aussi

$$\begin{aligned}\exp(M+N) &= \sum_{k=0}^{p+q-1}\frac{1}{k!}(M+N)^k=\sum_{k=0}^{p+q-1}\frac{1}{k!}\sum_{j=0}^{k}\binom{k}{j}M^jN^{k-j}\\ &= \sum_{k=0}^{p+q-1}\sum_{j=0}^{k}\frac{1}{j!}M^j\frac{1}{(k-j)!}N^{k-j}=\exp(M)\exp(N).\end{aligned}$$

b. Les matrices N et $-N$ sont nilpotentes et commutent donc, d'après la question précédente, $\exp(N)\exp(-N)=\exp(-N)\exp(N)=\mathrm{I}_n$. Par conséquent, $\exp N\in\mathrm{GL}_n(\mathbb{R})$ et $(\exp N)^{-1}=\exp(-N)$.

Problème 2

1. D'après la dernière propriété, si $m_{i,j}=1$ alors, pour tout indice k, $m_{i,k}=0$ ou $m_{j,k}=0$ c'est-à-dire si $m_{i,k}m_{j,k}=0$.
2. Notons $M^2=(a_{i,j})$.
a. ▹ Commençons par déterminer les termes diagonaux. Pour tout i,

$$a_{i,i}=\sum_{k=1}^{n}m_{i,k}m_{k,i}=\sum_{k=1}^{n}m_{i,k}^2=\delta$$

car la matrice M est symétrique et possède δ termes non-nuls égaux à 1 par ligne.
▹ Pour tous $i\neq j$, on utilise à nouveau que la matrice M est symétrique pour obtenir

$$a_{i,j}=\sum_{k=1}^{n}m_{i,k}m_{k,j}=\sum_{k=1}^{n}m_{i,k}m_{j,k}.$$

D'après la question précédente, si $m_{i,j}=0$, $a_{i,j}=1$ et si $m_{i,j}=1$, $a_{i,j}=0$.
b. Les coefficients diagonaux de M^2+M sont tous égaux à δ, tous les autres coefficients sont égaux à 1. D'où $M^2+M=J+(\delta-1)\mathrm{I}_n$.
3. a. L'endomorphisme ϕ envoie tous les vecteurs de la base canonique sur le vecteur v dont les coordonnées sont toutes égales à 1 : par conséquent, $\operatorname{Im}\phi=\operatorname{Vect}(v)$.
b. ▹ Soit X le vecteur colonne de $x\in\operatorname{Ker}(u-\delta\mathrm{Id})$. Comme $u(x)=\delta x$, on obtient $MX=\delta X$. D'après la question 2, $(\delta^2+1)X=(M^2+M+(1-\delta)\mathrm{I}_n)X=JX$ d'où $X=J(\frac{1}{\delta^2+1}X)$ et $x\in\operatorname{Im}\phi=\operatorname{Vect}(v)$. On a donc $\operatorname{Ker}(u-\delta\mathrm{Id})\subset\operatorname{Vect}(v)$.
▹ L'inclusion réciproque est une simple vérification : la matrice M admet δ coefficients égaux à 1 par ligne donc $MV=\delta V$ donc $u(v)=\delta v$.
c. En appliquant la relation $M^2+M=J+(\delta-1)\mathrm{I}_n$ au vecteur colonne du vecteur v, on obtient

$$\delta^2v+\delta v=nv+(\delta-1)v$$

soit $n=\delta^2+1$.

Chapitre 17

Équivalence des matrices et systèmes linéaires

1. Matrices équivalentes — **2.** Équivalence en pratique — **3.** Applications aux systèmes linéaires — **4.** Introduction à la similitude

Objectifs et compétences du programme

- Appliquer les formules de changement de bases.
- Calculer le rang en échelonnant une matrice.
- Calculer image et noyau en échelonnant une matrice.
- Résoudre un système linéaire par la méthode du pivot de Gauss.
- Comprendre équivalence et similitude des matrices.

Gabriel Cramer, mathématicien suisse (1704-1752)
Dès l'âge de vingt ans, il devient titulaire de la chaire de mathématiques à l'académie de Genève. Cette responsabilité ne l'empêche pas de faire un tour d'Europe et de rencontrer de nombreux autres scientifiques comme les Bernoulli, Buffon, Clairaut, Euler ou D'Alembert.
En 1750 à Genève, il publie son unique ouvrage *Introduction à l'analyse des lignes courbes algébriques.* C'est pour cette étude qu'il s'intéresse aux méthodes de résolution des systèmes linéaires. On trouve dans ces pages les formules désormais appelées de Cramer qui reposent sur l'utilisation des déterminants (même si cette dernière notion n'a pas encore été développée sous ce nom).

COURS

Ce chapitre aborde une même notion sous plusieurs angles différents. Partant de la formule de changement de bases qui permet de relier les matrices d'une même application linéaire relativement à différentes bases, on exhibe la notion de matrices équivalentes. La relation d'équivalence est alors étudiée en détail et caractérisée par un critère simple : l'égalité du rang.
Un deuxième temps dans cette étude consiste à fournir un algorithme simple (le pivot de Gauss) permettant de calculer le rang : il s'agit, étape après étape, de se ramener à une matrice équivalente échelonnée. Outre le rang ce procédé permet de calculer l'image, le noyau d'une matrice et, lorsqu'il existe, son inverse.
Cet algorithme correspond exactement à la méthode de résolution développée pour les systèmes linéaires qui sont l'objet de la troisième section.
On termine ce chapitre avec l'étude partielle d'une relation plus fine sur les matrices carrées : la similitude.

1. Matrices équivalentes

1.1. Formulaire de changement de bases

Définition 17.1.

Soit E un espace vectoriel de dimension finie, deux bases $\mathscr{B}$ et $\mathscr{B}'$ de E. La matrice de passage de $\mathscr{B}$ à $\mathscr{B}'$ est la matrice $\mathrm{mat}_{\mathscr{B}',\mathscr{B}}(\mathrm{Id}_E) \in \mathscr{M}_n(\mathbb{K})$. On la note $P_{\mathscr{B}}^{\mathscr{B}'}$.
En d'autres termes, en notant $\mathscr{B} = (e_1, e_2, \dots, e_n)$, $\mathscr{B}' = (e'_1, e'_2, \dots, e'_n)$, la matrice de passage de $\mathscr{B}$ à $\mathscr{B}'$ est donnée par

$$P_{\mathscr{B}}^{\mathscr{B}'} = \begin{array}{c} \begin{array}{cccc} e'_1 & e'_2 & \cdots & e'_n \\ \downarrow & \downarrow & & \downarrow \end{array} \\ \left[\begin{array}{cccc} a_{1,1} & a_{1,2} & \cdots & a_{1,p} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,p} \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,p} \end{array}\right] \end{array} \begin{array}{c} \\ \\ \to e_1 \\ \to e_2 \\ \vdots \\ \vdots \\ \to e_n \end{array}$$

La définition permet d'obtenir le résultat suivant d'inversibilité des matrices de passages.

Proposition 17.2.

Soit E un espace vectoriel de dimension finie, $\mathscr{B}$ et $\mathscr{B}'$ deux bases de E. Alors, la matrice de passage $P_{\mathscr{B}}^{\mathscr{B}'}$ est inversible et son inverse est la matrice de passage $P_{\mathscr{B}'}^{\mathscr{B}}$.

Démonstration
On utilise l'interprétation du produit matriciel comme la composition des applications linéaires

$$\mathrm{mat}_{\mathscr{B}',\mathscr{B}}(\mathrm{Id}_E)\,\mathrm{mat}_{\mathscr{B},\mathscr{B}'}(\mathrm{Id}_E) = \mathrm{mat}_{\mathscr{B},\mathscr{B}}(\mathrm{Id}_E) = \mathrm{I}_n,$$

donc $P_{\mathscr{B}}^{\mathscr{B}'} P_{\mathscr{B}'}^{\mathscr{B}} = \mathrm{I}_n$. ■

Notons $\mathscr{B}=(e_1,\dots,e_n)$ et $\mathscr{B}'=(e'_1,\dots,e'_n)$. Par définition, la matrice de passage de $\mathscr{B}$ à $\mathscr{B}'$ est la matrice dont la $j^{\text{ième}}$ colonne est $\mathrm{mat}_{\mathscr{B}}(e'_j)$ pour tout $j\in[\![1,n]\!]$.

Exemple

Considérons $\mathbb{R}^3$ muni de la base canonique $\mathscr{B}$ et définissons $\mathscr{B}'=(e'_1,e'_2,e'_3)$ avec

$$\begin{cases} e'_1 &= (1,2,3)\\ e'_2 &= (4,5,6)\\ e'_3 &= (0,7,8).\end{cases}$$

On vérifie rapidement que la famille $\mathscr{B}'$ est libre donc que c'est une base car son cardinal est $3=\dim\mathbb{R}^3$. La matrice de passage de $\mathscr{B}$ à $\mathscr{B}'$ est alors, par définition,

$$P_{\mathscr{B}}^{\mathscr{B}'}=\mathrm{mat}_{\mathscr{B}',\mathscr{B}}(\mathrm{Id}_E)=\begin{bmatrix}1&4&0\\2&5&7\\3&6&8\end{bmatrix}.$$

Proposition 17.3.

Soit E un espace vectoriel de dimension finie, $\mathscr{B}$ et $\mathscr{B}'$ deux bases de E. Alors, pour tout $x\in E$,

$$\mathrm{mat}_{\mathscr{B}}(x)=P_{\mathscr{B}}^{\mathscr{B}'}\,\mathrm{mat}_{\mathscr{B}'}(x).$$

Démonstration

On utilise l'interprétation du produit entre une matrice et un vecteur colonne

$$P_{\mathscr{B}}^{\mathscr{B}'}\,\mathrm{mat}_{\mathscr{B}'}(x)=\mathrm{mat}_{\mathscr{B}',\mathscr{B}}(\mathrm{Id}_E)\,\mathrm{mat}_{\mathscr{B}'}(x)=\mathrm{mat}_{\mathscr{B}}(\mathrm{Id}_E(x))=\mathrm{mat}_{\mathscr{B}}(x).$$

■

Proposition 17.4.

Soit E, F deux espaces vectoriels de dimension finie, $\mathscr{B}$ et $\mathscr{B}'$ deux bases de E, $\mathscr{C}$ et $\mathscr{C}'$ deux bases de F. Alors, pour tout $u\in\mathscr{L}(E,F)$,

$$\mathrm{mat}_{\mathscr{B}',\mathscr{C}'}(u)=(P_{\mathscr{C}}^{\mathscr{C}'})^{-1}\,\mathrm{mat}_{\mathscr{B},\mathscr{C}}(u)P_{\mathscr{B}}^{\mathscr{B}'}.$$

Démonstration

On utilise l'interprétation du produit matriciel en terme de composition ; illustrons la démarche avec les deux parcours joignant $(E,\mathscr{B}')$ et $(F,\mathscr{C}')$ sur le diagramme suivant

$$\begin{tikzcd}
(E,\mathscr{B}') \arrow[r, "u"] \arrow[d, "\mathrm{Id}_E"'] & (F,\mathscr{C}') \\
(E,\mathscr{B}) \arrow[r, "u"] & (F,\mathscr{C}) \arrow[u, "\mathrm{Id}_F"']
\end{tikzcd}$$

$$\begin{aligned}(P_{\mathscr{C}}^{\mathscr{C}'})^{-1}\,\mathrm{mat}_{\mathscr{B},\mathscr{C}}(u)P_{\mathscr{B}}^{\mathscr{B}'}&=\mathrm{mat}_{\mathscr{C},\mathscr{C}'}(\mathrm{Id}_F)\,\mathrm{mat}_{\mathscr{B},\mathscr{C}}(u)\,\mathrm{mat}_{\mathscr{B}',\mathscr{B}}(\mathrm{Id}_E)\\&=\mathrm{mat}_{\mathscr{B}',\mathscr{C}'}(\mathrm{Id}_F\circ u\circ\mathrm{Id}_E)=\mathrm{mat}_{\mathscr{B}',\mathscr{C}'}(u).\end{aligned}$$

■

Corollaire 17.5.

Soit E un espace vectoriel de dimension finie, $\mathscr{B}$ et $\mathscr{B}'$ deux bases de E. Alors, pour tout $u \in \mathscr{L}(E)$,

$$\mathrm{mat}_{\mathscr{B}'}(u) = (P_{\mathscr{B}}^{\mathscr{B}'})^{-1}\,\mathrm{mat}_{\mathscr{B}}(u) P_{\mathscr{B}}^{\mathscr{B}'}.$$

On retient souvent les formules de changement de bases sous la forme $X = PX'$ et $M' = Q^{-1}MP$ mais il faut impérativement savoir ce que signifie chacun de ces termes.

Exemple

Considérons u l'endomorphisme canoniquement associé à la matrice

$$\begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{bmatrix}$$

et cherchons l'expression de sa matrice dans la base $\mathscr{B}' = (e_1', e_2', e_3')$ avec

$$\begin{cases} e_1' &= (1,-1,0) \\ e_2' &= (1,0,-1) \\ e_3' &= (1,1,1). \end{cases}$$

La matrice de passage de $\mathscr{B}$ à $\mathscr{B}'$ est alors, par définition,

$$P_{\mathscr{B}}^{\mathscr{B}'} = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \\ 0 & -1 & 1 \end{bmatrix}$$

et son inverse est (après calculs)

$$(P_{\mathscr{B}}^{\mathscr{B}'})^{-1} = \frac{1}{3}\begin{bmatrix} 1 & -2 & 1 \\ 1 & 1 & -2 \\ 1 & 1 & 1 \end{bmatrix}.$$

Ainsi, la matrice de u dans la base $\mathscr{B}'$ est

$$\frac{1}{3}\begin{bmatrix} 1 & -2 & 1 \\ 1 & 1 & -2 \\ 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \\ 0 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

On peut, en effet, vérifier que l'image du vecteur $e_1 + e_2 + e_3$ est bien $-e_1 - e_2 - e_3$.

1.2. Équivalence et rang

Définition 17.6.

▹ L'image d'une matrice $A \in \mathscr{M}_{n,p}(\mathbb{K})$ est le sous-espace vectoriel de $\mathscr{M}_{n,1}(\mathbb{K})$ défini par

$$\mathrm{Im}\,A = \mathrm{Vect}(C_1(A), \dots, C_p(A)).$$

Le rang de la matrice A, noté $\mathrm{rg}\,A$, est la dimension de $\mathrm{Im}\,A$. Autrement dit, $\mathrm{rg}\,A$ est le nombre maximal de colonnes de A linéairement indépendantes.

▹ Le noyau d'une matrice $A \in \mathscr{M}_{n,p}(\mathbb{K})$ est le sous-espace vectoriel de $\mathscr{M}_{p,1}(\mathbb{K})$ défini par

$$\mathrm{Ker}\,A = \{X \in \mathscr{M}_{p,1}(\mathbb{K}),\ AX = 0_n\}.$$

Proposition 17.7.

Le rang d'une matrice est le rang de n'importe quelle application linéaire associée.

Démonstration

Soit A une matrice et $u \in \mathscr{L}(E,F)$ l'application linéaire associée à A dans les bases $\mathscr{B} = (e_1, \dots, e_p)$ de E et $\mathscr{B}' = (f_1, \dots, f_n)$ de F. Comme l'application $\left\{ \begin{array}{ccc} F & \to & \mathscr{M}_{n,1}(\mathbb{K}) \\ y & \mapsto & \mathrm{mat}_{\mathscr{B}'}(y) \end{array} \right.$ est un isomorphisme,

$$\mathrm{rg}(u(e_1), \dots, u(e_p)) = \mathrm{rg}(C_1(A), \dots, C_p(A))$$

donc $\mathrm{rg}\, u = \mathrm{rg}\, A$. ■

Remarque

Une fois les bases $\mathscr{B}$ et $\mathscr{C}$ fixées, on peut dresser un parallèle plus précis entre une application linéaire $u \in \mathscr{L}(E,F)$ et sa matrice $A = \mathrm{mat}_{\mathscr{B},\mathscr{C}}(u)$.

$y = u(x)$	$Y = AX$
$x \in \mathrm{Ker}\, u$ $y \in \mathrm{Im}\, u$	$X \in \mathrm{Ker}\, A$ $Y \in \mathrm{Im}\, A$
$\mathrm{rg}\, u$	$\mathrm{rg}\, A$

en notant $X = \mathrm{mat}_{\mathscr{B}}(x)$ et $Y = \mathrm{mat}_{\mathscr{C}}(y)$.

Des résultats sur les applications linéaires, on déduit les résultats suivants sur les matrices.

Proposition 17.8. Formule du rang

Pour tout $A \in \mathscr{M}_{n,p}(\mathbb{K})$,

$$\mathrm{rg}\, A + \dim \mathrm{Ker}\, A = p.$$

En particulier, $\mathrm{rg}\, A \leq \min(n,p)$.

Proposition 17.9.

Soit $A \in \mathscr{M}_n(\mathbb{K})$. Les trois propriétés suivantes sont équivalentes

- $A \in \mathrm{GL}_n(\mathbb{K})$,
- $\mathrm{rg}\, A = n$,
- $\mathrm{Ker}\, A = \{0_{n,1}\}$.

Proposition 17.10.

Le rang est invariant par produit par matrices inversibles. Plus précisément, si $A \in \mathscr{M}_{n,p}(\mathbb{K})$, $U \in \mathrm{GL}_n(\mathbb{K})$ et $V \in \mathrm{GL}_p(\mathbb{K})$, alors $\mathrm{rg}(UAV) = \mathrm{rg}\, A$.

Proposition 17.11.

Soit $A \in \mathscr{M}_{n,p}(\mathbb{K})$ de rang r. Alors, il existe $U \in \mathrm{GL}_n(\mathbb{K})$ et $V \in \mathrm{GL}_p(\mathbb{K})$ telles que

$$UAV = J_r = \begin{bmatrix} \mathrm{I}_r & 0 \\ 0 & 0 \end{bmatrix}.$$

Démonstration
Soit $u \in \mathscr{L}(E,F)$ l'application linéaire canoniquement associée. Soit $(e_{r+1},\dots,e_p)$ une base de $\operatorname{Ker} u$ que l'on complète en une base $\mathscr{B}=(e_1,\dots,e_p)$ de E. La famille $(u(e_1),\dots,u(e_r))$ est de rang r (car elle engendre $\operatorname{Im} u$) donc libre : complétons-la en une base $\mathscr{C}$ de F. Dans les bases $\mathscr{B}$ et $\mathscr{C}$, la matrice de u est J_r. D'après la formule de changement de bases, il existe des matrices de passage $U \in \mathrm{GL}_n(\mathbb{K})$ et $V \in \mathrm{GL}_p(\mathbb{K})$ telles que $UAV = J_r$. ■

Remarque
La réciproque est une simple conséquence de la proposition précédente.

Corollaire 17.12.

Pour tout $A \in \mathscr{M}_{n,p}(\mathbb{K})$, $\operatorname{rg} A^{\mathrm{T}} = \operatorname{rg} A$.

Démonstration
Soit $r = \operatorname{rg} A$. D'après la proposition précédente, il existe $U \in \mathrm{GL}_n(\mathbb{K})$ et $V \in \mathrm{GL}_p(\mathbb{K})$ telles que $UAV = J_r$ donc $V^{\mathrm{T}}A^{\mathrm{T}}U^{\mathrm{T}} = {J_r}^{\mathrm{T}}$ qui est encore de la forme J_r mais de taille p,n. D'après la caractérisation précédente, $\operatorname{rg} A^{\mathrm{T}} = r$. ■

Pour reformuler ces propriétés, précisons une définition.

Définition 17.13.

Deux matrices A, $B \in \mathscr{M}_{n,p}(\mathbb{K})$ sont équivalentes s'il existe $P \in \mathrm{GL}_n(\mathbb{K})$ et $Q \in \mathrm{GL}_p(\mathbb{K})$ telles que $B = PAQ$.

Voici donc la caractérisation des classes d'équivalence démontrée précédemment.

Proposition 17.14.

Deux matrices sont équivalentes si, et seulement si, elles ont même rang.

1.3. Matrices de rang donné

Proposition 17.15.

Soit $M \in \mathscr{M}_n(\mathbb{K})$ de rang 1. Alors, il existe des colonnes X, $Y \in \mathscr{M}_{n,1}(\mathbb{K})$ toutes deux non nulles telles que $M = X^{\mathrm{T}}Y$.

Démonstration
Notons $X \in \mathscr{M}_{n,1}(\mathbb{K})$ un vecteur-colonne qui engendre $\operatorname{Im} M$ (donc non nul). Par définition, pour tout $k \in [\![1,n]\!]$, il existe $y_k \in \mathbb{K}$ tel que $C_k(M) = y_k X$.
Ainsi, en notant Y la colonne des $y_1, \dots, y_n$, $M = XY^{\mathrm{T}}$. ■

Remarque
Si on ajoute l'hypothèse que les colonnes X et Y sont toutes les deux non nulles, la réciproque est vraie.

Exemple

Notons E_i les vecteurs de la base canonique de $\mathscr{M}_{n,1}(\mathbb{K})$ et $E_{i,j}$ les vecteurs de la base canonique de $\mathscr{M}_n(\mathbb{K})$. Alors

$$E_{i,j} = E_i {E_j}^{\mathrm{T}}.$$

Exemple

Pour la matrice $J \in \mathscr{M}_n(\mathbb{K})$ dont tous les coefficients sont égaux à 1, $J = XX^{\mathrm{T}}$ où X désigne la colonne dont tous les coefficients sont égaux à 1.

Remarque

Ce résultat sur les matrices de rang 1 permet aisément de calculer leurs puissances. En effet, si $M = XY^{\mathrm{T}}$, alors

$$M^2 = XY^{\mathrm{T}}XY^{\mathrm{T}} = (Y^{\mathrm{T}}X)XY^{\mathrm{T}} = (Y^{\mathrm{T}}X)M$$

en remarquant que la matrice $Y^{\mathrm{T}}X$ est de taille 1×1 donc un scalaire. On peut encore simplifier cette écriture en remarquant que $Y^{\mathrm{T}}X = \mathrm{tr}(Y^{\mathrm{T}}X) = \mathrm{tr}(XY^{\mathrm{T}}) = \mathrm{tr}(M)$. Ainsi, $M^2 - \mathrm{tr}(M)M = 0_n$ (ce qui donne un polynôme annulateur de degré 2 pour une telle matrice).

Ce résultat se généralise au cas de matrices de rang r fixé.

Proposition 17.16.

Une matrice $M \in \mathscr{M}_n(\mathbb{K})$ est de rang $r \leq n$ si, et seulement s'il existe deux familles libres $(X_1, \ldots, X_r)$ et $(Y_1, \ldots, Y_r)$ de $\mathscr{M}_{n,1}(\mathbb{K})$ telles que

$$M = \sum_{k=1}^{r} X_k {Y_k}^{\mathrm{T}}.$$

Démonstration

($\Rightarrow$) Il existe des matrices U et $V \in \mathrm{GL}_n(\mathbb{K})$ telles que $M = UJ_rV$. Alors,

$$M = U\sum_{i=1}^{r} E_i {E_i}^{\mathrm{T}} V = \sum_{i=1}^{r} UE_i(V^{\mathrm{T}}E_i)^{\mathrm{T}} = \sum_{i=1}^{r} C_i(U)C_i(V^{\mathrm{T}})^{\mathrm{T}}.$$

Comme les matrices U et V^{T} sont inversibles, leurs colonnes forment des familles libres, d'où le résultat.

($\Leftarrow$) Complétons les familles $(X_1, \ldots, X_r)$ et $(Y_1, \ldots, Y_r)$ en des bases $(X_1, \ldots, X_n)$ et $(Y_1, \ldots, Y_n)$ de $\mathscr{M}_{n,1}(\mathbb{K})$. Considérons les matrices U et V dont les colonnes sont respectivement $(X_1, \ldots, X_n)$ et $(Y_1, \ldots, Y_n)$. Par construction, ces matrices sont inversibles et

$$M = U\sum_{i=1}^{r} E_i {E_i}^{\mathrm{T}} V^{\mathrm{T}} = UJ_rV^{\mathrm{T}}.$$

■

2. Équivalence en pratique

Dans cette partie, on indique une méthode algorithmique pour déterminer le rang d'une matrice.

2.1. Interprétation matricielle

Définition 17.17.

Soit $i, j \in [\![1, n]\!]$ et $\lambda \in \mathbb{K}$.

▹ La matrice de permutation d'indices $i \neq j$ est la matrice de $\mathscr{M}_n(\mathbb{K})$ définie par

$$P_{i,j} = \mathrm{I}_n - E_{i,i} - E_{j,j} + E_{i,j} + E_{j,i}.$$

▹ La matrice de dilatation d'indice i et de rapport $\lambda \neq 0$ est la matrice de $\mathscr{M}_n(\mathbb{K})$ définie par
$$D_{i,\lambda} = \mathrm{I}_n + (\lambda - 1)E_{i,i}.$$

▹ La matrice de transvection d'indices $i \neq j$ et de rapport λ est la matrice de $\mathscr{M}_n(\mathbb{K})$ définie par
$$T_{i,j,\lambda} = \mathrm{I}_n + \lambda E_{i,j}.$$

Proposition 17.18.

Les matrices de permutation, dilatation et transvection sont inversibles. Plus précisément,
$$(P_{i,j})^{-1} = P_{i,j}, \qquad (D_{i,\lambda})^{-1} = D_{i,\frac{1}{\lambda}}, \qquad (T_{i,j,\lambda})^{-1} = T_{i,j,-\lambda}.$$

Démonstration

Il suffit d'effectuer les produits entre une matrice et le candidat pour son inverse. Par exemple, pour une matrice de transvection, on observe
$$\begin{aligned} T_{i,j,\lambda}.T_{i,j,-\lambda} &= (\mathrm{I}_n + \lambda E_{i,j})(\mathrm{I}_n - \lambda E_{i,j}) \\ &= \mathrm{I}_n - \lambda E_{i,j} + \lambda E_{i,j} - \lambda^2 E_{i,j}E_{i,j} \\ &= \mathrm{I}_n. \end{aligned}$$

■

Proposition 17.19.

La multiplication (à droite ou à gauche) par les matrices de permutation, dilatation et transvection correspond aux opérations élémentaires (sur les colonnes ou sur les lignes). Plus précisément,

M	opération pour MA	opération pour AM
$P_{i,j}$	$L_i \longleftrightarrow L_j$	$C_i \longleftrightarrow C_j$
$D_{i,\lambda}$	$L_i \leftarrow \lambda L_i$	$C_i \leftarrow \lambda C_i$
$T_{i,j,\lambda}$	$L_i \leftarrow L_i + \lambda L_j$	$C_j \leftarrow C_j + \lambda C_i$

Démonstration

Tous les cas se traitent à l'identique ; calculons, par exemple, les colonnes de $AT_{i,j,\lambda}$.

▹ Pour tout $k \neq j$, $C_k(AT_{i,j,\lambda}) = AC_k(T_{i,j,\lambda}) = AC_k(\mathrm{I}_n) = C_k(A)$.

▹ Pour $k = j$,
$$\begin{aligned} C_k(AT_{i,j,\lambda}) = C_j(AT_{i,j,\lambda}) = AC_j(T_{i,j,\lambda}) &= AC_j(\mathrm{I}_n) + \lambda AC_j(E_{i,j}) \\ &= C_j(A) + \lambda C_j(AE_{i,j}) = C_j(A) + \lambda C_i(A). \end{aligned}$$

■

Remarque

Comme les matrices d'opérations élémentaires sont inversibles, on peut en déduire que pour toute matrice A et toute matrice d'opération élémentaire M, $\operatorname{Ker} MA = \operatorname{Ker} A$ et $\operatorname{Im} AM = \operatorname{Im} A$.

2.2. Matrices échelonnées

Définition 17.20.

Soit $A=(a_{i,j})\in\mathscr{M}_{n,p}(\mathbb{K})$. Notons, pour tout $i\in[\![1,n]\!]$,
- $\ell_i=p+1$ si la ligne $L_i(A)$ est nulle,
- $\ell_i=\min\{j\in[\![1,p]\!],\ a_{i,j}\neq 0\}$, l'indice de colonne du premier terme non nul de la ligne i, sinon.

La matrice A est échelonnée en lignes si, pour tout $i\in[\![1,n-1]\!]$,
- si $\ell_i=p+1$, alors $\ell_{i+1}=p+1$,
- sinon, $\ell_i<\ell_{i+1}$.

Autrement dit, une matrice est échelonnée en lignes s'il y a « de plus en plus de zéros en début de ligne » à mesure qu'on parcourt les lignes vers le bas.

On définit de même les matrices échelonnées en colonnes.

Exemple

Parmi les matrices suivantes, la première est échelonnée en lignes ; la seconde en colonnes.

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 5 & 6 & 7 \\ 0 & 0 & 0 & 8 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 \\ 3 & 5 & 0 & 0 & 0 \\ 4 & 6 & 0 & 0 & 0 \end{bmatrix}$$

Remarque

Les matrices diagonales (et plus généralement triangulaires supérieures) ne sont pas forcément échelonnées en lignes ; par exemple, la matrice suivante n'est pas échelonnée

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Remarque

La somme de deux matrices échelonnées en lignes n'est pas nécessairement échelonnée en lignes ; par exemple,

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}+\begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}=\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

La proposition suivante peut paraître élémentaire mais est importante dans la détermination du rang.

Proposition 17.21.

Le rang d'une matrice échelonnée en lignes (respectivement en colonnes) est le nombre de lignes (respectivement de colonnes) non nulles.

2.3. Pivot de Gauss

Proposition 17.22.

Soit $A \in \mathscr{M}_{n,p}(\mathbb{K})$.
▹ Il existe $U \in \mathrm{GL}_n(\mathbb{K})$ telle que UA soit échelonnée en lignes.
▹ Il existe $V \in \mathrm{GL}_p(\mathbb{K})$ telle que AV soit échelonnée en colonnes.

Donnons une démonstration du premier point (l'autre se démontre de la même manière). L'idée est d'utiliser les opérations élémentaires à gauche pour agir sur les lignes.

Démonstration

Montrons, par récurrence sur $n \in \mathbb{N} \setminus \{0\}$, que, pour tout $p \in \mathbb{N} \setminus \{0\}$ et toute matrice $A \in \mathscr{M}_{n,p}(\mathbb{K})$, il existe $U \in \mathrm{GL}_n(\mathbb{K})$ telle que UA soit échelonnée en lignes.

▹ Si $n = 1$, alors la matrice A est une matrice ligne donc échelonnée en lignes.

▹ Soit $n \in \mathbb{N} \setminus \{0\}$. Supposons que le résultat soit vrai pour toute matrice à n lignes et considérons A une matrice à $n+1$ lignes et p colonnes. Si A est la matrice nulle, le résultat est évident ; sinon, considérons l'entier $r \in \mathbb{N}$ tel que les colonnes $C_1(A), \dots, C_r(A)$ soient nulles et $C_{r+1}(A)$ soit non nulle.

Notons $A' \in \mathscr{M}_{n+1,p-r}(\mathbb{K})$ la matrice composée des colonnes d'indices $r+1$ à p de A, c'est-à-dire telle que

$$A = \begin{bmatrix} 0_{n+1,r} & A' \end{bmatrix}.$$

Par définition, il existe $i \in [\![1, n+1]\!]$ tel que $a_{i,r+1} \neq 0$. La matrice $B = P_{i,1}A' = (b_{i,j})_{i,j}$ a un coefficient non nul en position $1,1$. Toujours avec les opérations élémentaires, la matrice

$$T_{2,1,-\frac{b_{2,1}}{b_{1,1}}} \cdots T_{n+1,1,-\frac{b_{n+1,1}}{b_{1,1}}} B$$

est de la forme

$$\begin{bmatrix} b_{1,1} & L \\ 0_{n,1} & A'' \end{bmatrix},$$

avec $L \in \mathscr{M}_{1,p-r-1}(\mathbb{K})$ et $A'' \in \mathscr{M}_{n,p-r-1}(\mathbb{K})$.

▹ Si B est une matrice colonne (c'est-à-dire si $p - r = 1$), nous avons obtenu une forme échelonnée en lignes pour B donc pour A.

▹ Sinon, par hypothèse de récurrence, il existe $U'' \in \mathrm{GL}_n(\mathbb{K})$ telle que $U''A''$ soit échelonnée en lignes.

Considérons la matrice de $\mathrm{GL}_{n+1}(\mathbb{K})$ définie par

$$U = \begin{bmatrix} 1 & 0_{1,n} \\ 0_{n,1} & U'' \end{bmatrix}.$$

Alors, les matrices

$$UA' = \begin{bmatrix} b_{1,1} & L \\ 0_{n,1} & U''A'' \end{bmatrix} \qquad \text{et} \qquad UA = \begin{bmatrix} U0_{n+1,r} & UA' \end{bmatrix} = \begin{bmatrix} 0_{1,r} & b_{1,1} & L \\ 0_{n,r} & 0_{n,1} & U''A'' \end{bmatrix}$$

sont échelonnées en lignes. ■

Cette preuve nous donne un algorithme récursif pour la manipulation des matrices appelé algorithme du pivot de Gauss.

- Si tous les coefficients (sauf éventuellement le premier) de la première colonne sont nuls, on passe à la matrice extraite en supprimant la première ligne et la première colonne.
- Sinon, on ramène un coefficient non nul en position $1,1$ puis on retranche un multiple de la première ligne à chacune des autres lignes de sorte à annuler les premiers coefficients.

Pour conserver la matrice U ou la matrice V, il suffit de conserver la trace des opérations élémentaires sur une seconde matrice initialement égale à l'identité.

Exemple

Cherchons la matrice U pour $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 1 & -2 & 3 \\ 3 & 3 & 6 \end{bmatrix}$.

On part de $\left[\begin{array}{ccc|cccc} 1 & 2 & 3 & 1 & 0 & 0 & 0 \\ 4 & 5 & 6 & 0 & 1 & 0 & 0 \\ 1 & -2 & 3 & 0 & 0 & 1 & 0 \\ 3 & 3 & 6 & 0 & 0 & 0 & 1 \end{array}\right]$ et on applique sur les deux termes (représentés ici comme deux blocs d'une même matrice « augmentée ») les opérations de l'algorithme du pivot de Gauss pour la première matrice.

$$\left[\begin{array}{ccc|cccc} 1 & 2 & 3 & 1 & 0 & 0 & 0 \\ 0 & -3 & -6 & -4 & 1 & 0 & 0 \\ 0 & -4 & 0 & -1 & 0 & 1 & 0 \\ 0 & -3 & -3 & -3 & 0 & 0 & 1 \end{array}\right] \qquad \begin{array}{l} L_2 \leftarrow L_2 - 4L_1 \\ L_3 \leftarrow L_3 - L_1 \\ L_4 \leftarrow L_4 - 3L_1 \end{array}$$

$$\left[\begin{array}{ccc|cccc} 1 & 2 & 3 & 1 & 0 & 0 & 0 \\ 0 & -3 & -6 & -4 & 1 & 0 & 0 \\ 0 & 0 & 24 & 13 & -4 & 3 & 0 \\ 0 & 0 & 3 & 1 & -1 & 0 & 1 \end{array}\right] \qquad \begin{array}{l} L_3 \leftarrow 3L_3 - 4L_2 \\ L_4 \leftarrow L_4 - L_2 \end{array}$$

$$\left[\begin{array}{ccc|cccc} 1 & 2 & 3 & 1 & 0 & 0 & 0 \\ 0 & -3 & -6 & -4 & 1 & 0 & 0 \\ 0 & 0 & 24 & 13 & -4 & 3 & 0 \\ 0 & 0 & 0 & -5 & -4 & -3 & 8 \end{array}\right] \qquad L_4 \leftarrow 8L_4 - L_3$$

donc $U = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -4 & 1 & 0 & 0 \\ 13 & -4 & 3 & 0 \\ -5 & -4 & -3 & 8 \end{bmatrix}$.

On vérifie que $UA = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -4 & 1 & 0 & 0 \\ 13 & -4 & 3 & 0 \\ -5 & -4 & -3 & 8 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 1 & -2 & 3 \\ 3 & 3 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 0 & 24 \\ 0 & 0 & 0 \end{bmatrix}$.

Exemple

Échelonnons cette même matrice A en colonnes. Pour cela, adoptons une autre disposition (non canonique) qui consiste à placer la matrice de départ au-dessus d'une matrice identité et d'effectuer les opérations sur les colonnes.

$$\left[\begin{array}{ccc} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 1 & -2 & 3 \\ 3 & 3 & 6 \\ \hline 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right] \quad \left[\begin{array}{ccc} 1 & 0 & 0 \\ 4 & -3 & -6 \\ 1 & -4 & 0 \\ 3 & -3 & -3 \\ \hline 1 & -2 & -3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right] \quad \left[\begin{array}{ccc} 1 & 0 & 0 \\ 4 & -3 & 0 \\ 1 & -4 & 8 \\ 3 & -3 & 3 \\ \hline 1 & -2 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{array}\right]$$

$$C_2 \leftarrow C_2 - 2C_1 \qquad C_3 \leftarrow C_3 - 2C_2$$
$$C_3 \leftarrow C_3 - 3C_1$$

Ainsi, la matrice $V = \begin{bmatrix} 1 & -2 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix}$ vérifie

$$AV = \begin{bmatrix} 1 & 0 & 0 \\ 4 & -3 & 0 \\ 1 & -4 & 8 \\ 3 & -3 & 3 \end{bmatrix}.$$

On retrouve alors la proposition déjà établie dans la première partie de ce chapitre.

Corollaire 17.23.

Soit $A \in \mathscr{M}_{n,p}(\mathbb{K})$. Alors, il existe $V \in \mathrm{GL}_p(\mathbb{K})$ et $U \in \mathrm{GL}_n(\mathbb{K})$ telle que $UAV = J_r$ où $r = \mathrm{rg}\, A$.

Remarque

Comme le rang d'une matrice échelonnée est facile à lire, on calcule le rang d'une matrice en l'échelonnant en lignes ou en colonnes avec l'algorithme de Gauss.

Par exemple, la matrice

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 1 & -2 & 3 \\ 3 & 3 & 6 \end{bmatrix}$$

des exemples précédents est de rang 3 car on déduit par opérations élémentaires les matrices suivantes de rang 3 à partir de A (une seule suffirait évidemment pour obtenir le rang) :

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 0 & 24 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ 4 & -3 & 0 \\ 1 & -4 & 8 \\ 3 & -3 & 3 \end{bmatrix}.$$

2.4. Détermination du noyau et de l'image

Proposition 17.24.

Soit $A \in \mathscr{M}_{n,p}(\mathbb{K})$ de rang r, $V \in \mathrm{GL}_p(\mathbb{K})$ telle que AV soit échelonnée en colonnes. Alors,
▹ $(C_1(AV), \dots, C_r(AV))$ est une base de $\mathrm{Im}\, A$;
▹ $(C_{r+1}(V), \dots, C_p(V))$ est une base de $\mathrm{Ker}\, A$.

Démonstration

▹ La famille $(C_1(AV), \dots, C_r(AV))$ est composée de vecteurs de $\mathrm{Im}\, A$ et est de rang $r = \dim \mathrm{Im}\, A$ (la famille est libre car la matrice AV est échelonnée de rang r) donc c'est une base de $\mathrm{Im}\, A$.

▹ La famille $(C_{r+1}(V), \dots, C_p(V))$ est une famille de $\mathrm{Ker}\, A$ (car, pour tout $k > r$, $AC_k(V) = C_k(AV) = 0_{n,1}$), libre (les colonnes de V forment une famille libre car V est inversible) et de cardinal $p-r$; or, la dimension de $\mathrm{Ker}\, A$ est $p-r$ d'après le théorème du rang. Ainsi, $(C_{r+1}(V), \dots, C_p(V))$ est une base de $\mathrm{Ker}\, A$. ■

Exemple

▹ D'après les calculs précédents pour la matrice $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 1 & -2 & 3 \\ 3 & 3 & 6 \end{bmatrix}$, on obtient

$$\operatorname{Im} A = \operatorname{Vect}\left(\begin{bmatrix} 1 \\ 4 \\ 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ -3 \\ -4 \\ -3 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 8 \\ 3 \end{bmatrix}\right), \qquad \operatorname{Ker} A = \left\{\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}\right\}.$$

▹ De même, avec la matrice $A = \begin{bmatrix} 1 & 2 & 1 \\ 4 & 5 & 1 \\ 1 & -2 & -3 \\ 3 & 3 & 0 \end{bmatrix}$, on obtient

$$\operatorname{Im} A = \operatorname{Vect}\left(\begin{bmatrix} 1 \\ 4 \\ 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ -3 \\ -4 \\ -3 \end{bmatrix}\right), \qquad \operatorname{Ker} A = \operatorname{Vect}\left(\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}\right).$$

2.5. Calcul de l'inverse

Supposons dans cette section que la matrice considérée est inversible ; on peut alors pousser l'algorithme du pivot de Gauss plus loin :

- on commence comme précédemment par trouver $U \in \mathrm{GL}_n(\mathbb{K})$ telle que UA soit triangulaire supérieure à coefficients diagonaux égaux à 1 ;
- on se sert de la dernière équation pour annuler les derniers coefficients des lignes précédentes puis on recommence avec la matrice privée des dernières ligne et colonne.

La matrice des opérations élémentaires qui permet de passer de A à la matrice identité est l'inverse de A.

Exemple

Appliquons l'algorithme sur la matrice $A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$. Les étapes successives sont

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & -2 & -1 & -2 & 1 & 0 \\ 0 & 0 & -1 & -1 & 0 & 1 \end{array}\right] \qquad \begin{array}{l} L_2 \leftarrow L_2 - 2L_1 \\ L_3 \leftarrow L_3 - L_1 \end{array}$$

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & \frac{1}{2} & 1 & -\frac{1}{2} & 0 \\ 0 & 0 & 1 & 1 & 0 & -1 \end{array}\right] \qquad \begin{array}{l} L_2 \leftarrow -\frac{1}{2}L_2 \\ L_3 \leftarrow -L_3 \end{array}$$

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 & 1 & 0 & -1 \end{array}\right] \qquad \begin{array}{l} L_1 \leftarrow L_1 - L_3 \\ L_2 \leftarrow L_2 - \frac{1}{2}L_3 \end{array}$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 & 1 & 0 & -1 \end{array}\right] \qquad L_1 \leftarrow L_1 - L_2$$

D'où $A^{-1} = \begin{bmatrix} -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ 1 & 0 & -1 \end{bmatrix}$.

Remarque
Attention à utiliser uniquement des opérations sur les lignes ou uniquement des opérations sur les colonnes.

En passant, on a obtenu le résultat suivant.

Proposition 17.25.

Toute matrice de $\mathrm{GL}_n(\mathbb{K})$ peut s'écrire comme un produit de matrices d'opérations élémentaires.

3. Applications aux systèmes linéaires

3.1. Systèmes linéaires

Commençons par quelques définitions sur les systèmes linéaires sans utiliser le calcul matriciel.

Définition 17.26.

▹ L'équation linéaire d'inconnues $x_1, x_2, \ldots, x_p$, de coefficients $a_1, a_2, \ldots, a_p \in \mathbb{C}$ et de second membre $b \in \mathbb{C}$ est

$$\sum_{k=1}^{p} a_k x_k = b.$$

▹ Un système linéaire de n équations à p inconnues est la donnée de n équations linéaires à p inconnues. Les coefficients (respectivement les seconds membres) du système sont les coefficients (respectivement les seconds membres) de chacune des équations qui le composent.

Remarque
Le système linéaire suivant à 3 équations et 4 inconnues

$$\begin{cases} 1x + 2y + 3z + 4t = 2 \\ 5x + 6y + 7z + 8t = 5 \\ 4x + 4y + 4z + 4t = 3. \end{cases}$$

se réécrit $AX = B$ avec

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 4 & 4 & 4 & 4 \end{bmatrix}, \qquad X = \begin{bmatrix} x \\ y \\ z \\ t \end{bmatrix}, \qquad B = \begin{bmatrix} 2 \\ 5 \\ 3 \end{bmatrix}.$$

On peut donc exprimer un système linéaire comme une équation matricielle, chaque équation de l'écriture initiale correspondant à une ligne de la nouvelle.

Forts de cette remarque, reprenons les définitions avec le calcul matriciel.

Définition 17.27.

Soit $A \in \mathscr{M}_{n,p}(\mathbb{K})$ et $B \in \mathscr{M}_{n,1}(\mathbb{K})$.
▹ Le système linéaire de matrice A et de second membre B est l'équation $AX = B$ d'inconnue le vecteur-colonne $X \in \mathscr{M}_{p,1}(\mathbb{K})$.
▹ Le rang du système linéaire $AX = B$ est le rang de sa matrice A.
▹ Le système linéaire $AX = B$ est carré si sa matrice A est carrée.
▹ Le système linéaire $AX = B$ est homogène si son second membre B est nul.

Définition 17.28.

▹ Un système linéaire est compatible s'il admet au moins une solution.
▹ Un système linéaire carré est de Cramer s'il admet une unique solution.

Remarque

Un système homogène est toujours compatible puisque le vecteur nul est solution.

Proposition 17.29.

Un système linéaire carré est de Cramer si, et seulement si, sa matrice est inversible.

Démonstration

($\Leftarrow$) L'unique solution du système $AX = B$ avec A inversible est $X = A^{-1}B$ (en multipliant à gauche par A^{-1}).

($\Rightarrow$) Par contraposition, si la matrice carrée A n'est pas inversible, alors il existe un vecteur X_0 non nul tel que $AX_0 = 0$. Ainsi, si X est une solution de $AX = B$, alors $X + \lambda X_0$ est une solution du même système pour tout $\lambda \in \mathbb{K}$: le système n'est pas de Cramer. ■

3.2. Ensemble des solutions

Commençons par rappeler ce qui découle directement de la définition du noyau.

Proposition 17.30.

L'ensemble des solutions du système linéaire homogène de matrice $A \in \mathscr{M}_{n,p}(\mathbb{K})$ est le sous-espace vectoriel $\operatorname{Ker} A$ (donc de dimension $p - \operatorname{rg} A$).

Proposition 17.31.

L'ensemble $\mathscr{S}$ des solutions du système linéaire de matrice A et de second membre B est soit vide, soit un sous-espace affine de direction $\operatorname{Ker} A$. Plus précisément,
▹ $\mathscr{S}$ est non vide si, et seulement si, $B \in \operatorname{Im} A$.
▹ Soit $X_0 \in \mathscr{S}$. Alors, $X \in \mathscr{S}$ si, et seulement si, $X - X_0 \in \operatorname{Ker} A$.

Remarque

Une équation linéaire (à coefficients non tous nuls) s'interprète comme l'équation d'un hyperplan (vectoriel si le second membre est nul ou tout au moins affine) ; avec cette interprétation géométrique, un système linéaire correspond aux équations de l'intersection d'une famille d'hyperplans.

3.3. Exemples de résolutions

Pour résoudre un système linéaire de matrice $A \in \mathscr{M}_{n,p}(\mathbb{K})$ et de second membre $B \in \mathscr{M}_{n,1}(\mathbb{K})$, on commence par échelonner A en lignes, c'est-à-dire à trouver une matrice $U \in \mathrm{GL}_n(\mathbb{K})$ telle que UA soit échelonnée en lignes.

En notant r le rang de A, on obtient que la condition de compatibilité est que les coefficients en position $k \in [\![r+1, n]\!]$ de UB doivent être nuls. Si cette condition est réalisée, on peut exprimer r coordonnées de X en fonction des autres. Prenons un exemple.

Exemple

Résolvons le système $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 4 & 4 & 4 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ t \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \\ 3 \end{bmatrix}$.

Pour cela échelonnons la matrice $A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 4 & 4 & 4 & 4 \end{bmatrix}$:

$$\left[\begin{array}{rrrr|rrr} 1 & 2 & 3 & 4 & 1 & 0 & 0 \\ 0 & -4 & -8 & -12 & -5 & 1 & 0 \\ 0 & -4 & -8 & -12 & -4 & 0 & 1 \end{array}\right] \qquad \begin{array}{l} L_2 \leftarrow L_2 - 5L_1 \\ L_3 \leftarrow L_3 - 4L_1 \end{array}$$

$$\left[\begin{array}{rrrr|rrr} 1 & 2 & 3 & 4 & 1 & 0 & 0 \\ 0 & -4 & -8 & -12 & -5 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 1 \end{array}\right] \qquad L_3 \leftarrow L_3 - L_2$$

Ainsi, on obtient pour la matrice $U = \begin{bmatrix} 1 & 0 & 0 \\ -5 & 1 & 0 \\ 1 & -1 & 1 \end{bmatrix}$, que $UA = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & -4 & -8 & -12 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

puis, après calculs, $UB = \begin{bmatrix} 2 \\ -5 \\ 0 \end{bmatrix}$.

La condition de compatibilité est vérifiée et l'ensemble des solutions $\mathscr{S}$ admet pour équations

$$\left\{ \begin{array}{rcrcrcrcr} x & + & 2y & + & 3z & + & 4t & = & 2 \\ & - & 4y & - & 8z & - & 12t & = & -5 \end{array} \right.$$

Ainsi, $\begin{bmatrix} x \\ y \\ z \\ t \end{bmatrix} \in \mathscr{S}$ si, et seulement si, $y = \frac{5}{4} - 2z - 3t$ et $x = -\frac{1}{2} + z + 2t$, soit

$$\begin{bmatrix} x \\ y \\ z \\ t \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} \\ \frac{5}{4} \\ 0 \\ 0 \end{bmatrix} + z \begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 2 \\ -3 \\ 0 \\ 1 \end{bmatrix}.$$

Donc, l'ensemble des solutions est

$$\mathscr{S} = \begin{bmatrix} -\frac{1}{2} \\ \frac{5}{4} \\ 0 \\ 0 \end{bmatrix} + \mathrm{Vect}\left(\begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ -3 \\ 0 \\ 1 \end{bmatrix} \right).$$

Remarque

L'intérêt de ce formalisme consiste à pouvoir conclure rapidement pour un autre système linéaire de même matrice. Par exemple, avec la matrice de l'exemple précédent et le second membre $B = \begin{bmatrix} 2 \\ 5 \\ 4 \end{bmatrix}$, on obtient directement $UB = \begin{bmatrix} 2 \\ -5 \\ 1 \end{bmatrix}$ et le système n'est pas compatible.

Il est possible de conserver une écriture « non-matricielle » pour la résolution d'un système linéaire. Reprenons ainsi l'exemple précédent.

Exemple

Résolvons le système $\left\{\begin{array}{rcrcrcrcr} 1x & + & 2y & + & 3z & + & 4t & = & 2 \\ 5x & + & 6y & + & 7z & + & 8t & = & 5 \\ 4x & + & 4y & + & 4z & + & 4t & = & 3. \end{array}\right.$

On procède par des opérations sur les équations (voir la présentation précédente) :

$$\left\{\begin{array}{rcrcrcrcr} 1x & + & 2y & + & 3z & + & 4t & = & 2 \\ & - & 4y & - & 8z & - & 12t & = & -5 \\ & - & 4y & - & 8z & - & 12t & = & -5 \end{array}\right. \qquad \begin{array}{l} \\ L_2 \leftarrow L_2 - 5L_1 \\ L_3 \leftarrow L_3 - 4L_1 \end{array}$$

$$\left\{\begin{array}{rcrcrcrcr} 1x & + & 2y & + & 3z & + & 4t & = & 2 \\ & - & 4y & - & 8z & - & 12t & = & -5 \\ & & & & & & 0 & = & 0 \end{array}\right. \qquad \begin{array}{l} \\ \\ L_3 \leftarrow L_3 - L_2 \end{array}$$

On conclut de la même façon que précédemment.

Exemple

Pour résoudre un système linéaire de matrice $\begin{bmatrix} \lambda-3 & 1 & 1 \\ 1 & \lambda-3 & 0 \\ 1 & 0 & \lambda+1 \end{bmatrix}$ avec $\lambda \in \mathbb{R}$, il convient de discuter selon les valeurs du paramètre λ ; cette discussion apparaît naturellement lors des étapes du pivot de Gauss : lorsque l'on effectue une opération de dilatation, il faut, pour ne pas perdre l'équivalence, s'assurer que le coefficient est non nul. On peut quelquefois utiliser des matrices de permutation pour échanger les lignes et éviter la discussion.

$$\begin{bmatrix} 1 & 0 & \lambda+1 \\ 1 & \lambda-3 & 0 \\ \lambda-3 & 1 & 1 \end{bmatrix} \quad L_1 \longleftrightarrow L_3 \quad \begin{bmatrix} 1 & 0 & \lambda+1 \\ 0 & \lambda-3 & -\lambda-1 \\ 0 & 1 & -\lambda^2+2\lambda+4 \end{bmatrix} \quad \begin{array}{l} L_2 \leftarrow L_2 - L_1 \\ L_3 \leftarrow L_3 - (\lambda-3)L_1 \end{array}$$

$$\begin{bmatrix} 1 & 0 & \lambda+1 \\ 0 & 1 & -\lambda^2+2\lambda+4 \\ 0 & \lambda-3 & -\lambda-1 \end{bmatrix} \quad L_2 \longleftrightarrow L_3 \quad \begin{bmatrix} 1 & 0 & \lambda+1 \\ 0 & 1 & -\lambda^2+2\lambda+4 \\ 0 & 0 & \lambda^3-5\lambda^2+\lambda+11 \end{bmatrix} \quad L_3 \leftarrow L_3 - (\lambda-3)L_2$$

Ainsi, si $\lambda^3 - 5\lambda^2 + \lambda + 11 \neq 0$, le système est de Cramer ; sinon, le système est de rang 2.

Remarque

Avant de terminer, remarquons que la complexité du pivot de Gauss (c'est-à-dire le nombre d'opérations

arithmétiques sur les coefficients dans le pire des cas) pour une matrice carrée de taille n est dominée par n^3 puisqu'il suffit de $n(n-1)$ opérations arithmétiques pour passer d'un système linéaire de taille n à un système linéaire de taille $n-1$.

4. Introduction à la similitude

4.1. Définitions et premiers exemples

Définition 17.32.

Deux matrices $A, B \in \mathscr{M}_n(\mathbb{K})$ sont semblables s'il existe $P \in \mathrm{GL}_n(\mathbb{K})$ telle que $B = P^{-1}AP$.

D'après la formule de changement de bases, deux matrices sont semblables si, et seulement si, elles représentent le même endomorphisme dans des bases différentes.

Remarque

La relation de similitude est réflexive, transitive et symétrique : c'est une relation d'équivalence. On peut donc parler de classe de similitude d'une matrice, c'est-à-dire de l'ensemble des matrices qui lui sont semblables.

Exemple

Soit $\lambda \in \mathbb{K}$. La seule matrice semblable à $\lambda \mathrm{I}_n$ est $\lambda \mathrm{I}_n$ car, pour tout $P \in \mathrm{GL}_n(\mathbb{K})$,

$$P(\lambda \mathrm{I}_n)P^{-1} = \lambda P P^{-1} = \lambda \mathrm{I}_n.$$

Autrement dit, une matrice scalaire est seule dans sa classe de similitude. Réciproquement, une matrice seule dans sa classe de similitude est scalaire car elle commute avec toutes les matrices inversibles donc avec toutes les matrices.

Conseils méthodologiques

Pour établir que deux matrices A, $B \in \mathscr{M}_n(\mathbb{K})$ sont semblables, on considère u l'endomorphisme canoniquement associé à A et on cherche une base dans laquelle la matrice de u est B ; on conclut d'après la formule de changement de bases pour les endomorphismes.

Exemple

Soit $a_1, \ldots, a_n \in \mathbb{K}$ et $\sigma : [\![1,n]\!] \to [\![1,n]\!]$ une application bijective. Les matrices diagonales $\mathrm{diag}(a_1, \ldots, a_n)$ et $\mathrm{diag}(a_{\sigma(1)}, \ldots, a_{\sigma(n)})$ sont semblables.
En effet, notons u l'endomorphisme de $\mathbb{K}^n$ associé à la première dans la base $\mathscr{B} = (e_1, \ldots, e_n)$. Alors, la seconde matrice est la matrice de u dans la base $\mathscr{B} = (e_{\sigma(1)}, \ldots, e_{\sigma(n)})$ ce qui établit la similitude des deux matrices.

Exemple

Montrons que les deux matrices A et $B \in \mathscr{M}_2(\mathbb{R})$ définies par

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$

sont semblables.

Soit u l'endomorphisme de $\mathbb{R}^2$ associé à A dans la base (e_1, e_2). Considérons alors la base $\mathscr{B}' = (e_1, e_1+e_2)$. Comme $u(e_1) = e_1$ et $u(e_1+e_2) = e_1+(e_1+2e_2) = 2(e_1+e_2)$, la matrice de u dans la base $\mathscr{B}'$ est B ce qui établit la similitude des deux matrices.

Exemple Endomorphismes nilpotents d'ordre n

Une matrice $A \in \mathscr{M}_n(\mathbb{C})$ telle que $A^n = 0_n$ et $A^{n-1} \neq 0_n$ est semblable à la matrice $\in \mathscr{M}_n(\mathbb{C})$ avec des 1 sur la « surdiagonale » et des 0 partout ailleurs (c'est-à-dire la matrice dont les coefficients sont $n_{i,j} = \delta_{j,i+1}$).
En effet, si u est l'endomorphisme canoniquement associé à A, alors $u^n = 0$ et $u^{n-1} \neq 0$: il existe $x \in \mathbb{K}^n$ tel que $u^{n-1}(x) \neq 0$. La famille $(u^{n-1}(x), \dots, u(x), x)$ est libre (il suffit d'écrire une combinaison linéaire nulle et de composer suffisamment avec u) donc c'est une base. Dans cette base, u a pour matrice N ce qui établit la similitude des deux matrices A et N.
En considérant la base $(x, u(x), \dots, u^{n-1}(x))$, on obtient de même que la matrice A est semblable à N^{T}.

4.2. Conditions nécessaires à la similitude

Caractériser la similitude de deux matrices est compliqué (davantage que l'équivalence). On se limitera ici à quelques conditions nécessaires à la similitude.

Conseils méthodologiques

Pour montrer que deux matrices ne sont pas semblables, on pourra essayer de montrer que l'une des caractéristiques ci-dessous n'est pas préservée en passant d'une matrice à l'autre.

4.2.1. Trace

Proposition 17.33.

Deux matrices semblables ont la même trace.

Démonstration
Il suffit de remarque que d'après les propriétés de la trace, pour tout $A \in \mathscr{M}_n(\mathbb{K})$ et tout $P \in \mathrm{GL}_n(\mathbb{K})$,

$$\mathrm{tr}(P^{-1}AP) = \mathrm{tr}(APP^{-1}) = \mathrm{tr}\, A.$$

■

Cela permet en particulier de définir la trace d'un endomorphisme comme la trace de sa matrice dans n'importe quelle base fixée.

Exemple

Déduisons de cette propriété que le rang d'une matrice A de projection p est égal à sa trace.
En effet, dans une base associée à la décomposition $E = \mathrm{Im}\, p \oplus \mathrm{Ker}\, p$, la matrice de p est

$$J_r = \begin{bmatrix} \mathrm{I}_r & 0 \\ 0 & 0 \end{bmatrix}.$$

Comme $\mathrm{rg}(J_r) = r = \mathrm{tr}(J_r)$ et que A et J_r sont semblables, $\mathrm{rg}(A) = \mathrm{tr}(A)$.

4.2.2. Rang

Proposition 17.34.

Deux matrices semblables ont même rang.

Démonstration
Deux matrices semblables sont équivalentes d'où le résultat. ■

Remarque
En revanche, deux matrices de même rang ne sont pas nécessairement semblables comme on peut le voir avec I_n et $2\mathrm{I}_n$.

4.2.3. Polynômes annulateurs

Proposition 17.35.

Soit A, $B \in \mathscr{M}_n(\mathbb{K})$ deux matrices semblables et $P \in \mathbb{K}[X]$. Alors, les matrices $P(A)$ et $P(B)$ sont semblables.

Démonstration
Par définition, il existe $U \in \mathrm{GL}_n(\mathbb{K})$ telle que $A = UBU^{-1}$. Remarquons que, pour tout $k \in \mathbb{N}$, $A^k = UB^kU^{-1}$ donc par linéarité $P(A) = UP(B)U^{-1}$. ■

Avec le point précédent, on obtient immédiatement le résultat suivant.

Corollaire 17.36.

Soit A, $B \in \mathscr{M}_n(\mathbb{K})$ deux matrices semblables et $P \in \mathbb{K}[X]$. Alors, les matrices $P(A)$ et $P(B)$ sont de même rang.

Exemple

Soit $A = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ et $B = \mathrm{diag}(\mu_1, \dots, \mu_n)$ deux matrices diagonales semblables. Alors, pour tout $\lambda \in \mathbb{K}$, les matrices $A - \lambda\mathrm{I}_n = \mathrm{diag}(\lambda_1 - \lambda, \dots, \lambda_n - \lambda)$ et $B - \lambda\mathrm{I}_n = \mathrm{diag}(\mu_1 - \lambda, \dots, \mu_n - \lambda)$ sont de même rang (et c'est le nombre de termes non nuls sur leur diagonale). Ainsi, il y a autant de λ_i égaux à λ que de μ_i égaux à λ.
Par conséquent, les termes diagonaux de ces deux matrices sont identiques à permutation près ce qui permet d'en déduire la similitude (en considérant la matrices permutant correctement les vecteurs de la base canonique).

Corollaire 17.37.

Soit A, $B \in \mathscr{M}_n(\mathbb{K})$ deux matrices semblables et $P \in \mathbb{K}[X]$. Alors, P est annulateur de A si, et seulement si, P est annulateur de B.

Exemple

Les matrices de $\mathscr{M}_2(\mathbb{R})$ définies par

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

ne sont pas semblables.
En effet, pour $P(X) = X - 1$, $P(A) = 0_2$ alors que $P(B) \neq 0_2$.

Remarque

Tous ces critères, même considérés conjointement, ne suffisent pas à caractériser les classes de similitude.

Afin de construire un contre-exemple, introduisons la matrice $N_r = (n_{i,j})_{i,j\in[1,r]} \in \mathscr{M}_r(\mathbb{C})$ telle que $n_{i,j} = \delta_{j,i+1}$ puis considérons les matrices par blocs

$$A = \operatorname{diag}(N_4, N_2, N_2), \qquad B = \operatorname{diag}(N_4, N_3, N_1)$$

c'est-à-dire

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \qquad B = . \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Ces deux matrices de $\mathscr{M}_8(\mathbb{C})$ sont de rang 5 (il y a 5 colonnes non nulles et échelonnées), de trace 0 (tous les coefficients sur la diagonale sont nuls) et admettent pour polynômes annulateurs les polynômes multiples de X^4 (ce dernier point s'obtient par calculs par blocs avec N_r nilpotente d'indice r).

Pourtant, les deux matrices ne sont pas semblables car

$$\operatorname{rg}(A^2) = 2 \neq 3 = \operatorname{rg}(B^2).$$

FICHE

SYNTHÈSE

Ce chapitre est consacré à l'algorithme d'échelonnement des matrices (et à son pendant, la résolution de systèmes linéaires).

Formule de changement de bases

Soit E et F deux espaces vectoriels de dimension finie, $\mathscr{B}$ et $\mathscr{B}'$ deux bases de E, $\mathscr{C}$ et $\mathscr{C}'$ deux bases de F.

▹ La matrice de passage de $\mathscr{B}$ à $\mathscr{B}'$ est la matrice notée $P_{\mathscr{B}}^{\mathscr{B}'}$ égale à $\mathrm{mat}_{\mathscr{B}',\mathscr{B}}(\mathrm{Id}_E)$.

▹ Pour tout $x \in E$, $\mathrm{mat}_{\mathscr{B}}(x) = P_{\mathscr{B}}^{\mathscr{B}'}\,\mathrm{mat}_{\mathscr{B}'}(x)$.

▹ Pour tout $u \in \mathscr{L}(E,F)$,

$$\mathrm{mat}_{\mathscr{B}',\mathscr{C}'}(u) = (P_{\mathscr{C}}^{\mathscr{C}'})^{-1}\,\mathrm{mat}_{\mathscr{B},\mathscr{C}}(u)P_{\mathscr{B}}^{\mathscr{B}'}.$$

Ces formules de changements de bases amènent à introduire deux relations d'équivalence : l'équivalence et la similitude.

Équivalence et similitude

▹ Deux matrices A, $B \in \mathscr{M}_{n,p}(\mathbb{K})$ sont équivalentes s'il existe $P \in \mathrm{GL}_n(\mathbb{K})$ et $Q \in \mathrm{GL}_p(\mathbb{K})$ telle que

$$B = PAQ.$$

Deux matrices sont équivalentes si, et seulement si, elles ont même rang.
En particulier, une matrice est de rang r si, et seulement si, elle est équivalente à la matrice J_r.

▹ Deux matrices A, $B \in \mathscr{M}_n(\mathbb{K})$ sont semblables s'il existe une matrice $P \in \mathrm{GL}_n(\mathbb{K})$ telle que

$$B = P^{-1}AP.$$

Deux matrices semblables sont équivalentes et ont même rang. La réciproque est fausse.

L'algorithme du pivot de Gauss permet d'obtenir le résultat suivant d'échelonnement.

Échelonnement

Soit $A \in \mathscr{M}_{n,p}(\mathbb{K})$.

▹ Il existe $U \in \mathrm{GL}_n(\mathbb{K})$ telle que UA soit échelonnée en lignes.

▹ Il existe $V \in \mathrm{GL}_p(\mathbb{K})$ telle que AV soit échelonnée en colonnes.

L'obtention pratique de cette forme échelonnée permet la détermination du rang, de l'image, du noyau d'une matrice voire de son inverse si la matrice appartient au groupe linéaire.

Les outils essentiels sont les opérations élémentaires dont voici la réalisation matricielle.

Opérations élémentaires

Soit $i, j \in [\![1, n]\!]$ et $\lambda \in \mathbb{K}$.
▹ La matrice de permutation d'indices $i \neq j$ est la matrice de $\mathscr{M}_n(\mathbb{K})$ définie par

$$P_{i,j} = \mathrm{I}_n - E_{i,i} - E_{j,j} + E_{i,j} + E_{j,i}.$$

▹ La matrice de dilatation d'indice i et de rapport $\lambda \neq 0$ est la matrice de $\mathscr{M}_n(\mathbb{K})$ définie par

$$D_{i,\lambda} = \mathrm{I}_n + (\lambda - 1)E_{i,i}.$$

▹ La matrice de transvection d'indices $i \neq j$ et de rapport λ est la matrice de $\mathscr{M}_n(\mathbb{K})$ définie par

$$T_{i,j,\lambda} = \mathrm{I}_n + \lambda E_{i,j}.$$

M	opération pour MA	opération pour AM
$P_{i,j}$	$L_i \longleftrightarrow L_j$	$C_i \longleftrightarrow C_j$
$D_{i,\lambda}$	$L_i \leftarrow \lambda L_i$	$C_i \leftarrow \lambda C_i$
$T_{i,j,\lambda}$	$L_i \leftarrow L_i + \lambda L_j$	$C_j \leftarrow C_j + \lambda C_i$

Les systèmes d'équations linéaires admettent une réécriture matricielle et l'algorithme du pivot de Gauss fournit alors une méthode générale de résolution.

Systèmes linéaires

Soit $A \in \mathscr{M}_{n,p}(\mathbb{K})$ et $B \in \mathscr{M}_{n,1}(\mathbb{K})$.
▹ Le système linéaire de matrice A et de second membre B est l'équation $AX = B$ d'inconnue $X \in \mathscr{M}_{p,1}(\mathbb{K})$.
▹ Le rang du système linéaire $AX = B$ est le rang de sa matrice A.
▹ Le système linéaire $AX = B$ est carré si sa matrice A est carrée.
▹ Le système linéaire $AX = B$ est homogène si son second membre B est nul.
▹ Le système linéaire $AX = B$ est compatible s'il admet au moins une solution.
Un système homogène est compatible.
Dès qu'un système est compatible, l'ensemble des solutions est un sous-espace affine dont la direction est le sous-espace des solutions du système homogène associé.
▹ Le système linéaire $AX = B$ carré est de Cramer s'il admet une unique solution. Cela équivaut à l'inversibilité de A.

EXERCICES

Vrai ou faux ?

		Vrai	Faux
a)	Il existe une base de $\mathscr{M}_n(\mathbb{R})$ composée de matrices de rang 1.	☐	☐
b)	Pour $A, B \in \mathscr{M}_n(\mathbb{R})$, il existe $P \in \mathrm{GL}_n(\mathbb{R})$ telle que $A = PB$ si, et seulement si, $\operatorname{Ker} A = \operatorname{Ker} B$.	☐	☐
c)	Pour $A, B \in \mathscr{M}_n(\mathbb{R})$, il existe $P \in \mathrm{GL}_n(\mathbb{R})$ telle que $A = BP$ si, et seulement si, $\operatorname{rg} A = \operatorname{rg} B$.	☐	☐
d)	Pour $A, B \in \mathscr{M}_n(\mathbb{R})$ telles qu'il existe $P \in \mathscr{M}_n(\mathbb{R})$ vérifiant $A = PB$, $\operatorname{rg} A \leq \operatorname{rg} B$.	☐	☐
e)	Pour $A \in \mathscr{M}_n(\mathbb{R})$, $\operatorname{rg} A = \operatorname{rg} A^{\mathrm{T}}$.	☐	☐
f)	Pour $A, B \in \mathscr{M}_n(\mathbb{R})$, $\operatorname{rg}(AB) = \operatorname{rg}(BA)$.	☐	☐
g)	Pour $A, B \in \mathscr{M}_n(\mathbb{R})$, $\operatorname{rg}(AB) \leq \operatorname{rg} B$.	☐	☐
h)	Pour $A, B \in \mathscr{M}_n(\mathbb{R})$, $\operatorname{rg}(AB) = \operatorname{rg}(A)\operatorname{rg}(B)$.	☐	☐
i)	Pour $A \in \mathscr{M}_n(\mathbb{R})$ de rang 1, $A^2 = \operatorname{tr}(A)A$.	☐	☐
j)	Toute matrice de rang 1 est équivalente à $J = (1)_{i,j}$.	☐	☐
k)	Multiplier A à droite par une matrice d'opération élémentaire fait agir l'opération élémentaire correspondante sur ses colonnes.	☐	☐
l)	Une matrice de permutation peut être obtenue comme produit de matrices de transvection et de matrices de dilatation.	☐	☐
m)	Le produit de deux matrices de transvection est une matrice de transvection.	☐	☐
n)	Une matrice échelonnée est inversible.	☐	☐
o)	La matrice $\begin{bmatrix} 1 & -1 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & 3 \end{bmatrix}$ n'est pas inversible.	☐	☐
p)	Pour obtenir une base de $\operatorname{Ker} A$, il suffit d'échelonner A en colonnes et de considérer les dernières colonnes de la matrice d'échelonnement.	☐	☐
q)	Pour tout $A \in \mathscr{M}_n(\mathbb{R})$, il existe P, Q inversibles telles que PAQ soit diagonale.	☐	☐
r)	Pour tout $A \in \mathscr{M}_n(\mathbb{R})$, il existe P inversible telle que $P^{-1}AP$ soit diagonale.	☐	☐
s)	Si le système $AX = B$ admet des solutions, alors A est inversible.	☐	☐

t)	Les solutions du système $AX = B$ forment, quand il en existe, un espace vectoriel.	☐	☐
u)	Si $AX = B$ admet une unique solution, alors $AX = C$ admet aussi une unique solution.	☐	☐
v)	Si $AX = B$ admet une unique solution, alors $AX = C$ admet au plus solution.	☐	☐
w)	Si $A \in \mathscr{M}_{p,q}(\mathbb{R})$ est de rang r, l'espace vectoriel des solutions de $AX = 0_{q,1}$ est de dimension $q - r$.	☐	☐

Exercices d'application du chapitre 17

EXERCICES

●○○ **Exercice 1**

Déterminer les matrices $M \in \mathscr{M}_n(\mathbb{K})$ telles que $\mathrm{rg}(M J_{n-1}) = \mathrm{rg}(M) - 1$.

●●○ **Exercice 2**

Soit $A \in \mathscr{M}_n(\mathbb{R})$.

1. Montrer que la colonne $Y \in \mathscr{M}_{n,1}(\mathbb{R})$ est nulle si, et seulement si, $Y^{\mathrm{T}} Y = 0$.
2. Montrer que $\mathrm{Ker}(A^{\mathrm{T}} A) = \mathrm{Ker}(A)$.
3. En déduire que $\mathrm{rg}(A^{\mathrm{T}} A) = \mathrm{rg}(A)$ puis que $\mathrm{rg}(AA^{\mathrm{T}}) = \mathrm{rg}(A)$.

●●○ **Exercice 3**

Déterminer les $(A, B) \in \mathscr{M}_n(\mathbb{C})^2$ tels que, pour tout $M \in \mathscr{M}_n(\mathbb{C})$, $AMB = 0_n$.

●●○ **Exercice 4**

Soit $A \in \mathscr{M}_n(\mathbb{R})$.

1. Montrer qu'il existe $U \in \mathscr{M}_n(\mathbb{R})$ telle que $AUA = A$.
2. Montrer qu'il existe $U \in \mathrm{GL}_n(\mathbb{R})$ telle que UA est une matrice de projection.

●●○ **Exercice 5**

Soit $\Diamond$ la relation binaire sur $\mathscr{M}_n(\mathbb{K})$ définie par $A \Diamond B$ si

$$\exists U, V \in \mathscr{M}_n(\mathbb{K}), \qquad A = UBV.$$

Montrer que $A \Diamond B$ si, et seulement si, $\mathrm{rg}\, A \leq \mathrm{rg}\, B$.

●●● **Exercice 6**

1. Soit u un endomorphisme d'un espace vectoriel E de dimension finie, $F \subset G$ deux sous-espaces vectoriels de E. Montrer que

$$\dim u(G) - \dim u(F) \leq \dim G - \dim F.$$

2. Soit A, B et $C \in \mathscr{M}_n(\mathbb{R})$. Montrer que

$$\mathrm{rg}(AB) + \mathrm{rg}(BC) \leq \mathrm{rg}(B) + \mathrm{rg}(ABC).$$

●○○ **Exercice 7**

Résoudre, pour $\lambda \in \mathbb{C}$,

$$\left\{\begin{array}{ccccccc} \lambda x & + & y & + & z & = & 1 \\ x & + & \lambda y & + & z & = & \lambda \\ x & + & y & + & \lambda z & = & \lambda^2 \end{array}\right.$$

●○○ **Exercice 8**

Soit $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{bmatrix}$, $B = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$.

Montrer que le système $AX = B$ est compatible si, et seulement si, $c - b = b - a$.

●○○ **Exercice 9**

Résoudre le système

$$\left\{\begin{array}{ccccccccccc} x_1 & + & x_2 & + & x_3 & + & \dots & + & x_n & = & 1 \\ x_1 & + & 2x_2 & + & 2x_3 & + & \dots & + & 2x_n & = & 1 \\ x_1 & + & 2x_2 & + & 3x_3 & + & \dots & + & 3x_n & = & 1 \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ x_1 & + & 2x_2 & + & 3x_3 & + & \dots & + & nx_n & = & 1 \end{array}\right.$$

●○○ **Exercice 10**

Résoudre, pour $\lambda \in \mathbb{C}$,

$$\left\{\begin{array}{ccccccccc} \lambda x & + & y & + & z & + & t & = & 1 \\ x & + & \lambda y & + & z & + & t & = & \lambda \\ x & + & y & + & \lambda z & + & t & = & \lambda + 1 \end{array}\right.$$

●○○ **Exercice 11**

Résoudre le système, pour $\lambda, \mu \in \mathbb{C}$,

$$\left\{\begin{array}{ccccccc} \lambda x & + & \mu y & + & z & = & 1 \\ x & + & \lambda\mu y & + & z & = & \mu \\ x & + & \mu y & + & \lambda z & = & 1 \end{array}\right.$$

●●○ **Exercice 12**

Résoudre dans $\mathbb{R}_+^*$ le système non linéaire suivant

$$\left\{\begin{array}{ccc} xyz & = & 1 \\ xy^2z^4 & = & 2 \\ xy^3z^9 & = & 3 \end{array}\right.$$

●●○ **Exercice 13**

Soit $A \in \mathscr{M}_{n,p}(\mathbb{K})$ de rang r, $U \in \mathrm{GL}_n(\mathbb{K})$ telle que UA soit échelonnée en lignes.

1. Montrer que $X \in \operatorname{Ker} A$ si, et seulement si, $L_1(UA)X = 0, \dots, L_r(UA)X = 0$.
2. Montrer que $Y \in \operatorname{Im} A$ si, et seulement si, $L_{r+1}(U)Y = 0, \dots, L_n(U)Y = 0$.

●●● **Exercice 14**

1. Soit $X \in \mathscr{M}_n(\mathbb{Z})$ et d le PGCD des coefficients de X. Il existe $U \in \mathrm{GL}_n(\mathbb{R})$ à coefficients entiers et d'inverse à coefficients entiers telle que
$$UX = \begin{bmatrix} d \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$
2. Soit $A = (a_{i,j}) \in \mathscr{M}_{n,p}(\mathbb{Z})$. Il existe $U \in \mathrm{GL}_n(\mathbb{R})$ à coefficients entiers et d'inverse à coefficients entiers telle que UA soit échelonnée en ligne.

Exercices d'approfondissement du chapitre 17

●●○ **Exercice A**

Soit $A \in \mathscr{M}_{3,2}(\mathbb{R})$ et $B \in \mathscr{M}_{2,3}(\mathbb{R})$ telles que
$$AB = \begin{bmatrix} 0 & -1 & -1 \\ -1 & 0 & -1 \\ 1 & 1 & 2 \end{bmatrix}.$$

1. Montrer que $\mathrm{rg}(A) = \mathrm{rg}(B) = 2$.
2. En déduire $BA = \mathrm{I}_2$.

●●○ **Exercice B**

Soit $\alpha \in \mathbb{R}$, $A \in \mathscr{M}_{3,2}(\mathbb{R})$, $B \in \mathscr{M}_{2,3}(\mathbb{R})$ telles que
$$AB = \begin{bmatrix} 1 & 4 & 2 \\ 2 & \alpha & 2 \\ 3 & 10 & 2 \end{bmatrix}.$$
Trouver α.

●●● **Exercice C**

1. Soit $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ deux à deux distincts et $A = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$. Déterminer le noyau et l'image de l'endomorphisme
$$\varphi : \left\{ \begin{array}{ccc} \mathscr{M}_n(\mathbb{C}) & \to & \mathscr{M}_n(\mathbb{C}) \\ M & \mapsto & AM - MA \end{array} \right.$$
2. Soit $M \in \mathscr{M}_n(\mathbb{C})$ une matrice de trace nulle.
   a. Montrer que M est semblable à une matrice à diagonale nulle.
   b. Montrer qu'il existe $A, B \in \mathscr{M}_n(\mathbb{C})$ telles que $M = AB - BA$.

●●● **Exercice D**

Soit H un hyperplan de $\mathscr{M}_n(\mathbb{R})$.

1. Montrer qu'il existe $A \in \mathscr{M}_n(\mathbb{R})$ tel que $H = \{M \in \mathscr{M}_n(\mathbb{R}),\ \mathrm{tr}(AM) = 0\}$.
2. En déduire que H contient une matrice inversible.

EXERCICES

●●● **Exercice E**

Considérons $M = (m_{i,j})_{i,j} \in \mathscr{M}_{n+2}(\mathbb{R})$ vérifiant

- $m_{i,j} = 0$ si $i \in \{1, n+2\}$ ou $j \in \{1, n+2\}$;
- $m_{i,j} = \frac{1}{4}(m_{i-1,j} + m_{i+1,j} + m_{i,j-1} + m_{i,j+1})$ sinon.

Déterminer M.

●●● **Exercice F**

1. Soit M une matrice de $\mathscr{M}_n(\mathbb{R})$. Montrer qu'il existe $P \in \mathrm{GL}_n(\mathbb{R})$ telle que les coefficients de PMP^{-1} aux positions $(i,1)$ pour $i \geq 3$ soient nuls.
2. Conclure que toute matrice de $\mathscr{M}_n(\mathbb{R})$ est semblable à une matrice $M = (m_{i,j}) \in \mathscr{M}_n(\mathbb{R})$ telle que $m_{i,j} = 0$ dès que $i > j+1$. Ces matrices sont appelées matrices de Hessenberg.

Problèmes du chapitre 17

●○○ **Problème 1 Commutant d'une matrice**

Pour toute matrice $A \in \mathscr{M}_n(\mathbb{C})$, le commutant de A est l'ensemble $\mathscr{C}(A)$ défini par

$$\mathscr{C}(A) = \{M \in \mathscr{M}_n(\mathbb{C}),\ AM = MA\}.$$

Partie A – Préliminaires et premiers exemples

1. Vérifier que, pour tout $A \in \mathscr{M}_n(\mathbb{C})$, $\mathscr{C}(A)$ est un sous-espace vectoriel de $\mathscr{M}_n(\mathbb{C})$ de dimension supérieure ou égal à 1. Est-ce une sous-algèbre de $\mathscr{M}_n(\mathbb{C})$?
2. Préciser le commutant de I_n.
3. Calculer le commutant de la matrice diagonale $D = \mathrm{diag}(1, 2, \ldots, n)$.
4. Considérons les matrices

$$T = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}, \qquad A = \begin{bmatrix} 1 & 4 & -2 \\ 0 & 6 & -3 \\ -1 & 4 & 0 \end{bmatrix}, \qquad P = \begin{bmatrix} 1 & 4 & -2 \\ 1 & 3 & 0 \\ 1 & 4 & -1 \end{bmatrix}.$$

   a. Déterminer le commutant $\mathscr{C}(T)$ de T et préciser sa dimension.
   Indication : on pourra échapper à quelques calculs en remarquant qu'une matrice de $\mathscr{C}(T)$ laisse stable les sous-espaces $\mathrm{Ker}(T - 3\mathrm{I}_3)$, $\mathrm{Ker}(T - 2\mathrm{I}_3)$ et $\mathrm{Ker}(T - 2\mathrm{I}_3)^3$.
   b. Calculer P^{-1} puis $P^{-1}AP$.
   c. Montrer que l'application $M \mapsto P^{-1}MP$ est un automorphisme de $\mathscr{M}_3(\mathbb{R})$. En déduire la dimension de $\mathscr{C}(A)$.
   d. Existe-t-il un polynôme annulateur de A de degré inférieur ou égal à 2 ?
   e. Montrer que $\mathscr{C}(A) = \mathrm{Vect}(\mathrm{I}_3, A, A^2)$ puis que cet ensemble est exactement l'ensemble des polynômes en A.

Partie B – Dimension maximale

1. Soit $A \in \mathscr{M}_n(\mathbb{C})$ et $B \in \mathscr{M}_p(\mathbb{C})$ des matrices non proportionnelles à la matrice identité et F un sous-espace vectoriel de $\mathscr{M}_{n,p}(\mathbb{C})$ tel que

$$\forall M \in F, \qquad AM = MB.$$

   On cherche à majorer la dimension de F.
   a. Soit $P \in \mathrm{GL}_n(\mathbb{C})$ et $Q \in \mathrm{GL}_p(\mathbb{C})$. Montrer que l'application $f : M \mapsto PMQ$ est un automorphisme de $\mathscr{M}_{n,p}(\mathbb{C})$.

b. Justifier que $\dim f(F) = \dim F$ puis que l'on peut remplacer A et B par des matrices qui leurs sont semblables.

c. Montrer que B est semblable à une matrice de première colonne est

$$\begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Justifier que A est semblable à une matrice de première ligne est

$$\begin{bmatrix} 0 & 1 & 0 & \dots & 0 \end{bmatrix}.$$

Conformément à ce que suggèrent les questions précédentes, on remplace A et B par ces matrices qui leurs sont semblables (que l'on appelle encore A et B) et on considère l'application

$$\varphi : \left\{ \begin{array}{ccc} \mathscr{M}_{n,p}(\mathbb{C}) & \to & \mathscr{M}_{n,p}(\mathbb{C}) \\ M & \mapsto & AM - MB \end{array} \right.$$

d. Montrer qu'une matrice non nulle définie par la structure par blocs

$$\begin{bmatrix} 0 & 0 & 0_{1,p-2} \\ \alpha & \beta & L \\ 0_{n-2,1} & C & 0_{n-2,p-2} \end{bmatrix}$$

avec $L \in \mathscr{M}_{1,p-2}(\mathbb{C})$, $C \in \mathscr{M}_{n-2,1}(\mathbb{C})$ et $\alpha, \beta \in \mathbb{C}$, n'appartient pas à $\operatorname{Ker}\varphi$.

e. En déduire que $\dim \operatorname{Ker}\varphi \leq np - (n+p-2)$.

f. En déduire que $\dim F \leq (n-1)(p-1)+1$.

2. Montrer que si $A \in \mathscr{M}_n(\mathbb{C})$ n'est pas proportionnelle à I_n, alors

$$\dim \mathscr{C}(A) \leq (n-1)^2 + 1.$$

●●○ **Problème 2 Théorème de Flanders**

Soit un entier $r \leq n$ et F un sous-espace vectoriel de $\mathscr{M}_n(\mathbb{R})$ composé de matrices de rang au plus r. On admet le résultat suivant.

Proposition 17.38.

Si F contient la matrice J_r, alors toute matrice $M \in F$ est de la forme par blocs

$$\begin{bmatrix} M_{1,1} & M_{1,2} \\ M_{2,1} & 0_{n-r} \end{bmatrix}$$

avec $M_{1,1} \in \mathscr{M}_r(\mathbb{R})$, $M_{1,2} \in \mathscr{M}_{r,n-r}(\mathbb{R})$, $M_{2,1} \in \mathscr{M}_{n-r,r}(\mathbb{R})$ et $M_{2,1}M_{1,2} = 0_{n-r}$.

Partie A

On suppose dans cette partie que $J_r \in F$.

1. Soit $M, N \in F$ de la forme

$$\begin{bmatrix} M_{1,1} & M_{1,2} \\ M_{2,1} & 0_{n-r} \end{bmatrix}, \qquad \begin{bmatrix} N_{1,1} & N_{1,2} \\ N_{2,1} & 0_{n-r} \end{bmatrix}.$$

Montrer que $M_{2,1}N_{1,2} + N_{2,1}M_{1,2} = 0_{n-r}$.

PROBLÈMES

2. Considérons l'application

$$f : \left\{ \begin{array}{ccc} F & \to & \mathscr{M}_{r,n}(\mathbb{R}) \\ \begin{bmatrix} M_{1,1} & M_{1,2} \\ M_{2,1} & 0_{n-r} \end{bmatrix} & \mapsto & \begin{bmatrix} M_{1,1} & M_{1,2} \end{bmatrix} \end{array} \right.$$

a. Montrer que les matrices de $\operatorname{Ker} f$ sont de la forme

$$\begin{bmatrix} 0_r & 0_{r,n-r} \\ M_{2,1} & 0_{n-r} \end{bmatrix}$$

avec $M_{2,1}N_{1,2} = 0_{n-r}$ pour toute matrice

$$\begin{bmatrix} N_{1,1} & N_{1,2} \\ N_{2,1} & 0_{n-r} \end{bmatrix} \in F.$$

Notons $k = \dim \operatorname{Ker} f$. Considérons l'application

$$g : \left\{ \begin{array}{ccc} \operatorname{Ker} f & \to & \mathscr{M}_{r,n}(\mathbb{R})^* \\ \begin{bmatrix} 0_r & 0_{r,n-r} \\ M_{2,1} & 0_{n-r} \end{bmatrix} & \mapsto & \begin{bmatrix} X_{1,1} & X_{1,2} \end{bmatrix} \mapsto \operatorname{tr}(M_{2,1}X_{1,2}) \end{array} \right.$$

b. Montrer que l'application linéaire g est injective ; en déduire que $\dim g(\operatorname{Ker} f) = k$.

c. Montrer que

$$\operatorname{Im} f \subset \bigcap_{M \in \operatorname{Ker} f} \operatorname{Ker} g(M);$$

en déduire que $\dim \operatorname{Im} f \leq rn - k$.

d. Conclure que F est de dimension au plus rn.

Partie B

Montrer que même si le sous-espace F ne contient pas J_r, alors $\dim F \leq rn$.

CORRIGÉS

Corrigés des Vrai/Faux

a) Vrai avec la base canonique.

b) Vrai.

c) Faux pour le sens retour.

d) Vrai.

e) Vrai.

f) Faux : on a exhibé dans le chapitre précédent des matrices A, B telles que $AB = 0$ mais $BA \neq 0$.

g) Vrai.

h) Faux, par exemple $A = B = \mathrm{I}_n$.

i) Vrai.

j) Vrai, car $\mathrm{rg}(J) = 1$.

k) Vrai.

l) Vrai.

m) Faux.

n) Faux, la matrice nulle est échelonnée mais évidemment pas inversible.

o) Vrai.

p) Vrai.

q) Vrai, toute matrice est équivalente à la matrice diagonale J_r où r est son rang.

r) Faux.

s) Faux.

t) Faux, seulement si B est la colonne nulle (et c'est alors $\mathrm{Ker}\, A$).

u) Faux, rien ne garantit que $C \in \mathrm{Im}\, A$.

v) Vrai.

w) Vrai.

Corrigés des exercices d'application du chapitre 17

Exercice 1

Après calculs, les colonnes de la matrice MJ_{n-1} sont $C_1(M), \dots, C_{n-1}(M)$ et $0_{n,1}$.
Ainsi, $\mathrm{rg}(MJ_{n-1}) = \mathrm{rg}(M) - 1$ si, et seulement si,

$$\dim \mathrm{Vect}(C_1(M), \dots, C_{n-1}(M)) = \dim \mathrm{Vect}(C_1(M), \dots, C_{n-1}(M), C_n(M)) - 1$$

soit encore

$$\dim \mathrm{Vect}(C_1(M), \dots, C_{n-1}(M), C_n(M)) = \dim \mathrm{Vect}(C_1(M), \dots, C_{n-1}(M)) + 1$$

soit enfin

$$\dim \mathrm{Vect}(C_1(M), \dots, C_{n-1}(M)) + \mathrm{Vect}(C_n(M)) = \dim \mathrm{Vect}(C_1(M), \dots, C_{n-1}(M)) + 1.$$

La condition est donc $C_n(M) \notin \mathrm{Vect}(C_1(M), \dots, C_{n-1}(M))$ (pour que la somme de sous-espace à gauche soit directe).

Exercice 2

1. Il suffit de remarquer que l'unique coefficient de $Y^{\mathrm{T}}Y$ est la somme des carrés des coefficients de Y : ainsi, $Y^{\mathrm{T}}Y$ est nul si, et seulement si, la colonne Y n'a que des coefficients nuls.
2. ▹ Soit $X \in \mathrm{Ker}(A)$. Alors, $AX = 0_{n,1}$ donc $A^{\mathrm{T}}AX = A^{\mathrm{T}}0_{n,1} = 0_{n,1}$. Ainsi, $\mathrm{Ker}(A) \subset \mathrm{Ker}(A^{\mathrm{T}}A)$.
▹ Soit $X \in \mathrm{Ker}(A^{\mathrm{T}}A)$. Alors, $A^{\mathrm{T}}AX = 0_{n,1}$ donc $X^{\mathrm{T}}A^{\mathrm{T}}AX = X^{\mathrm{T}}0_{n,1} = 0$. Par conséquent, $(AX)^{\mathrm{T}}(AX) = 0$ et, d'après la première question $AX = 0_{n,1}$. Ainsi, $\mathrm{Ker}(A^{\mathrm{T}}A) \subset \mathrm{Ker}(A)$.
3. ▹ D'après la deuxième question, $\dim \mathrm{Ker}(A^{\mathrm{T}}A) = \dim \mathrm{Ker}(A)$ donc, en utilisant la formule du rang, $\mathrm{rg}(A^{\mathrm{T}}A) = \mathrm{rg}(A)$.
▹ Appliquons le résultat établi au point précédent à la matrice A^{T} et utilisons que $(A^{\mathrm{T}})^{\mathrm{T}} = A$: on obtient que $\mathrm{rg}(AA^{\mathrm{T}}) = \mathrm{rg}(A^{\mathrm{T}})$. Utilisons dorénavant que $\mathrm{rg}(A^{\mathrm{T}}) = \mathrm{rg}(A)$ pour conclure que $\mathrm{rg}(AA^{\mathrm{T}}) = \mathrm{rg}(A)$.

Exercice 3

▹ Notons, pour tout $r \in [\![0, n]\!]$, J_r la matrice diagonale dont les coefficients aux positions (i, i) sont égaux à 1 si $i \leq r$, à 0 sinon.
Si r et r' sont les rangs de matrices A et B convenant, alors il existe P, P', Q, Q' inversibles telles que $A = PJ_rQ$, et $B = P'J_{r'}Q'$. Alors, pour $M = Q^{-1}P'^{-1}$, la matrice $ABM = PJ_rJ_{r'}Q' = PJ_{\min(r,r')}Q'$ est de rang $\min(r, r')$; elle est donc nulle si, et seulement si, le rang $\min(r, r')$ est nul, c'est-à-dire si $r = 0$ ou $r' = 0$: donc $A = 0_n$ ou $B = 0_n$.
▹ Réciproquement, si l'une (au moins) des matrices A ou B est nulle, alors, pour toute matrice $M \in \mathscr{M}_n(\mathbb{C})$, $AMB = 0_n$.

Exercice 4

1. Notons r le rang de A. Il existe des matrices $P, Q \in \mathrm{GL}_n(\mathbb{R})$ telles que la matrice PAQ soit égale à la matrice par blocs définie par

$$\begin{bmatrix} \mathrm{I}_r & 0_{r,n-r} \\ 0_{n-r,r} & 0_{n-r} \end{bmatrix}.$$

Considérons la matrice par blocs définie par

$$U = Q\begin{bmatrix} M_1 & M_2 \\ M_3 & M_4 \end{bmatrix}P.$$

Alors, $AUA = A$ si, et seulement si,

$$\begin{bmatrix} \mathrm{I}_r & 0_{r,n-r} \\ 0_{n-r,r} & 0_{n-r} \end{bmatrix}\begin{bmatrix} M_1 & M_2 \\ M_3 & M_4 \end{bmatrix}\begin{bmatrix} \mathrm{I}_r & 0_{r,n-r} \\ 0_{n-r,r} & 0_{n-r} \end{bmatrix}$$

est égale à

$$\begin{bmatrix} \mathrm{I}_r & 0_{r,n-r} \\ 0_{n-r,r} & 0_{n-r} \end{bmatrix},$$

soit $M_1 = \mathrm{I}_r$. On a ainsi exhibé une matrice U qui convient.
2. Avec la question précédente, il existe U tel que $AUA = A$ donc $(AU)^2 = AUAU = AU$: ainsi, AU est une matrice de projection.

Exercice 5

▹ Supposons qu'il existe des matrices U, $V \in \mathscr{M}_n(\mathbb{K})$ telles que $A = UBV$. Alors, $\operatorname{Im} A \subset \operatorname{Im}(UB)$ donc $\operatorname{rg}(A) \leq \operatorname{rg}(UB)$. Par ailleurs, $\operatorname{Im}(UB) = U\operatorname{Im}(B)$ donc $\operatorname{rg}(UB) \leq \operatorname{rg}(B)$.
Combinant ces deux résultats, on obtient $\operatorname{rg}(A) \leq \operatorname{rg}(B)$.
▹ Notons $r = \operatorname{rg}(A)$ et $s = \operatorname{rg}(B)$ et supposons $r \leq s$. Il existe des matrices inversibles P, Q, P' et Q' telles que $A = PJ_rQ$ et $B = P'J_sQ'$. Posons $U = PJ_r(P')^{-1}$ et $V = (Q')^{-1}Q$. Alors,

$$\begin{aligned} UBV &= PJ_r(P')^{-1}P'J_sQ'(Q')^{-1}Q \\ &= PJ_rJ_sQ \\ &= PJ_rQ = A \end{aligned}$$

d'où le résultat.

Exercice 6

1. Soit $f_1, \ldots, f_r$ une base d'un supplémentaire de $u(F)$ dans $u(G)$ (donc $r = \dim u(G) - \dim u(F)$) et $e_1, \ldots, e_r$ des antécédents dans G de ces vecteurs. Montrons que le sous-espace $\operatorname{Vect}(e_1, \ldots, e_r)$ est de dimension r et en somme directe avec F dans G. Pour cela, cherchons les scalaires $\lambda_1, \ldots, \lambda_r$ tels que $\sum_{k=1}^r \lambda_k e_k \in F$ (ce qui inclus le cas où la combinaison linéaire est le vecteur nul) ; alors, en composant par u, $\sum_{k=1}^r \lambda_k f_k \in u(F)$ donc, par définition de la famille $(f_1, \ldots, f_r)$, $\sum_{k=1}^r \lambda_k f_k = 0_E$ et par liberté de cette même famille, $\lambda_k = 0$ pour tout $k \leq r$. Ainsi, $F \oplus \operatorname{Vect}(e_1, \ldots, e_r) \subset G$ ce qui en passant à la dimension indique que $\dim F + r \leq \dim G$ ce qui est bien l'inégalité annoncée.
2. Notons a, b et c les endomorphismes de $E = \mathbb{R}^n$ canoniquement associés aux matrices A, B et C. Appliquons le lemme précédent à $u = a$, $F = \operatorname{Im} b \circ c$ et $G = \operatorname{Im} b$. Alors,

$$\dim a(\operatorname{Im} b) - \dim a(\operatorname{Im} b \circ c) \leq \dim \operatorname{Im} b - \dim \operatorname{Im} b \circ c,$$

soit l'inégalité désirée $\operatorname{rg} AB - \operatorname{rg} ABC \leq \operatorname{rg} B - \operatorname{rg} BC$.

Exercice 7

$$\left\{\begin{array}{ccccccccl} x & + & \lambda y & + & z & = & \lambda & & L_1 \longleftrightarrow L_2 \\ \lambda x & + & y & + & z & = & 1 & & \\ x & + & y & + & \lambda z & = & \lambda^2 & & \end{array}\right.$$

$$\left\{\begin{array}{ccccccccl} x & + & \lambda y & + & z & = & \lambda & & \\ & & (1-\lambda^2)y & + & (1-\lambda)z & = & 1-\lambda^2 & & L_2 \leftarrow L_2 - \lambda L_1 \\ & & (1-\lambda)y & + & (\lambda-1)z & = & \lambda^2-\lambda & & L_3 \leftarrow L_3 - L_1 \end{array}\right.$$

$$\left\{\begin{array}{ccccccccc} x & + & \lambda y & + & z & = & \lambda & & \\ & & (1-\lambda)y & + & (\lambda-1)z & = & \lambda^2-\lambda & & L_2 \longleftrightarrow L_3 \\ & & (1-\lambda^2)y & + & (1-\lambda)z & = & 1-\lambda^2 & & \end{array}\right.$$

$$\left\{\begin{array}{ccccccccc} x & + & \lambda y & + & z & = & \lambda & & \\ & & (1-\lambda)y & + & (\lambda-1)z & = & \lambda^2-\lambda & & \\ & & & & (1-\lambda)(2+\lambda)z & = & (1-\lambda)^2(1+\lambda) & & L_3 \leftarrow L_3-(1+\lambda)L_2 \end{array}\right.$$

Ainsi, si $\lambda = 1$, le système se réduit à l'équation $x+y+z=1$; si $\lambda=-2$, il est incompatible à cause de la dernière équation ; sinon, il y a un unique triplet solution.

Exercice 8

Résolvons le système $AX = B$ par la méthode du pivot de Gauss.

$$\left\{\begin{array}{ccccccc} x & + & 2y & + & 3z & = & a \\ 2x & + & 3y & + & 4z & = & b \\ 3x & + & 4y & + & 5z & = & c \end{array}\right.$$

$$\left\{\begin{array}{ccccccccc} x & + & 2y & + & 3z & = & a & & \\ & - & y & - & 2z & = & b-2a & & L_2 \leftarrow L_2-2L_1 \\ & - & 2y & - & 4z & = & c-3a & & L_3 \leftarrow L_3-3L_1 \end{array}\right.$$

$$\left\{\begin{array}{ccccccccc} x & + & 2y & + & 3z & = & a & & \\ & - & y & - & 2z & = & b-2a & & \\ & & & & 0 & = & c+a-2b & & L_3 \leftarrow L_3-2L_2 \end{array}\right.$$

Ainsi, le système admet des solutions si, et seulement si, $a+c-2b=0$.

Dans cet exercice, nous avons en fait déterminé une équation de l'image de A.

Exercice 9

En effectuant $L_k \leftarrow L_k - L_{k-1}$ pour k allant de n à 2, on obtient le système

$$\left\{\begin{array}{ccccccccccc} x_1 & + & x_2 & + & x_3 & + & \dots & + & x_n & = & 1 \\ & & x_2 & + & x_3 & + & \dots & + & x_n & = & 0 \\ & & & & x_3 & + & \dots & + & x_n & = & 0 \\ & & & & & & \ddots & & \vdots & & \vdots \\ & & & & & & & & x_n & = & 0 \end{array}\right.$$

dont l'unique solution est $(1,0,0,\dots,0)$.

Exercice 10

Le système est équivalent à

$$\left\{\begin{array}{ccccccccccc} \lambda x & + & y & + & z & + & t & = & 1 & & \\ x & + & y & + & \lambda z & + & t & = & \lambda+1 & & L_2 \longleftrightarrow L_3 \\ x & + & \lambda y & + & z & + & t & = & \lambda & & \end{array}\right.$$

$$\left\{\begin{array}{ccccccccccc} \lambda x & + & y & + & z & + & t & = & 1 & & \\ (1-\lambda)x & & & + & (\lambda-1)z & & & = & \lambda & & L_2 \leftarrow L_2-L_1 \\ (1-\lambda)x & + & (\lambda-1)y & & & & & = & \lambda-1 & & L_3 \leftarrow L_3-L_1 \end{array}\right.$$

Si $\lambda = 1$, la seconde équation (donc le système) n'admet pas de solution. Sinon, les solutions sont

$$x\,(1,1,1,-2-\lambda)+\left(0,1,\tfrac{\lambda}{\lambda-1},-\tfrac{\lambda}{\lambda-1}\right),$$

avec $x \in \mathbb{C}$.

Exercice 11

$$\left\{\begin{array}{ccccccc} \mu y & + & \lambda x & + & z & = & 1 \\ \lambda\mu y & + & x & + & z & = & \mu \\ \mu y & + & x & + & \lambda z & = & 1 \end{array}\right.$$

$$\left\{\begin{array}{ccccccccc} \mu y & + & \lambda x & + & z & = & 1 & & \\ & & (1-\lambda^2)x & + & (1-\lambda)z & = & \mu-\lambda & & L_2 \leftarrow L_2-\lambda L_1 \\ & & (1-\lambda)x & + & (\lambda-1)z & = & 0 & & L_3 \leftarrow L_3-L_1 \end{array}\right.$$

$$\left\{\begin{array}{ccccccccc} \mu y & + & \lambda x & + & z & = & 1 & & \\ & & (1-\lambda^2)x & + & (1-\lambda)z & = & \mu-\lambda & & \\ & & (2-\lambda-\lambda^2)x & & & = & \mu-\lambda & & L_3 \leftarrow L_3+L_2 \end{array}\right.$$

- Si $\lambda=-2$, la dernière équation devient $\lambda=\mu$ et la précédente $-3x+3z=0$. Les solutions sont donc
$$\left\{\left(z,-\tfrac{1+z}{2},z\right),\ \text{avec } z\in\mathbb{C}\right\}.$$
- Si $\lambda=1$, les deux dernières équations donnent $\lambda=\mu$ et la première $x+y+z=1$.
- Sinon, le système est de Cramer et il y a une unique solution.

Exercice 12

Comme tous les termes sont strictement positifs, passons au logarithme et on obtient le système linéaire en $X=\ln x$, $Y=\ln y$ et $Z=\ln z$:

$$\left\{\begin{array}{ccccccc} X & + & Y & + & Z & = & 0 \\ X & + & 2Y & + & 4Z & = & \ln 2 \\ X & + & 3Y & + & 9Z & = & \ln 3 \end{array}\right.$$

Après résolution par la méthode du pivot de Gauss, on obtient

$$(X,Y,Z)=\left(\frac{1}{2}\ln\frac{9}{64},\frac{1}{2}\ln\frac{256}{27},\frac{1}{2}\ln\frac{3}{4}\right),$$

et donc $(x,y,z)=(\frac{3}{8},\frac{16}{3\sqrt{3}},\frac{\sqrt{3}}{2})$.

Exercice 13

1. $X\in\operatorname{Ker}A$ si, et seulement si, $UAX=0$ (car U inversible) donc si, et seulement si, pour tout $k\in[\![1,n]\!]$, $L_k(UA)X=0$. Cette dernière condition est équivalente (car les dernières lignes de UA sont nulles) à $L_k(UA)X=0$ pour tout $k\in[\![1,r]\!]$.
2. $Y\in\operatorname{Im}A$ si, et seulement s'il existe X tel que $Y=AX$, c'est-à-dire tel que $UY=UAX$ (car U inversible). Cela entraîne $L_k(U)Y=0$ pour tout $k\in[\![r+1,n]\!]$. Réciproquement, cette condition suffit car les r premières lignes de UA forment une matrice injective.

Exercice 14

1. Procédons par récurrence sur $n\in\mathbb{N}\setminus\{0\}$.

▹ Le résultat est évident pour tout vecteur $X\in\mathscr{M}_{1,1}(\mathbb{Z})$.

▹ Soit $X=\begin{bmatrix} a \\ b \end{bmatrix}\in\mathscr{M}_{2,1}(\mathbb{Z})$. Notons $d=a\wedge b$, $a=a'd$, $b=b'd$ et considérons des entiers $u,v\in\mathbb{Z}$ tels que $ua+vb=d$. Alors

$$\begin{bmatrix} u & v \\ b' & -a' \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix}=\begin{bmatrix} d \\ 0 \end{bmatrix}$$

et

$$\begin{bmatrix} u & v \\ b' & -a' \end{bmatrix}^{-1} = \begin{bmatrix} -a' & -v \\ -b' & u \end{bmatrix} \in \mathscr{M}_2(\mathbb{Z}).$$

▹ Soit $n \in \mathbb{N} \setminus \{0\}$ tel que la propriété soit vraie au rang n. Considérons $X = \begin{bmatrix} x_1 \\ \vdots \\ x_{n+1} \end{bmatrix} \in \mathscr{M}_{n+1,1}(\mathbb{Z})$. D'après l'hypothèse de récurrence, il existe une matrice $U \in \mathrm{GL}_n(\mathbb{R})$ à coefficients entiers et d'inverse à coefficients entiers telle que

$$U \begin{bmatrix} x_2 \\ x_3 \\ \vdots \\ x_{n+1} \end{bmatrix} = \begin{bmatrix} x_2 \wedge x_3 \wedge \ldots \wedge x_{n+1} \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Par ailleurs, d'après la propriété au rang 2, il existe $U' \in \mathrm{GL}_2(\mathbb{R})$ à coefficients entiers et d'inverse à coefficients entiers telle que

$$U' \begin{bmatrix} x_1 \\ x_2 \wedge x_3 \wedge \ldots \wedge x_{n+1} \end{bmatrix} = \begin{bmatrix} x_1 \wedge x_2 \wedge \ldots \wedge x_{n+1} \\ 0 \end{bmatrix}.$$

Notons

$$U'' = \begin{bmatrix} U' & 0_{2,1} & \ldots & 0_{2,1} \\ 0_{1,2} & & & \\ \vdots & & \mathrm{I}_{n-1} & \\ 0_{1,2} & & & \end{bmatrix} \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & & & \\ \vdots & & U & \\ 0 & & & \end{bmatrix}$$

Alors, U'' est à coefficients dans $\mathbb{Z}$, d'inverse à coefficients dans $\mathbb{Z}$ car, en faisant les calculs par blocs, on obtient

$$(U'')^{-1} = \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & & & \\ \vdots & & U^{-1} & \\ 0 & & & \end{bmatrix} \begin{bmatrix} (U')^{-1} & 0_{2,1} & \ldots & 0_{2,1} \\ 0_{1,2} & & & \\ \vdots & & \mathrm{I}_{n-1} & \\ 0_{1,2} & & & \end{bmatrix}.$$

Alors

$$U'' \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n+1} \end{bmatrix} = \begin{bmatrix} U' & 0_{2,1} & \ldots & 0_{2,1} \\ 0_{1,2} & & & \\ \vdots & & \mathrm{I}_{n-1} & \\ 0_{1,2} & & & \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \wedge x_3 \wedge \ldots \wedge x_{n+1} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} x_1 \wedge x_2 \wedge \ldots \wedge x_{n+1} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

2. Raisonnons par récurrence sur n le nombre de lignes de la matrice A.
▹ Si $A \in \mathscr{M}_{1,p}(\mathbb{Z})$ alors A est déjà échelonnée en lignes.
▹ Soit $n \in \mathbb{N} \setminus \{0\}$ tel que la propriété soit vérifiée pour toute matrice à n lignes et à coefficients entiers et $A \in \mathscr{M}_{n+1,p}(\mathbb{Z})$. D'après la première question, il existe $U \in \mathrm{GL}_{n+1}(\mathbb{R})$ à coefficients entiers et d'inverse à

coefficients entiers telle que $UC_1(A)$ soit nulle en dehors de la première ligne. Alors, UA est de la forme

$$\begin{bmatrix} d & \star & \dots & \star \\ 0 & & & \\ \vdots & & A' & \\ 0 & & & \end{bmatrix}.$$

Par hypothèse de récurrence, il existe $V \in \mathrm{GL}_n(\mathbb{R})$ à coefficients entiers et d'inverse à coefficients entiers telle que VA' soit échelonnée en lignes. Ainsi,

$$\begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & V & \\ 0 & & & \end{bmatrix} UA$$

est échelonnée en lignes.

Corrigés des exercices d'approfondissement du chapitre 17

Exercice A

1. ▹ Commençons par déterminer le rang de AB. Les deux premières colonnes de AB ne sont pas proportionnelles donc $\mathrm{rg}(AB) \geq 2$; comme la troisième est la somme des deux premières, $\mathrm{rg}(AB) = 2$.
▹ On a immédiatement que $\mathrm{Im}\, AB \subset \mathrm{Im}\, A$ donc $\mathrm{rg}(A) \geq 2$. Comme $\mathrm{rg}(A) \leq \min(3,2) = 2$, on conclut que $\mathrm{rg}(A) = 2$. Remarquons que cela entraîne notamment que

$$\mathrm{Im}(A) = \mathrm{Im}(AB) = \mathrm{Vect}\left(\begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}\right).$$

▹ De même, on utilise l'inclusion $\mathrm{Ker}(B) \subset \mathrm{Ker}(AB)$ pour obtenir $\dim \mathrm{Ker}\, B \leq 1$. D'après la formule du rang pour B, $\mathrm{rg}(B) \geq 3 - \dim \mathrm{Ker}(B) \geq 2$. En comparant à nouveau avec les dimensions, $\mathrm{rg}(B) \leq \min(2,3) = 2$ donc $\mathrm{rg}(B) = 2$.
2. La matrice BA appartient à $\mathscr{M}_2(\mathbb{R})$. Montrons que tout vecteur de $\mathrm{Im}\, A$ demeure inchangé par la matrice AB. D'après la détermination de l'image que l'on a trouvé à la question précédente, il suffit de remarquer cette propriété sur une base de l'image soit

$$\begin{bmatrix} 0 & -1 & -1 \\ -1 & 0 & -1 \\ 1 & 1 & 2 \end{bmatrix}\begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}, \qquad \begin{bmatrix} 0 & -1 & -1 \\ -1 & 0 & -1 \\ 1 & 1 & 2 \end{bmatrix}\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.$$

Par conséquent, $ABA = A$ puisque pour tout $X \in \mathscr{M}_{2,1}(\mathbb{R})$, $AX \in \mathrm{Im}\, A$ donc $ABAX = AX$.
Ainsi, $A(BA - \mathrm{I}_2) = 0_{3,2}$ et donc $\mathrm{Im}(BA - \mathrm{I}_2) \subset \mathrm{Ker}(A)$. Or, d'après la formule du rang pour la matrice 1 à 2 colonnes,

$$\dim \mathrm{Ker}(A) = 2 - \mathrm{rg}(A) = 0.$$

Ainsi, l'image de $BA - \mathrm{I}_2$ est réduite au sous-espace nul donc cette matrice est 0_2. En conclusion, $BA = \mathrm{I}_2$.

Exercice B

Remarquons tout d'abord que $\operatorname{rg}(AB) \leq \min(\operatorname{rg} A, \operatorname{rg} B) \leq 2$. Or, $C_1(AB)$ et $C_3(AB)$ ne sont pas colinéaires donc forment une base de $\operatorname{Im} AB$. Ainsi, il existe λ, $\mu \in \mathbb{R}$ tels que $C_2(AB) = \lambda C_1(AB) + \mu C_3(AB)$. On est donc ramené au système suivant

$$\left\{\begin{array}{rcrcl} \lambda & + & 2\mu & = & 4 \\ 2\lambda & + & 2\mu & = & \alpha \\ 3\lambda & + & 2\mu & = & 10 \end{array}\right.$$

En résolvant ce système d'inconnues λ et μ, on obtient la condition de compatibilité $\alpha = 7$.

Exercice C

1. ▹ $M \in \operatorname{Ker} \varphi$ si, et seulement si, pour tout (i, j),

$$\lambda_i m_{i,j} = \lambda_j m_{i,j}.$$

Par conséquent, pour tout $i \neq j$, $m_{i,j} = 0$ (car $\lambda_i \neq \lambda_j$). $\operatorname{Ker} \varphi$ est l'espace des matrices diagonales et $\dim \operatorname{Ker} \varphi = n$.
On peut remarquer que $\operatorname{Ker} \varphi$ est l'espace des polynômes en A en utilisant l'interpolation de Lagrange aux points $\lambda_1, \ldots, \lambda_n$.
▹ D'après la formule du rang, $\dim \operatorname{Im} \varphi = n^2 - n$; or, $\operatorname{Im} \varphi$ est inclus dans l'espace des matrices à diagonale nulle qui est de dimension $n^2 - n$. En conclusion, $\operatorname{Im} \varphi$ est l'espace des matrices à diagonale nulle.
2. a. Raisonnons par récurrence sur $n \in \mathbb{N} \setminus \{0\}$.
▹ Pour $n = 1$, une matrice de trace nulle est nulle donc à diagonale nulle.
▹ Soit $n \geq 1$ tel que la propriété soit vraie au rang n et $M \in \mathscr{M}_{n+1}(\mathbb{R})$ de trace nulle.
• Si, pour tout $X \in \mathscr{M}_{n,1}(\mathbb{C})$, MX et X sont colinéaires, alors M est une matrice d'homothétie et $M = 0_n$ car de trace nulle : le résultat est établi.
• Sinon, il existe un vecteur X tel que $MX \notin \operatorname{Vect} X$. Complétons (X, MX) en une base de $\mathscr{M}_{n,1}(\mathbb{C})$. D'après la formule du changement de bases, M est semblable à une matrice M' de la forme

$$\begin{bmatrix} 0 & L \\ 1 & \\ 0 & N \\ \vdots & \end{bmatrix}$$

Comme $\operatorname{tr} M' = \operatorname{tr} M = 0$, on en déduit que $\operatorname{tr} N = 0$. Appliquons l'hypothèse de récurrence à N, il existe $P \in GL_n(\mathbb{C})$ tel que PNP^{-1} est à diagonale nulle. Alors

$$\begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & & & \\ \vdots & & P & \\ 0 & & & \end{bmatrix} M' \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & & & \\ \vdots & & P^{-1} & \\ 0 & & & \end{bmatrix}$$

est à diagonale nulle. En conclusion, M est semblable à une matrice à diagonale nulle.
b. Écrivons $M = PNP^{-1}$ où N est à diagonale nulle.
Soit A une matrice diagonale à coefficients diagonaux deux à deux distincts et φ l'application associée comme dans la question 1. On a vu que $N \in \operatorname{Im} \varphi$ donc qu'il existe $B \in \mathscr{M}_n(\mathbb{C})$ telle que $N = AB - BA$; par conséquent,

$$M = P(AB - BA)P^{-1} = (PAP^{-1})(PBP^{-1}) - (PBP^{-1})(PAP^{-1}),$$

ce qui est le résultat attendu.

Exercice D

1. Soit φ une forme linéaire non nulle telle que $H = \operatorname{Ker}\varphi$. Considérons la matrice $A = (\varphi(E_{j,i}))_{i,j}$. Alors, pour tout (i, j), $\varphi(E_{i,j}) = [A]_{j,i} = \operatorname{tr}(AE_{i,j})$. Par linéarité, φ est l'application $M \mapsto \operatorname{tr}(AM)$ et donc $\varphi(M) = 0$ si, et seulement si, $\operatorname{tr}(AM) = 0$.
2. Discutons selon le rang de A.
- Si A est inversible, la matrice $A^{-1}\operatorname{diag}(1, 1 \ldots, 1, 1-n)$ appartient à H et est inversible.
- Si $\operatorname{rg} A = r < n$, alors il existe P, Q inversibles telles que $A = PK_rQ$ où $K_r = (\delta_{j,n+r-i})_{i,j}$ car deux matrices de même rang sont équivalentes. La matrice $M = Q^{-1}P^{-1}$ est inversible et appartient à H car

$$\operatorname{tr}(AM) = \operatorname{tr}(PK_rP^{-1}) = \operatorname{tr} K_r = 0.$$

Exercice E

Montrons par l'absurde que la matrice est nulle. Considérons (i_0, j_0) l'indice d'un des coefficients de valeur absolue maximale (donc $i_0 \notin \{1, n+2\}$ et $j_0 \notin \{1, n+2\}$) puis définissons

$$i_1 = \min\{i \leq i_0,\ |m_{i,j_0}| = |m_{i_0,j_0}|\}.$$

Par définition de i_1, $i_1 \neq 0$ (car $m_{0,j} = 0$ et $|m_{i_1,j_0}| = |m_{i_0,j_0}| > 0$) et $|m_{i_1-1,j_0}| < |m_{i_1,j_0}|$.
La seconde condition en (i_1, j_0) donne

$$|m_{i_1,j_0}| \leq \frac{1}{4}(|m_{i_1-1,j_0}| + |m_{i_1+1,j_0}| + |m_{i_1,j_0-1}| + |m_{i_1,j_0+1}|) \leq \frac{1}{4}(|m_{i_1-1,j_0}| + 3|m_{i_1,j_0}|) < |m_{i_1,j_0}|,$$

d'où la contradiction. En conclusion, la seule matrice vérifiant ce système d'équations linéaires est la matrice nulle 0_{n+2}.

Exercice F

1. Soit M une matrice de $\mathscr{M}_n(\mathbb{R})$. Si la première colonne de M est nulle ou ne possède que le premier coefficient non nul, le résultat est établi. Sinon, il existe au moins un coefficient non nul à une position notée $(k, 1)$ avec $k \geq 2$ que nous allons ramener en position $(2, 1)$ à l'aide de la matrice de permutation $P_{2,k} = \mathrm{I}_n - E_{k,k} - E_{2,2} + E_{2,k} + E_{k,2}$. En effet, la matrice $P_{2,k}MP_{2,k}$ est semblable à M (car $P_{2,k}^{-1} = P_{2,k}$) et se déduit de M en permutant les lignes et colonnes d'indices 2 et k ; intuitivement, seule la multiplication à gauche est utile pour ramener un coefficient non nul en position $(2, 1)$ mais on cherche à rester dans la classe de similitude de M ; on vérifie que l'opération sur les colonnes ne modifie pas la première colonne. On suppose désormais que $m_{2,1} \neq 0$. Pour chaque coefficient non nul en position $(k, 1)$ avec $k > 2$, on multiplie à gauche par $(\mathrm{I}_n - \frac{m_{k,1}}{m_{2,1}}E_{k,2})$ et à droite par l'inverse $(\mathrm{I}_n + \frac{m_{k,1}}{m_{2,1}}E_{k,2})$. Le seul coefficient de la première colonne à être modifié est en position $(k, 1)$ et il vaut désormais $m_{k,1} - \frac{m_{k,1}}{m_{2,1}}m_{2,1} = 0$. Ainsi, par opérations successives, on se ramène à une matrice M' semblable à M dont la première colonne est nulle sauf éventuellement les coefficients $m'_{1,1}$ et $m'_{2,1}$. Remarquons et c'est un point essentiel pour la suite qu'aucune des opérations élémentaires effectuées n'affectent pas le coefficient en position $1, 1$ puisqu'elles ne concernent que les dernières lignes et colonnes, c'est-à-dire que les coefficients de la matrice de passage sur la première ligne et la première colonne sont tous nuls sauf celui en position $1, 1$ qui est égal à 1.
2. Procédons par récurrence sur la taille de la matrice.
▹ Toutes les matrices de taille 1, 2 ou 3 vérifient l'hypothèse de récurrence.
▹ Soit $n \in \mathbb{N} \setminus \{0\}$; supposons le résultat établi pour les matrices de taille n et considérons une matrice de taille $n+1$. D'après la question précédente, cette matrice est semblable à une matrice de la forme

$$A = \begin{bmatrix} m_{1,1} & L \\ m_{2,1} & \\ 0 & \\ \vdots & A' \\ 0 & \end{bmatrix},$$

CORRIGÉS

avec $L \in \mathscr{M}_{1,n}(\mathbb{R})$ et $A' \in \mathscr{M}_n(\mathbb{R})$.
Par hypohtèse de récurrence, il existe une matrice de passage $P \in \mathrm{GL}_n(\mathbb{R})$, dont tous les coefficients sur la première ligne et la première colonne sont tous nuls sauf celui en position $1,1$ qui est égal à 1, telle que la matrice $PA'P^{-1}$ soit de Hessenberg. En effectuant le produit par blocs :

$$\begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & P & \\ 0 & & & \end{bmatrix} A \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & P^{-1} & \\ 0 & & & \end{bmatrix} = \begin{bmatrix} m_{1,1} & & LP^{-1} & \\ m_{2,1} & & & \\ 0 & & & \\ \vdots & & PA'P^{-1} & \\ 0 & & & \end{bmatrix},$$

on obtient que A est semblable à une matrice de Hessenberg ce qui achève la récurrence.

Corrigés des problèmes du chapitre 17

Problème 1

Partie A – Préliminaires et premiers exemples

1. Remarquons tout d'abord que $\mathrm{I}_n \in \mathscr{C}(A)$. Par ailleurs, pour toutes matrices $M, N \in \mathscr{C}(A)$ et tout $\lambda \in \mathbb{C}$,

$$A(M+\lambda N) = AM + \lambda AN = MA + \lambda NA = (M+\lambda N)A$$
$$AMN = (AM)N = M(AN) = MNA.$$

D'où $M+\lambda N$ et $MN \in \mathscr{C}(A)$. En conclusion, $\mathscr{C}(A)$ est une sous-algèbre de $\mathscr{M}_n(\mathbb{C})$ (et comme elle contient la matrice non-nulle I_n, sa dimension est au moins 1).
2. Toute matrice commute avec I_n donc $\mathscr{C}(\mathrm{I}_n) = \mathscr{M}_n(\mathbb{C})$.
3. ▹ Soit $M = (m_{i,j})$ une matrice commutant avec D. La relation $[MD]_{i,j} = [DM]_{i,j}$ devient $i\,m_{i,j} = j\,m_{i,j}$. Par conséquent, pour tout $i \neq j$, $m_{i,j} = 0$: la matrice M est diagonale. *On aurait pu montrer ce résultat sans calcul en remarquant que si M commute avec D alors M laisse stable le sous-espace* $\mathrm{Ker}(D - i\mathrm{I}_n) = \mathrm{Vect}(e_i)$.
▹ Réciproquement, deux matrices diagonales commutent. En conclusion, le commutant de D est l'espace des matrices diagonales.
4. a. On peut encore calculer les matrices M commutant avec T. Remarquons pour se simplifier les calculs que si M commute avec T avec M laisse stable les sous-espaces $\mathrm{Ker}(T - 3\mathrm{I}_3) = \mathrm{Vect}(e_1)$, $\mathrm{Ker}(T - 2\mathrm{I}_3) = \mathrm{Vect}(e_2)$ et $\mathrm{Ker}(T - 2\mathrm{I}_3)^2 = \mathrm{Vect}(e_2, e_3)$. Ainsi, M est de la forme

$$\begin{bmatrix} a & 0 & 0 \\ 0 & b & c \\ 0 & 0 & d \end{bmatrix}.$$

En calculant $MT = TM$ (ce qui est maintenant plus simple), on obtient $b = d$. Ainsi, le commutant de T est l'espace des matrices de la forme

$$\begin{bmatrix} a & 0 & 0 \\ 0 & b & c \\ 0 & 0 & b \end{bmatrix}$$

qui est de dimension 3 avec une base donnée par

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

b. Après échelonnement, on obtient successivement

$$P^{-1} = \begin{bmatrix} 3 & 4 & -6 \\ -1 & -1 & 2 \\ -1 & 0 & 1 \end{bmatrix},$$

puis $P^{-1}AP = T$.
c. L'application $M \mapsto P^{-1}MP$ est linéaire ; sa réciproque est l'application $M \mapsto PMP^{-1}$. On vérifie rapidement que $M \in \mathscr{C}(A)$ si, et seulement si, $f(M) \in \mathscr{C}(T)$; en effet,

$$\begin{aligned} MA = AM &\Leftrightarrow MPTP^{-1} = PTP^{-1}M \\ &\Leftrightarrow P^{-1}MPT = TP^{-1}MP \\ &\Leftrightarrow f(M)T = Tf(M). \end{aligned}$$

Ainsi, $\mathscr{C}(T) = f(\mathscr{C}(A))$ et donc $\dim \mathscr{C}(A) = \dim \mathscr{C}(T) = 3$.
d. Soit α, β et $\gamma \in \mathbb{C}$ tels que $\alpha \mathrm{I}_3 + \beta A + \gamma A^2 = 0_3$. Alors, en factorisant par P et P^{-1}, $\alpha \mathrm{I}_3 + \beta T + \gamma T^2 = 0_3$. Comme

$$T^2 = \begin{bmatrix} 9 & 0 & 0 \\ 0 & 4 & 4 \\ 0 & 0 & 4 \end{bmatrix}$$

la relation donne le système $\alpha + 3\beta + 9\gamma = 0$, $\alpha + 2\beta + 4\gamma = 0$ et $\beta + 4\gamma = 0$ dont l'unique solution est $\alpha = \beta = \gamma = 0$. Il n'y a donc pas de polynôme annulateur de A de degré inférieur ou égal à 2.
e. On vérifie immédiatement que I_3, A et A^2 forment une famille libre de $\mathscr{C}(A)$ de cardinal $3 = \dim \mathscr{C}(A)$ donc en est une base. Ainsi, $\mathscr{C}(A) = \mathrm{Vect}(\mathrm{I}_3, A, A^2)$.
Pour conclure, remarquons que l'espace des polynômes en A est inclus dans $\mathscr{C}(A)$ et que $\mathrm{Vect}(\mathrm{I}_3, A, A^2)$ est inclus dans l'espace des polynômes en A.

Partie B – Dimension maximale

1. a. L'application $f : M \mapsto PMQ$ est clairement linéaire et est bijective de réciproque $M \mapsto P^{-1}MQ^{-1}$.
b. Comme f est un automorphisme, l'image par f d'une base de F est une base de $f(F)$ et donc $\dim f(F) = \dim F$. Remarquons que $Af(M) = f(M)B$ si, et seulement si,

$$(P^{-1}AP)M = M(QBQ^{-1}).$$

Ainsi, $f(F)$ est un sous-espace qui vérifie la propriété de F en remplaçant A et B par leurs semblables $P^{-1}AP$ et QBQ^{-1} de même dimension de F.
c. B n'est pas proportionnelle à l'identité donc il existe une colonne X telle que (X, BX) est libre. Dans une base complétée à partir de ces vecteurs, la matrice semblable à B à la première colonne recherchée. On recommence avec la transposée de A pour avoir le résultat.
d. Soit $L \in \mathscr{M}_{1,n-2}(\mathbb{C})$, $C \in \mathscr{M}_{p-2,1}(\mathbb{C})$, $\alpha, \beta \in \mathbb{C}$ et $M = \begin{bmatrix} 0 & 0 & 0_{1,n-2} \\ \alpha & \beta & L \\ 0_{p-2,1} & C & 0_{p-1,n-2} \end{bmatrix} \in \mathrm{Ker}\,\varphi$.

Alors, $L_1(\varphi(M)) = \begin{bmatrix} \alpha & \beta & L \end{bmatrix}$ donc $\alpha = \beta = 0$ et $L = 0_{1,n-2}$. Ensuite, $C_1(\varphi(M)) = \begin{bmatrix} 0 \\ \beta \\ C \end{bmatrix}$,

donc $C = 0_{p-2,1}$ et par conséquent $M = 0_{n,p}$.
e. L'espace des matrices de la forme précédente est en somme directe avec $\mathrm{Ker}\,\varphi$ (car leur intersection est réduite à la matrice nulle) et de dimension $n + p - 2$: ainsi

$$np = \dim \mathscr{M}_{n,p}(\mathbb{C}) \geq n + p - 2 + \dim \mathrm{Ker}\,\varphi.$$

D'où $\dim \mathrm{Ker}\,\varphi \leq np - (n + p - 2)$.
f. Comme $F \subset \mathrm{Ker}\,\varphi$, $\dim F \leq np - (n + p - 2) = (n-1)(p-1) + 1$.

2. En prenant $n = p$ et $A = B$ dans la question précédente, on obtient que les sous-espaces F composés de matrices commutant avec A sont de dimension au plus $(n-1)^2+1$. En particulier, avec $F = \mathscr{C}(A)$, $\dim \mathscr{C}(A) \leq (n-1)^2+1$.

Problème 2

Partie A

1. Comme $M+N \in F$, la proposition entraîne que $(M_{2,1}+N_{2,1})(M_{1,2}+N_{1,2}) = 0_{n-r}$ donc

$$M_{2,1}N_{1,2} + N_{2,1}M_{1,2} = 0_{n-r}$$

car $M_{2,1}M_{1,2} = N_{2,1}N_{1,2} = 0_{n-r}$.
2. a. La forme de l'énoncé est évidente ; la condition de produit vient de la question précédente.
b. Soit $M \in \operatorname{Ker} g$. Alors $g(M)(0_r {M_{2,1}}^{\mathrm{T}}) = \operatorname{tr}(M_{2,1}{M_{2,1}}^{\mathrm{T}}) = 0$. Ainsi, la somme des carrés des coefficients de $M_{2,1}$ est nulle donc $M_{2,1}$ est nulle. En conclusion, le noyau de g est réduit à la matrice nulle donc g est injective.
Par conséquent, g induit un isomorphisme entre $\operatorname{Ker} f$ et $g(\operatorname{Ker} f)$: ces deux espaces sont donc de même dimension.
c. ▹ L'inclusion $\operatorname{Im} f \subset \bigcap_{M \in \operatorname{Ker} f} \operatorname{Ker} g(M)$ est une conséquence de la proposition admise.
▹ Soit $g(X_1), \ldots, g(X_k)$ une base de $g(\operatorname{Ker} f)$. D'après ce qui précède, $\operatorname{Im} f \subset \cap_{j=1}^{k} \operatorname{Ker} g(X_j)$. Comme intersection de k hyperplans indépendants de l'espace de dimension rn, $\dim \operatorname{Im} f \leq rn-k$.
d. D'après la formule du rang appliquée à f donne

$$\dim F = \dim \operatorname{Ker} f + \dim \operatorname{Im} f \leq k + rn - k = rn.$$

Partie B

Soit $M \in F$ de rang maximal s. Il existe $P, Q \in \mathrm{GL}_n(\mathbb{R})$ tel que $PMQ = J_s$. L'application linéaire

$$\varphi : \left\{ \begin{array}{ccc} F & \to & \mathscr{M}_n(\mathbb{R}) \\ A & \mapsto & PAQ \end{array} \right.$$

est injective. Ainsi, $\dim F = \dim \varphi(F)$. Or $\varphi(F)$ est un sous-espace vectoriel constitué de matrices de rang au plus r (car $\operatorname{rg} PAQ = \operatorname{rg} A$) et contenant J_s. D'après la partie précédente,

$$\dim F = \dim \varphi(F) \leq sn \leq rn.$$

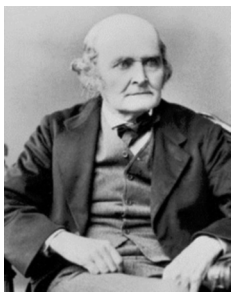
Chapitre 18

Groupe symétrique

1. Permutations — **2.** Décomposition en cycles disjoints — **3.** Signature

Objectifs et compétences du programme

- Calculer des compositions de permutations.
- Obtenir la décomposition en cycles à supports disjoints d'une permutation.
- Calculer l'ordre d'une permutation.
- Calculer une signature.
- Exploiter une partie génératrice de S_n.



Arthur Cayley, mathématicien britannique (1821-1895)
On le croise essentiellement selon deux entrées :

- d'une part, avec la structure algébrique de groupes qu'il a contribué à définir et à étudier (la table de Cayley d'un groupe est la table de "multiplication" de la loi interne, le théorème de Cayley est un exercice de ce chapitre) ;
- d'autre part, avec le produit matriciel qu'il est le premier à considérer (on lui attribue un résultat de réduction, théorème de Cayley-Hamilton).

Son œuvre mathématique ne se limite pourtant pas à ses domaines puisqu'il a également fait de la combinatoire et de la géométrie projective (où il a laissé son nom à une surface d'équation cubique non singulière sur laquelle on peut tracer 27 droites).

COURS

Ce chapitre admet deux objectifs : d'une part, introduire les permutations et les principales caractéristiques afin de disposer de cet outil pour définir le déterminant au chapitre suivant ; d'autre part, considérer le groupe symétrique dans son ensemble comme un exemple riche de groupe fini et ainsi permettre des révisions du chapitre 10 « Structures algébriques ».

1. Permutations

1.1. Définitions

Définition 18.1.

▹ Une permutation d'un ensemble E est une application de E dans E bijective.
▹ Le groupe symétrique de E est l'ensemble des permutations de E muni de la composition. Il est noté $S(E)$ ou plus simplement S_n lorsque $E = [\![1, n]\!]$.

On vérifie immédiatement que le groupe symétrique d'un ensemble E est bien un groupe.

Notation

On note ainsi une permutation $\sigma \in S_n$: $\begin{pmatrix} 1 & 2 & \dots & n \\ \sigma(1) & \sigma(2) & \dots & \sigma(n) \end{pmatrix}$.

Exemple

Soit $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 6 & 3 & 5 & 4 & 1 \end{pmatrix}$ et $\sigma' = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 3 & 1 & 5 & 4 & 6 \end{pmatrix}$. Alors,

$$\sigma \circ \sigma' = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 3 & 2 & 4 & 5 & 1 \end{pmatrix} \qquad \sigma' \circ \sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 6 & 1 & 4 & 5 & 2 \end{pmatrix}$$

$$\sigma^3 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 2 & 3 & 5 & 4 & 6 \end{pmatrix} \qquad \sigma^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 1 & 3 & 5 & 4 & 2 \end{pmatrix}$$

Définition 18.2.

▹ L'orbite d'un élément $x \in E$ pour $\sigma \in S(E)$ est l'ensemble $\{\sigma^k(x),\ k \in \mathbb{N}\}$.
▹ Un élément $x \in E$ est fixe pour $\sigma \in S(E)$ si $\sigma(x) = x$.
En d'autres termes, un élément est donc fixe si, et seulement si, son orbite est de cardinal 1.
▹ Le support de $\sigma \in S(E)$ est l'ensemble des éléments de E qui ne sont pas fixes pour σ.

Exemple

Le support de $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 5 & 4 & 3 & 2 \end{pmatrix}$ est $\{2, 3, 4, 5\}$. Cette permutation admet trois orbites : celle du point fixe $\{1\}$ et deux autres $\{2, 5\}$ et $\{3, 4\}$.

On suppose dorénavant que E est de cardinal fini.

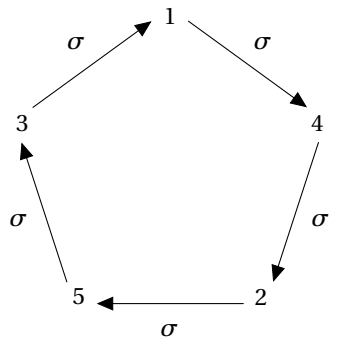
1.2. Cycles

Définition 18.3.

Une permutation circulaire de E est une permutation $\sigma \in S(E)$ qui n'admet qu'une seule orbite.

Exemple

La permutation $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 5 & 1 & 2 & 3 \end{pmatrix}$ est circulaire car $\sigma(1) = 4$, $\sigma^2(1) = \sigma(4) = 2$, $\sigma^3(1) = \sigma(2) = 5$, $\sigma^4(1) = \sigma(5) = 3$ et $\sigma^5(1) = \sigma(3) = 1$. On comprend l'appellation circulaire avec le petit croquis suivant (qui n'a aucun intérêt mathématique)



Notation

Si $\sigma \in S(E)$ est une permutation circulaire, on adopte la notation plus simple

$$\sigma = (x\ \sigma(x)\ \ldots \sigma^{n-1}(x)).$$

Remarque

Une permutation circulaire admet plusieurs écritures selon l'élément par lequel on commence à représenter l'orbite. Par exemple,

$$(1\ 3\ 2\ 4\ 5) = (4\ 5\ 1\ 3\ 2) = (5\ 1\ 3\ 2\ 4).$$

En revanche, cette permutation n'est pas égale à $(1\ 3\ 2\ 5\ 4)$.

Remarque

Un produit de permutations circulaires n'est pas forcément une permutation circulaire. Par exemple,

$$(1\ 2\ 3\ 4\ 5\ 6)^2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 4 & 5 & 6 & 1 & 2 \end{pmatrix}$$

admet deux orbites $\{1,3,5\}$ et $\{2,4,6\}$.

Définition 18.4.

▹ Un cycle de longueur l ou l-cycle de E est une permutation dont le support est de cardinal l et dont la restriction au support est une permutation circulaire de ce support.
▹ Un cycle de longueur 2 est appelé transposition.

Une transposition permute seulement deux éléments ; en particulier, une transposition est involutive.

Notation

On note un cycle comme la permutation circulaire du support correspondant. Par exemple, $(2\,3\,5)$ désigne un 3-cycle de S_{1789} (mais pourrait tout aussi bien désigner un 3-cycle de S_{1515}).

Effectuons quelques calculs de composition dans S_n.

Proposition 18.5.

Deux cycles de supports disjoints commutent.

Démonstration

Le résultat est immédiat car l'un des cycles laisse fixes les éléments du support de l'autre. ■

Exemple

Soit $\sigma = (x_1\ x_2\ \dots\ x_l)$ un cycle de longueur l et une entier $k \geq 1$.
La permutation σ^k est un cycle si, et seulement si, σ^k est un cycle de longueur l. Cette dernière condition se traduit par

$$\forall j \in [\![1, l-1]\!], \qquad (\sigma^k)^j \neq \mathrm{Id}$$

soit encore, en utilisant que σ commute avec les puissances de σ,

$$\forall j \in [\![1, l-1]\!], \qquad (\sigma^k)^j(x_1) \neq x_1.$$

En considérant les indices modulo l, $(\sigma^k)^j(x_1) = x_{1+jk}$ donc σ^k est un cycle si, et seulement si, $1 + jk \neq 1[l]$ pour tout $j \in [\![1, l-1]\!]$. D'après un raisonnement arithmétique simple, cela équivaut à la condition $k \wedge l = 1$.

Définition 18.6.

Deux permutations σ et $\sigma' \in S_n$ sont conjuguées (dans S_n) s'il existe $\rho \in S_n$ tel que

$$\sigma = \rho \circ \sigma' \circ \rho^{-1}.$$

Proposition 18.7. Conjugaison

Soit σ un élément de S_n. Alors, pour tout $x_1, \dots, x_l \in [\![1, n]\!]$ deux à deux distincts,

$$\sigma \circ (x_1\ x_2\ \dots x_l) \circ \sigma^{-1} = (\sigma(x_1)\ \sigma(x_2)\ \dots \sigma(x_l)).$$

En particulier, pour τ une transposition,

$$\tau \circ (x_1\ x_2\ \dots x_l) \circ \tau = (\tau(x_1)\ \tau(x_2)\ \dots \tau(x_l)).$$

Démonstration

▹ Un simple calcul montre qu'un élément $\sigma(x_k)$ est envoyé sur $\sigma(x_{k+1})$ par les deux permutations (avec la convention consistant à considérer les indices modulo l donc $\sigma(x_{l+1}) = \sigma(x_1)$) :

$$\sigma \circ (x_1\ x_2\ \dots x_l) \circ \sigma^{-1}(\sigma(x_k)) = \sigma \circ (x_1\ x_2\ \dots x_l)(x_k) = \sigma(x_{k+1}).$$

▹ Reste à montrer que les autres éléments sont fixes par la permutation $\sigma \circ (x_1\ x_2\ \dots x_l) \circ \sigma^{-1}$. Or, ces autres éléments sont de la forme $\sigma(y)$ avec $y \notin \{x_1, x_2, \dots, x_l\}$ et donc

$$\sigma \circ (x_1\ x_2\ \dots x_l) \circ \sigma^{-1}(\sigma(y)) = \sigma \circ (x_1\ x_2\ \dots x_l)(y) = \sigma(y).$$

■

Exemple

Montrons que toutes les transpositions sont conjuguées à la transposition $(1\ 2)$. ▹ Soit $i \neq j$. Considérons $\sigma \in S_n$ une permutation telle que $\sigma(1) = i$ et $\sigma(2) = j$. Alors,

$$\sigma \circ (1\ 2) \circ \sigma^{-1} = (i\ j).$$

▹ Donnons une démonstration alternative de cette démonstration en utilisant uniquement des transpositions.
Soit $i \neq j$. Supposons sans perte de généralité $j \neq 1$ et $i \neq 2$ (quitte à échanger les rôles de i et j) et calculons

$$(1\ j) \circ (1\ 2) \circ (1\ j) = (j\ 2) \quad \text{puis} \quad (2\ i) \circ (j\ 2) \circ (2\ i) = (i\ j).$$

Ainsi, en combinant les deux lignes, on obtient bien la relation de conjugaison

$$(i\ j) = (2\ i) \circ (1\ j) \circ (1\ 2) \circ (1\ j) \circ (2\ i)$$

car $((2\ i) \circ (1\ j))^{-1} = (1\ j) \circ (2\ i)$.

Remarque

On peut généraliser cet exemple : deux cycles de même longueur sont conjugués.

2. Décomposition en cycles disjoints

2.1. Résultat théorique

Proposition 18.8.

Soit E un ensemble fini. Toute permutation de $S(E)$, différente de l'identité, s'écrit comme composition de cycles de supports deux à deux disjoints.

Démonstration

Procédons par récurrence (forte) sur le cardinal de E.

▹ Le résultat est vrai pour les ensembles à (au plus) deux éléments.

▹ Soit $n \geq 2$. Supposons que le résultat est vrai pour tous les ensembles de cardinal au plus n et montrons-le pour un ensemble E de cardinal $n+1$ et $\sigma \in S(E)$.

Soit $x \in E$ et O l'orbite de x pour σ. Définissons un cycle s par

$$s(k) = \begin{cases} k & \text{si } k \notin O, \\ \sigma(k) & \text{sinon.} \end{cases}$$

Le support de $\sigma \circ s^{-1}$ est inclus dans $E \setminus O$ donc induit une permutation de cet ensemble (de cardinal inférieur ou égal à n). Par hypothèse de récurrence, cette permutation se décompose en une composition de cycles de supports deux à deux disjoints : $\sigma \circ s^{-1} = s_1 \circ \dots \circ s_l$. D'où $\sigma = s_1 \circ \dots \circ s_l \circ s$ et les supports sont bien deux à deux disjoints. ■

Remarque

Remarquons que cette décomposition (commutative) est unique à l'ordre près puisque chaque cycle qui apparaît correspond à l'une des orbites de la permutation que l'on décompose.

Cette preuve donne une méthode algorithmique pour obtenir la décomposition. On considère la permutation obtenue sur l'orbite du premier élément puis on se ramène à la permutation induite sur son complémentaire.

Exemple

Pour la permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 6 & 3 & 5 & 4 & 1 \end{pmatrix},$$

l'orbite de 1 est $\{1,2,6\}$ et le cycle correspondant est $(1\ 2\ 6)$. Le plus petit élément restant est 3 et il est fixe (donc seul dans son orbite). Le plus petit élément restant est 4, son orbite est $\{4,5\}$ et le cycle correspondant est $(4\ 5)$. En conclusion, $\sigma = (1\ 2\ 6)\circ(4\ 5)$.

Un simple calcul donne le résultat suivant de décomposition d'un cycle en produit de transpositions.

Proposition 18.9.

Soit $(x_1,\dots,x_l) \in [\![1,n]\!]^l$. Le cycle $(x_1\ x_2\ \dots x_l)$ se décompose des deux manières suivantes en composition de $l-1$ transpositions :

$$(x_1\ x_l)\circ(x_1\ x_{l-1})\circ\dots\circ(x_1\ x_2), \qquad (x_1\ x_2)\circ(x_2\ x_3)\circ\dots\circ(x_{l-1}\ x_l).$$

D'où l'on déduit facilement à l'aide de la décomposition en cycles le résultat suivant.

Proposition 18.10.

Toute permutation se décompose comme une composition de transpositions.
Autrement dit, les transpositions engendrent le groupe S_n.

Remarque

Par rapport à la décomposition précédente en cycle à support disjoints, l'écriture d'une permutation comme produit de transpositions n'est *a priori* pas commutative.

Corollaire 18.11.

La famille $\{(1\ k),\ k \in [\![2,n]\!]\}$ engendre le groupe S_n.

Démonstration

D'après la proposition précédente, il suffit de montrer que l'on peut obtenir toute transposition à partir de cette famille. Ceci découle immédiatement du calcul

$$(i\ j) = (1\ i)\circ(1\ j)\circ(1\ i),$$

pour tous $i \neq j$ (distincts de 1 sinon, il n'y a rien à montrer). ■

2.2. Ordre d'une permutation

Commençons par un rappel.

Définition 18.12.

L'ordre d'une permutation $\sigma \in S_n$ est le plus petit entier $k > 0$ tel que $\sigma^k = \mathrm{Id}$ (ou de manière équivalente, l'ordre du sous-groupe monogène engendré par σ).

Exemple

L'ordre d'un cycle est sa longueur. En effet, si $\sigma = (x_1\ x_2\ \dots\ x_l)$, alors,
- pour tout $k < l$, $\sigma^k(x_1) = x_{k+1} \neq x_1$ donc $\sigma^k \neq \mathrm{Id}$,
- $\sigma^l = \mathrm{Id}$.

Proposition 18.13.

Soit $\sigma \in S_n$. L'ordre de σ est le ppcm des ordres des cycles qui apparaissent dans sa décomposition en cycles à supports disjoints.

Démonstration

Notons $c_1, \dots, c_p$ les cycles à supports disjoints apparaissant dans la décomposition de σ. Comme l'écriture $\sigma = c_1 \circ c_2 \circ \dots \circ c_p$ est commutative, on obtient pour tout $k \in \mathbb{N}$, $\sigma^k = c_1^k \circ c_2^k \circ \dots \circ c_p^k$.

Les permutations $c_1^k, c_2^k, \dots, c_p^k$ sont encore à supports disjoints (mais ne sont plus forcément des cycles). Par conséquent, $\sigma^k = \mathrm{Id}$ si, et seulement si, pour tout $i \leq p$, $c_i^k = \mathrm{Id}$ soit k est un multiple de la longueur de c_i. En conclusion, l'ordre de σ est le plus petit commun multiple des longueurs des cycles c_i pour tout $i \leq p$. ■

Exemple

Pour la permutation $\sigma \in S_8$ définie par

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 2 & 1 & 8 & 5 & 7 & 4 & 6 & 3 \end{pmatrix},$$

la décomposition en cycles à supports disjoints est $\sigma = (1\ 2) \circ (3\ 8) \circ (4\ 5\ 7\ 6)$; par conséquent, son ordre est $2 \vee 2 \vee 4 = 4$.

Remarque

D'après la proposition précédente, on obtient qu'une permutation est d'ordre 2 (respectivement d'ordre p premier) si, et seulement si, elle peut s'écrire comme produit de transpositions (respectivement de p-cycles) à supports disjoints.

Remarque

L'ordre d'un élément de S_n peut être plus grand que n.

Par exemple, la permutation $(1\ 2) \circ (3\ 4\ 5) \in S_5$ est, d'après la proposition précédente, d'ordre égal à $2 \vee 3 = 6$.

En revanche, d'après le théorème de Lagrange, l'ordre d'un élément de S_n divise $|S_n| = n!$.

3. Signature

3.1. Définition

Définition 18.14.

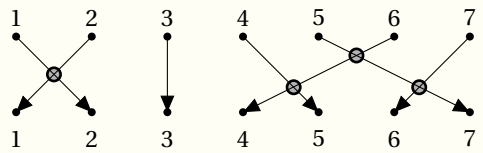
Soit $\sigma \in S_n$.
▹ Une paire $\{i, j\}$ est une inversion pour σ si $(i-j)(\sigma(i)-\sigma(j)) < 0$.
On note $\mathrm{Inv}(\sigma)$ le nombre d'inversions de σ.
▹ La signature de σ est le réel $\varepsilon(\sigma) = (-1)^{\mathrm{Inv}(\sigma)}$.

Remarque

Pour voir facilement le nombre d'inversions d'une permutation $\sigma \in S_n$, on peut adopter une représentation graphique adaptée. On place sur deux lignes des sommets représentant les entiers de $[\![1, n]\!]$ (dans l'ordre usuel) et on crée une arête entre le sommet i de la première ligne et le sommet j de la seconde si, et seulement si, $\sigma(i) = j$. Le nombre d'inversions est alors le nombre d'intersections entre les arêtes. Par exemple, pour la permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 2 & 1 & 3 & 5 & 7 & 4 & 6 \end{pmatrix},$$

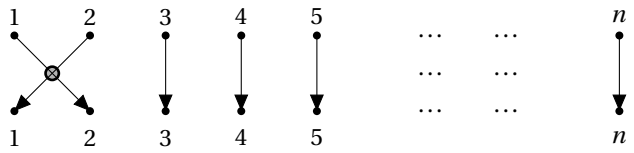
le graphique est



Il y a donc 4 inversions et la signature est $(-1)^4 = 1$.

Exemple

Si $\tau = (1\ 2)$, alors $\varepsilon(\tau) = -1$ puisque τ n'admet qu'une seule inversion.



Remarque

Comme, pour toute permutation σ, $\mathrm{Inv}(\sigma) = \mathrm{Inv}(\sigma^{-1})$, on en déduit que $\varepsilon(\sigma) = \varepsilon(\sigma^{-1})$.

3.2. Propriété de morphisme

Mentionnons une expression explicite de la signature qui permet une utilisation théorique.

Proposition 18.15.

Pour tout $\sigma \in S_n$,

$$\varepsilon(\sigma) = \prod_{\{i,j\}} \frac{\sigma(i)-\sigma(j)}{i-j},$$

où la notation $\prod_{\{i,j\}}$ signifie le produit sur toutes les paires $\{i,j\}$ (donc avec $i \neq j$).

Remarque

Quitte à ordonner les paires qui apparaissent dans le produit, la signature d'une permutation $\sigma \in S_n$ est aussi

$$\varepsilon(\sigma) = \prod_{i>j} \frac{\sigma(i)-\sigma(j)}{i-j}.$$

Démonstration

Soit $\sigma \in S_n$ et étudions la quantité $\prod_{\{i,j\}} \frac{\sigma(i)-\sigma(j)}{i-j}$. Commençons par évacuer le problème du signe : par définition des inversions,

$$\prod_{\{i,j\}} \frac{\sigma(i)-\sigma(j)}{i-j} = (-1)^{\mathrm{Inv}(\sigma)} \prod_{\{i,j\}} \frac{|\sigma(i)-\sigma(j)|}{|i-j|}.$$

Or, σ est une permutation de $[\![1,n]\!]$ donc $\{\sigma(i),\sigma(j)\}$ décrit l'ensemble des paires de $[\![1,n]\!]$ lorsque $\{i,j\}$ décrit l'ensemble des paires de $[\![1,n]\!]$. Par conséquent,

$$\prod_{\{i,j\}} |\sigma(i)-\sigma(j)| = \prod_{\{i,j\}} |i-j|,$$

et donc

$$\prod_{\{i,j\}} \frac{\sigma(i)-\sigma(j)}{i-j} = (-1)^{\mathrm{Inv}(\sigma)}.$$

■

Proposition 18.16.

L'application signature est un morphisme de groupes de S_n dans $\{-1,1\}$, c'est-à-dire

$$\forall \sigma_1, \sigma_2 \in S_n, \qquad \varepsilon(\sigma_1 \circ \sigma_2) = \varepsilon(\sigma_1)\varepsilon(\sigma_2).$$

En particulier, deux permutations σ_1, σ_2 conjuguées ont la même signature.

Démonstration

Avec l'expression de la signature obtenue précédemment, pour tous $\sigma_1, \sigma_2 \in S_n$,

$$\varepsilon(\sigma_1 \circ \sigma_2) = \prod_{\{i,j\}} \frac{\sigma_1 \circ \sigma_2(i) - \sigma_1 \circ \sigma_2(j)}{i-j} = \left[\prod_{\{i,j\}} \frac{\sigma_1 \circ \sigma_2(i) - \sigma_1 \circ \sigma_2(j)}{\sigma_2(i)-\sigma_2(j)}\right] \cdot \left[\prod_{\{i,j\}} \frac{\sigma_2(i)-\sigma_2(j)}{i-j}\right].$$

Par ailleurs, comme σ_2 est une bijection,

$$\prod_{\{i,j\}} \frac{\sigma_1 \circ \sigma_2(i) - \sigma_1 \circ \sigma_2(j)}{\sigma_2(i)-\sigma_2(j)} = \prod_{\{i,j\}} \frac{\sigma_1(i)-\sigma_1(j)}{i-j} = \varepsilon(\sigma_1).$$

En conclusion, $\varepsilon(\sigma_1 \circ \sigma_2) = \varepsilon(\sigma_1)\varepsilon(\sigma_2)$. ■

Montrons comment calculer en pratique la signature d'une permutation.

Exemple

▹ Rappelons que toutes les transpositions sont conjuguées à $(1\,2)$. Par conséquent, toutes les transpositions sont de signature $\varepsilon((1\,2)) = -1$.
▹ En utilisant la décomposition d'un cycle de longueur l en produit de $l-1$ transpositions, on obtient qu'un cycle de longueur l a pour signature $(-1)^{l-1}$.
▹ En utilisant la décomposition en cycles à supports disjoints et la propriété de morphisme, on obtient la signature de toute permutation.
Par exemple, la permutation $\sigma = (1\,5\,7\,8) \circ (2\,9\,3) \circ (4\,6)$ a pour signature

$$\varepsilon(\sigma) = \varepsilon((1\,5\,7\,8))\varepsilon((2\,9\,3))\varepsilon((4\,6)) = (-1)^{4-1}(-1)^{3-1}(-1)^{2-1} = 1.$$

On obtient la caractérisation suivante de l'application signature.

Proposition 18.17.

Soit $\varphi : S_n \to \mathbb{C}^*$ un morphisme de groupes différent de l'application constante égale à 1. Alors, φ est la signature.

Démonstration
Remarquons que φ attribue la même valeur à toutes les transpositions (car elles sont conjuguées).
Soit τ une transposition ; comme $\varphi(\tau)^2 = \varphi(\tau^2) = 1$, on obtient $\varphi(\tau) \in \{\pm 1\}$.

- Si $\varphi(\tau) = 1$, alors φ et la fonction constante égale à 1 coïncident sur les transpositions et toute permutation est un produit de transpositions donc $\varphi = 1$.
- Si $\varphi(\tau) = -1$, alors φ et la signature coïncident sur les transpositions et toute permutation est un produit de transpositions donc $\varphi = \varepsilon$.

■

3.3. Groupe alterné

Définition 18.18.

Le groupe alterné A_n est le noyau de la signature sur S_n.

Il est clair que A_n est un sous-groupe de S_n en tant que noyau d'un morphisme de groupes.

Proposition 18.19.

Le sous-groupe A_n est d'ordre d'ordre $\frac{n!}{2}$.

Démonstration
Soit τ une transposition. L'application $\varphi : \left\{ \begin{array}{ccc} A_n & \to & S_n \setminus A_n \\ \sigma & \mapsto & \sigma \circ \tau \end{array} \right.$ est une bijection (on connaît la bijection réciproque) d'où $|A_n| = |S_n \setminus A_n| = \frac{n!}{2}$. ■

Proposition 18.20.

Le sous-groupe A_n vérifie

$$\forall \sigma \in A_n,\ \forall \sigma' \in S_n, \qquad \sigma' \circ \sigma \circ \sigma'^{-1} \in A_n.$$

Autrement dit, A_n est stable par conjugaison par tout élément de S_n.

Démonstration

Soit $\sigma \in A_n$ et $\sigma' \in S_n$. Alors, $\varepsilon(\sigma' \circ \sigma \circ \sigma'^{-1}) = \varepsilon(\sigma) = 1$ et donc $\sigma' \circ \sigma \circ \sigma'^{-1} \in A_n$. ■

Exemple

Détaillons quelques groupes alternés de petits indices.

▹ Le groupe A_3 est composé de Id et des deux permutations circulaires $(1\ 2\ 3)$ et $(1\ 3\ 2)$.

▹ Le groupe A_4 contient Id, les 3-cycles $(1\ 2\ 3), (1\ 3\ 2), (1\ 2\ 4), (1\ 4\ 2), (1\ 3\ 4), (1\ 4\ 3), (2\ 3\ 4), (2\ 4\ 3)$ et les « bitranspositions » $(1\ 2)\circ(3\ 4), (1\ 3)\circ(2\ 4), (1\ 4)\circ(2\ 3)$; remarquons que $V_4 = \{\mathrm{Id}, (1\ 2)\circ(3\ 4), (1\ 3)\circ(2\ 4), (1\ 4)\circ(2\ 3)\}$ est un sous-groupe distingué de A_4.

▹ Le groupe A_5 est composé de 60 éléments qui se répartissent en :

ordre	nombre	description
1	1	Identité
2	$\frac{1}{2}\binom{5}{4}\binom{4}{2} = 15$	produits de deux transpositions à supports disjoints
3	$2.\binom{5}{3} = 20$	3-cycles
4	0	
5	$4! = 24$	5-cycles

Remarquons que A_n ne contient aucune transposition puisque celles-ci sont de signature -1.

Proposition 18.21.

La famille $\{(1\ 2\ k),\ k \in [\![3, n]\!]\}$ engendre A_n.

Démonstration

▹ Tout élément de A_n s'écrit comme la composition d'un nombre pair de transpositions (car les transpositions engendrent S_n et la signature doit être 1) ; plus précisément, tout élément de A_n s'écrit comme un nombre pair de transpositions de la forme $(1\ x)$ (en utilisant que cette famille engendre S_n).

▹ Il suffit de vérifier que toute composition de deux telles transpositions est obtenue avec les 3-cycles de la famille considérée.

Or, pour tous $x,\ y \geq 3$ différents, on obtient la relation suivante par un simple calcul

$$(1\ x)\circ(1\ 2) = (1\ 2\ x),$$

et on en déduit

$$(1\ 2)\circ(1\ y) = ((1\ y)\circ(1\ 2))^{-1} = (1\ 2\ y)^{-1},$$
$$(1\ x)\circ(1\ y) = ((1\ x)\circ(1\ 2))\circ((1\ 2)\circ(1\ y)) = (1\ 2\ x)\circ(1\ 2\ y)^2.$$

■

FICHE

SYNTHÈSE

Le groupe symétrique est un riche exemple de groupe fini (non abélien).

Définitions

▹ Une permutation de $[\![1,n]\!]$ est une application de $[\![1,n]\!]$ dans $[\![1,n]\!]$ bijective. Le groupe symétrique d'ordre n est l'ensemble des permutations de $[\![1,n]\!]$ muni de la composition. Il est noté S_n.
▹ Une permutation circulaire est une permutation qui n'admet qu'une seule orbite.
▹ Un cycle de longueur l ou l-cycle est une permutation dont le support est de cardinal l et dont la restriction au support est une permutation circulaire de ce support.
▹ Un cycle de longueur 2 est appelé transposition.

Un résultat important est la décomposition (commutative suivante). Elle permet entre autres de calculer les puissances, l'ordre, la conjugaison, la signature... d'une permutation.

Décomposition en cycles à supports disjoints

Toute permutation de S_n, différente de l'identité, s'écrit comme composition de cycles de supports deux à deux disjoints.

Un morphisme s'avère particulièrement important : la signature.

Signature

La signature ε est l'unique morphisme de groupes non constant de S_n dans $\mathbb{C}^*$. Son noyau est le sous-groupe alterné.

Au passage, on a établi l'existence de parties génératrices simples.

Parties génératrices

▹ Le groupe S_n est engendré par la famille de transpositions $\{(1\ k),\ k \in [\![2,n]\!]\}$.
▹ Le groupe A_n est engendré par la famille de 3-cycles $\{(1\ 2\ k),\ k \in [\![3,n]\!]\}$.

EXERCICES

Vrai ou faux ?

		Vrai	Faux
a)	$(1\ 2\ 3\ 4) = (4\ 1\ 2\ 3)$.	☐	☐
b)	$(1\ 2\ 3)^{-1} = (3\ 2\ 1)$.	☐	☐
c)	$((1\ 2)\circ(1\ 3))^{-1} = (1\ 2)\circ(1\ 3)$.	☐	☐
d)	Pour $i \neq j$, $(1\ i)\circ(1\ j) = (1\ j\ i)$.	☐	☐
e)	Deux permutations commutent si, et seulement si, elles sont à supports disjoints.	☐	☐
f)	Le produit de deux cycles à supports non disjoints est un cycle.	☐	☐
g)	Pour tout $\sigma \in S_n$, il existe $k \in [\![1, n]\!]$ tel que $\sigma^k = \mathrm{Id}$.	☐	☐
h)	L'ordre d'une permutation est la maximum des longueurs des cycles dans sa décomposition en cycles à supports disjoints.	☐	☐
i)	La signature de $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 5 & 4 & 1 & 2 & 6 \end{pmatrix}$ est 1.	☐	☐
j)	La signature d'un l-cycle est $(-1)^{l+1}$.	☐	☐
k)	Deux permutations conjuguées ont la même signature.	☐	☐

EXERCICES

Exercices d'application du chapitre 18

●○○ **Exercice 1**

1. Déterminer la décomposition en cycles à supports disjoints et la signature des permutations suivantes :

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 9 & 4 & 3 & 8 & 7 & 10 & 1 & 2 & 5 & 6 \end{pmatrix}$$

$$\sigma' = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 9 & 4 & 3 & 8 & 7 & 6 & 1 & 2 & 5 \end{pmatrix}.$$

2. Préciser σ^{1789} et σ'^{1789}.

●○○ **Exercice 2**

Montrer que les permutations de S_n qui commutent avec $\sigma = (1\ 2\ 3\ \dots\ n)$ sont les puissances de σ.

●●○ **Exercice 3**

Déterminer toutes les permutations de S_6 qui commutent avec $\sigma = (1\ 2\ 3) \circ (4\ 5\ 6)$.

●○○ **Exercice 4**

Montrer que l'ensemble G des permutations $\sigma \in S_n$ telles que

$$\forall k \in [\![1, n]\!], \qquad \sigma(n+1-k) + \sigma(k) = n+1$$

est un sous-groupe de S_n.

●○○ **Exercice 5**

Soit $\sigma \in S_n$. Déterminer le nombre de parties $A \subset [\![1, n]\!]$ telles que $\sigma(A) = A$.

●○○ **Exercice 6**

Soit $A \subset [\![1, n]\!]$ de cardinal p. Déterminer le nombre de permutations $\sigma \in S_n$ telles que $\sigma(A) = A$.

●●○ **Exercice 7**

1. Montrer que la famille $\{(k\ k+1),\ k \in [\![1, n]\!]\}$ engendre le groupe S_n.
2. Montrer que la famille $\{(1\ 2), (1\ 2\ 3\ \dots\ n)\}$ engendre le groupe S_n.

●●● **Exercice 8 Théorème de Cayley**

Soit G un groupe d'ordre n. Montrer que G est isomorphe à un sous-groupe de S_n.

●○○ **Exercice 9**

Déterminer la signature de la permutation $\sigma \in S_n$ vérifiant $\sigma(k) = n+1-k$ pour tout $k \in [\![1, n]\!]$.

●●○ **Exercice 10**

Exprimer la signature d'une permutation de S_n en fonction de son nombre d'orbites (et de n).

●●○ **Exercice 11**

Pour tout $\sigma \in S_n$, définissons la matrice $P_\sigma \in \mathscr{M}_n(\mathbb{R})$ dont le coefficient en position i, j est 1 si $i = \sigma(j)$, 0 sinon.

1. Calculer, pour tous $\sigma, \rho \in S_n$, $P_\sigma P_\rho$.
2. Justifier que, pour tout $\sigma \in S_n$, $P_\sigma \in \mathrm{GL}_n(\mathbb{R})$.
3. En déduire que l'application $\sigma \mapsto P_\sigma$ est un morphisme de groupes injectif de S_n dans $\mathrm{GL}_n(\mathbb{R})$.

●●○ **Exercice 12**

Pour tout $\sigma \in S_n$, définissons l'endomorphisme u_σ de $\mathbb{R}_n$ défini sur la base canonique $(e_1, \dots, e_n)$ par

$$\forall j \in [\![1, n]\!], \qquad u_\sigma(e_j) = e_{\sigma(j)}.$$

Posons enfin

$$p = \frac{1}{n!} \sum_{\sigma \in S_n} u_\sigma.$$

1. Montrer que p est un projecteur de $\mathbb{R}^n$.
2. Montrer que l'image de p est engendrée par le vecteur $e_1 + e_2 + \dots + e_n$.
3. Déterminer le noyau de p.

Exercices d'approfondissement du chapitre 18

●●○ **Exercice A**

1. Soit $n \geq 4$. Montrer que le groupe S_n est engendré par les 4-cycles.
2. Soit $n \geq p$. Déterminer le sous-groupe engendré par les p-cycles.
   Indication : on pourra calculer $(a_2\ a_1\ a_3\ a_4\ \dots\ a_p) \circ (a_2\ a_1\ a_p\ a_{p-1}\ \dots\ a_3)$.

●●○ **Exercice B**

Montrer qu'il existe un morphisme de groupes injectif de S_n dans A_{n+2}.

●●● **Exercice C**

Soit $(x_k)_{k \leq n}$ et $(y_k)_{k \leq n}$ deux suites finies strictement décroissantes. Posons, pour tout $\sigma \in S_n$,

$$\Sigma_\sigma = \sum_{k=1}^{n} x_k\, y_{\sigma(k)}.$$

Montrer que, pour tout $\sigma \in S_n$, $\Sigma_{\mathrm{Id}} \geq \Sigma_\sigma$.

Problèmes du chapitre 18

●○○ **Problème 1 Jeu de taquin**

Le taquin est un jeu solitaire composé de 15 petits carreaux numérotés de 1 à 15 qui glissent dans un cadre carré prévu pour 16. Il consiste à remettre dans l'ordre les 15 carreaux à partir d'une configuration initiale. Dans cet exercice, nous considérons la configuration initiale suivante (où 16 représente l'emplacement vide) :

1	2	3	4
5	6	7	8
9	10	11	12
13	15	14	16

Les mouvements autorisés consistent à échanger la case 16 avec l'une de ses voisines (horizontalement ou verticalement). Le but du jeu est d'obtenir, après une succession de mouvements autorisés, la configuration finale suivante

1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	16

On définit une permutation de S_{16} pour chaque configuration S en associant à la case i dans la configuration finale le numéro l'occupant dans la configuration S. Ainsi, la configuration finale est associée à l'identité, la configuration initiale à la transposition (14 15).

1. Déterminer la permutation associée à la configuration

2	5	3	8
9	6	7	4
15	10	11	12
16	14	1	13

PROBLÈMES

2. Montrer qu'une solution comporte un nombre impair de mouvements.
3. Relier la somme des indices d'abscisse et d'ordonnée de la case 16 avec le nombre de mouvements dans une solution.
4. Conclure.

Problème 2 Simplicité de A_5

Définition 18.22.

Un sous-groupe H de S_5 (respectivement A_5) est distingué dans S_5 (respectivement A_5) s'il contient tous les conjugués de ses éléments par ceux de S_5 (respectivement A_5).

Soit H un sous-groupe distingué dans S_5.

1. Montrer que si $H \cap A_5 = A_5$, alors $H = A_5$ ou $H = S_5$.
2. Montrer que si $H \cap A_5 = \{\mathrm{Id}\}$, alors $H = \{\mathrm{Id}\}$.
   Indication : on peut commencer par remarquer que la restriction de la signature à H est injective et en déduire que $\operatorname{card} H \leq 2$.
3. a. Déterminer le nombre d'éléments de chaque ordre dans A_5.
   b. Montrer que si H contient un élément d'ordre 2 (respectivement 3 ou 5) de A_5, alors il les contient tous.
   c. Déterminer $H \cap A_5$.
4. En déduire tous les sous-groupes distingués de S_5.
5. Précisons le résultat de la question 3 en déterminant les sous-groupes distingués de A_5.
   Soit K un sous-groupe distingué de A_5.
   a. Montrer que si K contient un élément d'ordre 2 (respectivement 3) de A_5, alors il les contient tous.
   b. Soit σ et ρ deux éléments d'ordre 5 de A_5. Montrer que σ est un conjugué de ρ ou de ρ^2 par un élément de A_5.
   En déduire que si K contient un élément d'ordre 5 de A_5, alors il les contient tous.
   c. Déterminer les sous-groupes distingués de A_5.

CORRIGÉS

Corrigés des Vrai/Faux

a) Vrai.

b) Vrai.

c) Faux, $(1\ 2\ 3) \neq (3\ 2\ 1)$.

d) Vrai.

e) Vrai.

f) Faux : par exemple, $(1\ 2\ 3\ 4)^2 = (1\ 3) \circ (2\ 4)$.

g) Faux, $(1\ 2) \circ (3\ 4\ 5) \in S_5$ mais aucune puissance strictement inférieure à 6 ne donne l'identité.

h) Faux, $(1\ 2) \circ (3\ 4\ 5)$ n'est pas d'ordre 3.

i) Faux.

j) Vrai, $(-1)^{l-1} = (-1)^{l+1}$.

k) Vrai.

Corrigés des exercices d'application du chapitre 18

Exercice 1

1. ▹ Sous l'action de σ, 1 est envoyé sur 9, qui est envoyé sur 5, qui est envoyé sur 7, qui est envoyé sur 1 : le premier cycle de la décomposition de σ est donc $(1\ 9\ 5\ 7)$.
Le plus petit entier encore non considéré est 2 : 2 est envoyé sur 4, qui est envoyé sur 8, qui est envoyé sur 2 : le cycle suivant dans la décomposition est $(2\ 4\ 8)$.
Le plus petit entier encore non considéré est 3 qui est fixe.
Le plus petit entier encore non considéré est 6 : 6 est envoyé sur 10 , qui est envoyé sur 6 : le cycle suivant est la transposition $(6\ 10)$.
La décomposition de σ est

$$\sigma = (1\ 9\ 5\ 7) \circ (2\ 4\ 8) \circ (6\ 10).$$

▹ On procède de même pour σ' et on trouve

$$\sigma' = (1\ 9\ 5\ 7) \circ (2\ 4\ 8).$$

2. Exploitons les décompositions précédentes en notant que d'une part, ces décompositions sont commutatives (car les cycles sont à supports disjoints) et d'autre part que la puissance $l^{\text{ième}}$ d'un l-cycle est

l'identité. Ainsi,

$$\begin{aligned}\sigma^{1789} &= (1\ 9\ 5\ 7)^{1789} \circ (2\ 4\ 8)^{1789} \circ (6\ 10)^{1789}\\ &= (1\ 9\ 5\ 7)^{1} \circ (2\ 4\ 8)^{1} \circ (6\ 10)^{1} = \sigma,\\ \sigma'^{1789} &= (1\ 9\ 5\ 7)^{1789} \circ (2\ 4\ 8)^{1789}\\ &= (1\ 9\ 5\ 7)^{1} \circ (2\ 4\ 8)^{1} = \sigma'.\end{aligned}$$

Exercice 2

Soit $\rho \in S_n$. Alors, ρ commute avec σ si, et seulement si, $\rho \circ \sigma = \sigma \circ \rho$ ou encore $\rho \circ \sigma \circ \rho^{-1} = \sigma$. D'après la calcul de la conjugaison sur un cycle, cette condition se réécrit

$$(\rho(1)\ \rho(2)\ \dots\ \rho(n)) = (1\ 2\ \dots\ n).$$

Notons $k = \rho(1)$. Une récurrence immédiate donne, pour tout $i \leq n$, $\rho(i) = k + i - 1[n]$ d'après l'égalité précédente. Ainsi, $\rho = \sigma^{k-1}$.
En conclusion, une permutation commute avec σ est une puissance de σ (la réciproque est immédiate).

Exercice 3

Soit une permutation $f \in S_6$ telle que $f \circ \sigma = \sigma \circ f$. Alors, $f \circ \sigma \circ f^{-1} = \sigma$. Donc

$$\begin{aligned}(1\ 2\ 3) \circ (4\ 5\ 6) &= f \circ (1\ 2\ 3) \circ (4\ 5\ 6) \circ f^{-1}\\ &= f \circ (1\ 2\ 3) \circ f^{-1} \circ f \circ (4\ 5\ 6) \circ f^{-1}\\ &= (f(1)\ f(2)\ f(3)) \circ (f(4)\ f(5)\ f(6)).\end{aligned}$$

En utilisant l'unicité de la décomposition en cycles à supports disjoints, on obtient deux cas de figure
- $(f(1)\ f(2)\ f(3)) = (1\ 2\ 3)$ et $(f(4)\ f(5)\ f(6)) = (4\ 5\ 6)$ qui donne 9 permutations (3 choix pour $f(1)$ et 3 choix pour $f(4)$), $f_1, \dots, f_9$, décrites dans le premier tableau.
- $(f(1)\ f(2)\ f(3)) = (4\ 5\ 6)$ et $(f(4)\ f(5)\ f(6)) = (1\ 2\ 3)$ qui donne 9 permutations, $f'_1, \dots, f'_9$, décrites dans le second tableau.

k	1	2	3	4	5	6
$f_1(k)$	1	2	3	4	5	6
$f_2(k)$	1	2	3	5	6	4
$f_3(k)$	1	2	3	6	4	5
$f_4(k)$	2	3	1	4	5	6
$f_5(k)$	2	3	1	5	6	4
$f_6(k)$	2	3	1	6	4	5
$f_7(k)$	3	1	2	4	5	6
$f_8(k)$	3	1	2	5	6	4
$f_9(k)$	3	1	2	6	4	5

k	1	2	3	4	5	6
$f'_1(k)$	4	5	6	1	2	3
$f'_2(k)$	5	6	4	1	2	3
$f'_3(k)$	6	4	5	1	2	3
$f'_4(k)$	4	5	6	2	3	1
$f'_5(k)$	5	6	4	2	3	1
$f'_6(k)$	6	4	5	2	3	1
$f'_7(k)$	4	5	6	3	1	2
$f'_8(k)$	5	6	4	3	1	2
$f'_9(k)$	6	4	5	3	1	2

Exercice 4

Utilisons la caractérisation de sous-groupe.
▹ L'ensemble G est une partie non vide de S_n car il contient la permutation Id.
▹ Soit $\sigma \in G$, $k \in [\![1, n]\!]$. Notons $k' = \sigma^{-1}(k)$. Alors,

$$\begin{aligned}\sigma^{-1}(n+1-k) &= \sigma^{-1}(n+1-\sigma(k')\\ &= \sigma^{-1}(\sigma(n+1-k')) && \text{car } \sigma \in G\\ &= n+1-k' = n+1-\sigma^{-1}(k).\end{aligned}$$

Ainsi, $\sigma^{-1} \in G$.
▹ Soit σ, $\rho \in G$ et $k \in [\![1, n]\!]$. Alors,

$$\begin{aligned} \sigma \circ \rho(n+1-k) &= \sigma(n+1-\rho(k)) && \text{car } \rho \in G \\ &= n+1-\sigma(\rho(k)) && \text{car } \sigma \in G \\ &= n+1-\sigma \circ \rho(k). \end{aligned}$$

Ainsi, $\sigma \circ \rho \in G$.

Exercice 5

Il est clair que si la partie A contient un entier k, elle contient également l'entier $\sigma(k)$ et par une récurrence immédiate, elle contient toute l'orbite de k.
Ainsi, une partie A vérifie $\sigma(A) = A$ si, et seulement si, elle est réunion d'orbites de σ. Si σ admet o orbites, le nombre recherché est 2° (le nombre de parties de l'ensemble des orbites).

Exercice 6

Une telle permutation σ est uniquement déterminée par
- une permutation de A (et il y en a $p!$),
- une permutation de $[\![1, n]\!] \setminus A$ (et il y en a $(n-p)!$).

Ainsi, il y a donc $p!(n-p)!$ permutations σ telles que $\sigma(A) = A$.

Exercice 7

1. D'après le cours, il suffit de montrer que l'on peut atteindre toute transposition de la forme $(1\ k)$ à partir de cette famille (car cette famille est génératrice de S_n).
Fixons $\in [\![2, n]\!]$ et calculons successivement

$$\begin{gathered} (k-2\ k-1) \circ (k-1\ k) \circ (k-2\ k-1) = (k-2\ k) \\ (k-3\ k-2) \circ (k-1\ k) \circ (k-3\ k-2) = (k-3\ k) \\ \vdots \\ (1\ 2) \circ (2\ k) \circ (1\ 2) = (1\ k) \end{gathered}$$

On obtient donc que les transpositions $(k-2\ k)$, $(k-3\ k)$, ... et finalement $(1\ k)$ appartiennent bien au groupe engendré par la famille considérée.
2. D'après la question précédente, il suffit de montrer que l'on peut atteindre toute transposition de la forme $(k\ k+1))$ à partir de cette famille (car cette famille est génératrice de S_n) ce qui est élémentaire par conjugaison :

$$(1\ 2\ 3\ \dots\ n)^k \circ (1\ 2) \circ (1\ 2\ 3\ \dots\ n)^{-k} = (k\ k+1).$$

Remarque

Avec un raisonnement analogue à la solution de la deuxième question, on montre que S_n est engendré par la famille composée de la transposition $(1\ 2)$ et du $(n-1)$-cycle $(2\ 3\ \dots\ n)$.

Exercice 8

Nous allons montrer que G est isomorphe à un sous-groupe de $S(G)$ ce qui suffit puisque $S(G)$ est lui-même isomorphe à S_n.
Pour tout $g \in G$, posons

$$\varphi_g : \left\{ \begin{array}{ccc} G & \to & G \\ x & \mapsto & x \star g \end{array} \right.$$

Remarquons que, pour tous g et $g' \in G$, $\varphi_g \circ \varphi_{g'} = \varphi_{g \star g'}$. Par conséquent, pour tout $g \in G$, $\varphi_g \in S(G)$ (de réciproque $\varphi_{g^{-1}}$). De plus,

$$\varphi : \left\{ \begin{array}{ccc} G & \to & S(G) \\ g & \mapsto & \varphi_g \end{array} \right.$$

est un morphisme de groupes.
Pour conclure, il suffit de vérifier que φ est injective : $g \in \operatorname{Ker} \varphi$ équivaut à $\varphi_g = \mathrm{Id}$ soit, en évaluant en e_G, à $g = e_G$.

Exercice 9

Discutons selon la parité de n.
▹ Si $n = 2p$ avec $p \in \mathbb{N} \setminus \{0\}$, alors σ est la composition de p transpositions (à supports disjoints) :

$$\sigma = (1\ 2p) \circ (2\ 2p-1) \circ \ldots (p\ p+1)$$

donc $\varepsilon(\sigma) = (-1)^p$.
▹ Si $n = 2p+1$ avec $p \in \mathbb{N}$, alors σ laisse fixe l'élément $p+1$ et échange les $2p$ autres éléments : ainsi, σ est la composition de p transpositions (à supports disjoints)

$$\sigma = (1\ 2p+1) \circ (2\ 2p) \circ \ldots (p\ p+2)$$

donc $\varepsilon(\sigma) = (-1)^p$.
Dans tous les cas, on obtient que $\varepsilon(\sigma) = (-1)^{\lfloor \frac{n}{2} \rfloor}$.

Exercice 10

Soit $\sigma \in S_n$. Notons $l_1, \ldots, l_r$ les longueurs des différents cycles apparaissant dans la décomposition en cycles à supports disjoints de σ. D'après la propriété de morphisme de ε,

$$\varepsilon(\sigma) = \prod_{k=1}^{r} (-1)^{l_k - 1} = (-1)^{\sum_{k=1}^{r} l_k - r}.$$

Remarquons ensuite que $\sum_{k=1}^{r} l_k$ est la taille du support de σ donc $n - f$ où f est le nombre de points fixes de σ. Ainsi, $\varepsilon(\sigma) = (-1)^{n-r-f}$. Pour conclure, il suffit de remarquer que $n+f$ est le nombre o d'orbites de σ et donc $\varepsilon(\sigma) = (-1)^{n+o}$.

Exercice 11

1. Soit $\sigma, \rho \in S_n$. Le coefficient en position i, j de $P_\sigma P_\rho$ est

$$\sum_{k=1}^{n} [P_\sigma]_{i,k} [P_\rho]_{k,j} = [P_\sigma]_{i,\rho(j)}.$$

D'après la définition de P_σ, ce coefficient est 1 si $i = \sigma(\rho(j))$, 0 sinon.
En conclusion, $P_\sigma P_\rho = P_{\sigma \circ \rho}$.
2. Soit $\sigma \in S_n$. D'après la question précédente, $P_\sigma P_{\sigma^{-1}} = P_{\mathrm{Id}} = \mathrm{I}_n$ et $P_{\sigma^{-1}} P_\sigma = \mathrm{I}_n$. Par conséquent, P_σ est inversible d'inverse $P_{\sigma^{-1}}$.
3. La propriété de morphisme est établie à la première question. Montrons l'injectivité en considérant une permutation σ du noyau : si $P_\sigma = \mathrm{I}_n$, alors pour tout j, $\sigma(j) = j$ donc $\sigma = \mathrm{Id}$.

Remarque

Avec un argument supplémentaire (par exemple, la notion de polynôme caractéristique d'une matrice), on obtient que deux permutations $\sigma, \rho \in S_n$ sont conjuguées si, et seulement si, les matrices P_σ, P_ρ sont semblables.

Exercice 12

On vérifie que, pour tous σ et $\rho \in S_n$, $u_\sigma \circ u_\rho = u_{\sigma\circ\rho}$ (car ces endomorphismes coincident sur la base canonique).

1. Il est clair que p est un endomorphisme comme combinaison linéaire des endomorphismes u_σ. Calculons p^2

$$\begin{aligned} p^2 &= \frac{1}{n!^2}\sum_{\sigma\in S_n}\sum_{\rho\in S_n} u_\sigma \circ u_\rho \\ &= \frac{1}{n!^2}\sum_{\sigma\in S_n}\sum_{\rho\in S_n} u_{\sigma\circ\rho}. \end{aligned}$$

Or, $\sigma \mapsto \sigma\circ\rho$ est une bijection de S_n dans S_n d'où

$$\begin{aligned} p^2 &= \frac{1}{n!^2}\sum_{\sigma\in S_n}\sum_{\tau\in S_n} u_\tau \\ &= \frac{1}{n!}\sum_{\tau\in S_n} u_\tau = p. \end{aligned}$$

2. Remarquons que, pour tout k et tout i, il existe $(n-1)!$ permutations envoyant k sur i d'où

$$p(e_k) = \frac{1}{n!}\sum_{\sigma\in S_n} e_{\sigma(k)} = \frac{(n-1)!}{n!}\sum_{i=1}^{n} e_i.$$

Ainsi, $\operatorname{Im} p = \operatorname{Vect}(p(e_1),\dots,p(e_n)) = \operatorname{Vect}\left(\sum_{i=1}^n e_i\right)$.

3. Soit $x \in \mathbb{R}^n$ de coordonnées x_1, x_2,..., x_n. Alors,

$$p(x) = \sum_{i=1}^{n} x_i\, p(e_i) = \left(\sum_{i=1}^{n} x_i\right)\frac{1}{n}\sum_{j=1}^{n} e_j.$$

Ainsi, $x \in \operatorname{Ker} p$ si, et seulement si, $\sum_{i=1}^n x_i = 0$ (ce qui est bien l'équation d'un hyperplan).

Corrigés des exercices d'approfondissement du chapitre 18

Exercice A

1. ▹ Commençons par un simple calcul. Soit a, b, c et $d \in [\![1,n]\!]$ deux à deux distincts. Alors,

$$(a\ b\ c\ d)\circ(a\ b\ d\ c) = (a\ c\ b).$$

Ainsi, le groupe engendré par les 4-cycles contient tous les 3-cycles donc le groupe engendré par ceux-ci, à savoir A_n.

▹ Par ailleurs, le groupe engendré par les 4-cycles contient strictement A_n (puisqu'il contient également les 4-cycles qui sont des permutations impaires). Son ordre est donc un diviseur de $n!$ (théorème de Lagrange), strictement plus grand que $\frac{n!}{2}$: c'est $n!$ et le sous-groupe recherché est donc $n!$.

2. Soit a_1, a_2, ..., $a_p \in [\![1,n]\!]$ deux à deux distincts. La composition des deux p-cycles proposée dans l'indication donne le 3-cycle $(a_1\ a_2\ a_3)$. Ainsi, le groupe recherché contient tous les 3-cycles donc le sous-groupe qu'ils engendrent : A_n.

• Si p est impair, tous les p-cycles sont de signature 1 donc appartiennent à A_n. Le sous-groupe recherché est donc inclus dans A_n et avec la remarque liminaire, égal à A_n.

• Si p est pair, tous les p-cycles sont de signature -1 donc le sous-groupe recherché est strictement plus grand (au sens de l'inclusion) que A_n. En utilisant des considérations de cardinal comme dans la première question, on obtient qu'il est égal à S_n.

Exercice B

Considérons τ la transposition $(n+1\ \ n+2)$ et définissons l'application $\varphi : S_n \to S_{n+2}$ définie par

$$\varphi : \sigma \mapsto \begin{cases} \sigma & \text{si } \sigma \in A_n, \\ \sigma \circ \tau & \text{sinon.} \end{cases}$$

Il y a délibérément un léger abus de notation puisque dans l'ensemble d'arrivée, on note σ la permutation $\sigma' \in S_{n+2}$ qui laisse fixes $n+1$ et $n+2$ et telle que, pour tout $k \leq n$, $\sigma'(k) = \sigma(k)$.

▹ Remarquons tout d'abord que φ est à valeurs dans A_{n+2} puisque, pour tout $\sigma \in A_n$, $\varepsilon(\varphi(\sigma)) = \varepsilon(\sigma) = 1$ et, pour tout $\sigma \notin A_n$, $\varepsilon(\varphi(\sigma)) = \varepsilon(\sigma)\varepsilon(\tau) = (-1)^2 = 1$.

▹ Montrons que φ est un morphisme de groupes. Soit σ, $\rho \in S_n$. Discutons selon l'appartenance de σ et ρ à A_n et résumons les calculs de $\varphi(\sigma) \circ \varphi(\rho)$ dans le tableau suivant

	$\rho \in A_n$	$\rho \notin A_n$
$\sigma \in A_n$	$\sigma \circ \rho$	$\sigma \circ \rho \circ \tau$
$\sigma \notin A_n$	$\sigma \circ \tau \circ \rho \circ \tau = \sigma \circ \rho$	$\sigma \circ \tau \circ \rho = \sigma \circ \rho \circ \tau$

en utilisant que τ commute avec σ et ρ puisque son support ne rencontre pas $[\![1, n]\!]$.

On remarque en particulier que $\varphi(\sigma) \circ \varphi(\rho) = \varphi(\sigma \circ \rho)$. Ainsi, φ est bien un morphisme de groupes.

▹ Déterminons le noyau de φ. Soit $\sigma \in \operatorname{Ker} \varphi$.

• Si $\sigma \in A_n$, $\mathrm{Id} = \varphi(\sigma) = \sigma$.
• Si $\sigma \notin A_n$, $\varphi(\sigma)(n+1) = (n+2)$ donc $\sigma \notin \operatorname{Ker} \varphi$.

En conclusion, le noyau de φ est réduit à la permutation identité donc le morphisme φ est injectif.

Exercice C

Considérons une permutation σ telle que la quantité Σ_σ soit maximale et montrons par l'absurde que $\sigma = \mathrm{Id}$: pour cela introduisons j le plus petit entier tel que $\sigma(j) \neq j$. Par construction, $\sigma(j) > j$ et $\sigma^{-1}(j) > j$ (car en-deça de j tous les entiers sont fixes). Enfin définissons la permutation $\sigma' = \sigma \circ (\sigma^{-1}(j)\ j)$. Alors,

$$\begin{aligned} \Sigma_\sigma - \Sigma_{\sigma'} &= \sum_{k=1}^{n} x_k\, y_{\sigma(k)} - \sum_{k=1}^{n} x_k\, y_{\sigma'(k)} \\ &= x_j\, y_{\sigma(j)} + x_{\sigma^{-1}(j)} y_j - x_j\, y_j - x_{\sigma^{-1}(j)} y_{\sigma(j)} \\ &= (x_j - x_{\sigma^{-1}(j)})(y_{\sigma(j)} - y_j). \end{aligned}$$

Comme $\sigma(j) > j$ et $\sigma^{-1}(j) > j$, on a par monotonie des suites finies $x_j > x_{\sigma^{-1}(j)}$ et $y_{\sigma(j)} < y_j$ donc $\Sigma_\sigma < \Sigma_{\sigma'}$: contradiction avec l'hypothèse de maximalité. En conclusion, l'unique permutation réalisant le maximum de Σ_σ est $\sigma = \mathrm{Id}$.

Ce résultat est connu sous le nom d'« inégalité de réarrangement ».

Corrigés des problèmes du chapitre 18

Problème 1

1. On obtient :

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 \\ 2 & 5 & 3 & 8 & 9 & 6 & 7 & 4 & 15 & 10 & 11 & 12 & 16 & 14 & 1 & 13 \end{pmatrix}$$

2. Remarquons tout d'abord que la permutation associée à la configuration finale est paire, celle associée à la configuration initiale est impaire (c'est une transposition).
Or, chacun des mouvements autorisés permet de passer d'une permutation-configuration à une autre en composant par une transposition de la forme $(16\ k)$. Pour passer d'une permutation impaire à une permutation paire avec des transpositions, il convient de composer par un nombre impair de transpositions : en conclusion, toute succession de mouvements autorisés qui permet de passer de la configuration initiale à la configuration finale comporte un nombre impair de mouvements.
3. L'entier défini comme la somme des indices d'abscisse et d'ordonnée de la case 16 est égal modulo 2 au nombre de mouvements effectués (c'est vrai dans la position initiale puis chaque mouvement change d'une unité l'abscisse ou l'ordonnée). Dans la position finale, la case 16 a même abscisse et ordonnée donc la somme de ces deux coordonnées est paire : on ne peut donc y revenir avec un nombre impair de mouvements autorisés.
4. Il n'y a pas de solution à ce jeu puisque les deux résultats précédents sont incompatibles.

Problème 2

1. Supposons que $H \cap A_5 = A_5$ et $H \neq A_5$. Il existe une permutation $\sigma \in H \setminus A_5$. L'application $\rho \mapsto \sigma \circ \rho$ est une bijection de $H \cap A_5 = A_5$ vers $H \setminus A_5$. Ainsi H est de cardinal $2 \operatorname{card} A_5 = \operatorname{card} S_5$ donc $H = S_5$.
2. Le noyau de la restriction de la signature à H est inclus dans $H \cap A_5 = \{\mathrm{Id}\}$ donc cette application est injective et $\operatorname{card} H \leq \operatorname{card}\{\pm 1\} = 2$. Supposons que H contienne une permutation σ distincte de l'identité. Alors pour tout $\rho \in S_5$, $\rho \circ \sigma \circ \rho^{-1} = \sigma$ (car c'est une permutation impaire et elle est dans le sous-groupe distingué H) donc $\rho \circ \sigma = \sigma \circ \rho$: la permutation σ est dans le centre de S_5 donc égale à l'identité, contradiction. En conclusion, $H = \{\mathrm{Id}\}$.
3. a. Résumons le dénombrement des 60 éléments de A_5 dans le tableau suivant

ordre	nombre	description
1	1	Identité
2	$\frac{1}{2}\binom{5}{4}\binom{4}{2} = 15$	produits de deux transpositions à supports disjoints
3	$2.\binom{5}{3} = 20$	3-cycles
4	0	
5	$4! = 24$	5-cycles

b. Deux éléments admettant la même décomposition en cycles disjoints sont conjugués dans S_5 (il suffit de conjuguer par l'une des permutations qui associe aux éléments de la première décomposition un élément du cycle « correspondant » de la seconde). Ainsi, si H contient un élément d'ordre 2 (respectivement 3 ou 5) de A_5, alors il contient tous ses conjugués donc tous les éléments d'ordre 2 (respectivement 3 ou 5).

c. Remarquons que si $H \cap A_5 \neq A_5$, alors $\operatorname{card} H \cap A_5 \leq 30$ car c'est un diviseur strict de $\operatorname{card} A_5 = 60$ d'après le théorème de Lagrange. Ainsi, $H \cap A_5$ ne peut contenir des éléments d'ordre 2 et 3 ou 2 et 5 ou 3 et 5. Si $H \cap A_5$ ne contient que des éléments d'ordre 2 (respectivement 3 ou 5) et l'identité, alors son cardinal est 16 (respectivement 21 ou 25) qui ne divise par 60 : contradiction!
Ainsi, $H \cap A_5$ est égal soit à A_5 soit à $\{\mathrm{Id}\}$.
4. D'après les questions précédentes, les sous-groupes distingués de S_5 sont parmi $\{\mathrm{Id}\}$, A_5 et S_5, or ces trois groupes sont distingués ; vérifions cette propriété pour A_5. Pour tout $\rho \in A_5$ et $\sigma \in S_5$,

$$\varepsilon(\sigma \circ \rho \circ \sigma^{-1}) = \varepsilon(\sigma)\varepsilon(\rho)\varepsilon(\sigma^{-1}) = \varepsilon(\rho) = 1$$

donc $\sigma \circ \rho \circ \sigma^{-1} \in A_5$.
5. a. Montrons que les 3-cycles sont conjugués dans A_5 (on procède de même pour les produits de transpositions de supports disjoints). Considérons deux trois cycles $(a\ b\ c)$ et $(a\ d\ e)$ (leurs supports se rencontrent forcément en au moins un point car il n'y a que 5 éléments). Ils sont alors conjugués avec $(b\ d)\circ(c\ e)$ si $b \neq d$ et $c \neq e$, avec $(b\ e)\circ(c\ d)$ si $b = d$ ou $c = e$ et ces permutations sont paires.
Comme tous les éléments d'ordre 2 (respectivement 3) de A_5 sont conjugués, si un sous-groupe distingué en contient un, alors il les contient tous.
b. Soit σ et ρ deux éléments d'ordre 5 de A_5, c'est-à-dire deux 5-cycles. Alors σ est un conjugué de ρ dans S_5. Il existe donc $\tau \in S_5$ tel que $\sigma = \tau \circ \rho \circ \tau^{-1}$.
- Si τ est une permutation paire, le résultat est établi.
- Sinon, on remarque que ρ est le conjugué de ρ^2 par une permutation impaire puisque, pour tout $\{a, b, c, d, e\} = [\![1, 5]\!]$,

$$(a\ b\ c\ d\ e) = (b\ e\ d\ c)\circ(a\ c\ e\ b\ d)\circ(b\ c\ d\ e).$$

Ainsi, si τ est une permutation impaire, alors σ est un conjugué de ρ^2 dans A_5 en utilisant les deux relations et le fait que la composée de deux applications impaires est paire.
Si K contient un élément ρ d'ordre 5 de A_5, alors il contient aussi ρ^2 (car c'est un groupe) et leurs conjugués (car il est distingué) donc contient tous les éléments d'ordre 5 de A_5.
c. On conclut comme précédemment : les sous-groupes distingués de A_5 sont A_5 et $\{\mathrm{Id}\}$.

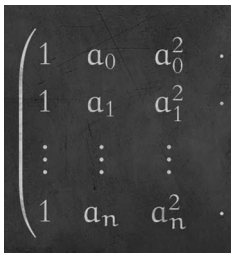
Chapitre 19

Déterminants

1. Déterminant de n vecteurs dans une base — **2.** Applications linéaires et matrices — **3.** Quelques méthodes de calcul — **4.** Applications

Objectifs et compétences du programme

- Calculer le déterminant d'une famille de vecteurs.
- Calculer le déterminant d'une matrice, d'un endomorphisme.
- Utiliser un calcul de déterminant pour caractériser une base ou une propriété d'inversibilité.
- Préciser le rang d'une matrice par les déterminants extraits.



Alexandre-Théophile Vandermonde, mathématicien français du siècle des lumières(1735-1796)
Il s'intéresse tardivement aux mathématiques et en particulier aux systèmes linéaires sous l'appellation « théorie de l'élimination » (sous-entendu, élimination des variables) ; c'est à ce titre qu'il va considérer ce que l'on appelle aujourd'hui les matrice et déterminant de Vandermonde (même si le déterminant n'est pas calculé explicitement dans ces travaux). On a également attribué son nom à une identité sur les coefficients binomiaux.

COURS

Dans ce chapitre, nous introduisons un nouvel outil : le déterminant. Cette forme n-linéaire sert à caractériser les familles de vecteurs qui sont des bases, les matrices inversibles et les automorphismes. Tout l'enjeu, une fois la construction achevée et ses caractérisations établies, consiste à savoir calculer les déterminants rencontrés : dans la troisième section, on développe alors différents algorithmes de calcul et quelques astuces pratiques.

1. Déterminant de n vecteurs dans une base

1.1. Formes n-linéaires

Définition 19.1.

Soit E et F deux $\mathbb{K}$-espaces vectoriels.
▹ Une application n-linéaire de E dans F est une application $f : E^n \to F$ telle que pour tout $k \in [\![1, n]\!]$ et tout $(x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_n) \in E^{n-1}$, l'application

$$x \mapsto f(x_1, \dots, x_{k-1}, x, x_{k+1}, \dots, x_n)$$

est linéaire de E dans F.
▹ Une forme n-linéaire sur E est une application n-linéaire de E dans son corps de base.

Proposition 19.2.

L'ensemble des applications n-linéaires entre deux $\mathbb{K}$-espaces vectoriels est un espace vectoriel.

Définition 19.3.

Soit f une application n-linéaire entre deux $\mathbb{K}$-espaces vectoriels E et F.
▹ f est symétrique si, pour toute transposition $\tau \in S_n$,

$$\forall (x_1, \dots, x_n) \in E^n, \qquad f(x_{\tau(1)}, \dots, x_{\tau(n)}) = f(x_1, \dots, x_n).$$

▹ f est antisymétrique si, pour toute transposition $\tau \in S_n$,

$$\forall (x_1, \dots, x_n) \in E^n, \qquad f(x_{\tau(1)}, \dots, x_{\tau(n)}) = -f(x_1, \dots, x_n).$$

▹ f est alternée si $f(x_1, \dots, x_n) = 0$ dès que deux des vecteurs $x_1, \dots, x_n$ sont égaux.

En utilisant le fait que toute permutation est un produit de transpositions, on obtient la proposition suivante.

Proposition 19.4.

Soit f une application n-linéaire entre deux $\mathbb{K}$-espaces vectoriels E et F.
▹ f est symétrique si, et seulement si, pour tout $\sigma \in S_n$,

$$\forall (x_1, \dots, x_n) \in E^n, \qquad f(x_{\sigma(1)}, \dots, x_{\sigma(n)}) = f(x_1, \dots, x_n).$$

▹ f est antisymétrique si, et seulement si, pour tout $\sigma \in S_n$,

$$\forall (x_1, \dots, x_n) \in E^n, \qquad f(x_{\sigma(1)}, \dots, x_{\sigma(n)}) = \varepsilon(\sigma) f(x_1, \dots, x_n),$$

où $\varepsilon(\sigma)$ est la signature de la permutation σ.

On peut également dans les cas standards montrer l'équivalence entre la propriété d'antisymétrie et le caractère alterné.

Proposition 19.5.

Soit f une application n-linéaire entre deux $\mathbb{K}$-espaces vectoriels E et F. Alors, f est alternée si, et seulement si, f est antisymétrique.

Démonstration

Écrivons la preuve pour une application 2-linéaire f (le passage à n arguments ne présente aucune difficulté autre que d'écriture et de lecture).

($\Rightarrow$) Si f est alternée, alors pour tout $x, y \in E$, on a $f(x+y, x+y) = 0$. D'où $f(x,y) + f(y,x) = 0$ et f est antisymétrique.

($\Leftarrow$) Si f est antisymétrique, alors, pour tout $x \in E$, $f(x,x) = -f(x,x)$ d'où $2f(x,x) = 0$ et le caractère alterné suit. ■

Résumons dans la proposition suivante le comportement d'une forme n-linéaire alternée lors d'une manipulation élémentaire des vecteurs en argument (on se limite à l'énoncé sur les premiers vecteurs pour des raisons de simplicité des notations).

Proposition 19.6.

Soit f n-linéaire alternée entre deux $\mathbb{K}$-espaces vectoriels E et F.
▹ Pour tout $(x_1, \dots, x_n) \in E^n$,

$$f(x_2, x_1, \dots, x_n) = -f(x_1, x_2, \dots, x_n).$$

▹ Pour tout $(x_1, \dots, x_n) \in E^n$ et tout $\lambda \in \mathbb{K}$,

$$f(\lambda.x_1, x_2, \dots, x_n) = \lambda f(x_1, x_2, \dots, x_n).$$

▹ Pour tout $(x_1, \dots, x_n) \in E^n$, tout $\lambda \in \mathbb{K}$ et tout $k \in [\![2, n]\!]$,

$$f(x_1 + \lambda x_k, \dots, x_n) = f(x_1, \dots, x_n).$$

Démonstration

Il suffit d'utiliser la linéarité, le caractère alterné ou antisymétrique. ■

1.2. Déterminant dans une base donnée

Proposition 19.7.

L'espace vectoriel des formes n-linéaires alternées sur un espace vectoriel de dimension n est de dimension 1.
Plus précisément, si E est un espace vectoriel de dimension n dont une base est $\mathscr{B} = (e_1, \dots, e_n)$, alors, pour toute forme n-linéaire alternée f, il existe un scalaire $\lambda \in \mathbb{K}$ tel que

$$\forall (x_1, \dots, x_n) \in E^n, \qquad f(x_1, \dots, x_n) = \lambda \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{j=1}^{n} x_{\sigma(j),j},$$

où, pour tout $j \in [\![1, n]\!]$, $x_j = \sum_{i=1}^{n} x_{i,j} e_i$ (c'est-à-dire les $(x_{i,j})_i$ sont les coordonnées de x_j dans la base $\mathscr{B}$).

Démonstration

Montrons tout d'abord qu'une forme n-linéaire alternée f est de la forme recherchée. Pour tout $(x_1, \dots, x_n) \in E^n$, les linéarités entraînent que

$$f(x_1, \dots, x_n) = \sum_{i_1=1}^{n} \dots \sum_{i_n=1}^{n} \prod_{j=1}^{n} x_{i_j,j} f(e_{i_1}, \dots, e_{i_n}).$$

Si deux des indices $i_1, \dots, i_n$ sont égaux, alors $f(e_{i_1}, \dots, e_{i_n}) = 0$ d'après le caractère alterné. Ainsi, chaque terme non nul $f(e_{i_1}, \dots, e_{i_n})$ correspond à une application $j \mapsto i_j$ bijective, c'est-à-dire une permutation de S_n. Ainsi

$$f(x_1, \dots, x_n) = \sum_{\sigma \in S_n} \prod_{j=1}^{n} x_{\sigma(j),j} f(e_{\sigma(1)}, \dots, e_{\sigma(n)}).$$

En utilisant l'antisymétrie, on obtient la forme désirée avec $\lambda = f(e_1, \dots, e_n)$.
Montrons désormais que l'application $f : (x_1, \dots, x_n) \mapsto \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{j=1}^{n} x_{\sigma(j),j}$ est alternée (la propriété de n-linéarité est évidente). Pour tout $(x_1, \dots, x_n) \in E^n$ et $\tau \in S_n$,

$$\begin{aligned}
f(x_{\tau(1)}, \dots, x_{\tau(n)}) &= \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{j=1}^{n} x_{\sigma(j),\tau(j)} \\
&= \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{j=1}^{n} x_{\sigma \circ \tau^{-1}(\tau(j)),\tau(j)} \\
&= \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{k=1}^{n} x_{\sigma \circ \tau^{-1}(k),k} \\
&= \sum_{\sigma' \in S_n} \varepsilon(\sigma' \circ \tau) \prod_{k=1}^{n} x_{\sigma'(k),k},
\end{aligned}$$

en posant $\sigma' = \sigma \circ \tau$ et en utilisant que $\sigma \mapsto \sigma \circ \tau^{-1}$ est une bijection de S_n dans S_n. Le résultat suit en utilisant la propriété de multiplicativité de la signature.

$$f(x_{\tau(1)}, \dots, x_{\tau(n)}) = \varepsilon(\tau) f(x_1, \dots, x_n).$$

■

Définition 19.8.

Soit $\mathscr{B}=(e_1,\dots,e_n)$ une base de E. Le déterminant en base $\mathscr{B}$ est l'unique forme n-linéaire alternée f telle que $f(e_1,\dots,e_n)=1$. On le note $\det_{\mathscr{B}}$.
D'après la proposition précédente,

$$\forall (x_1,\dots,x_n)\in E^n,\qquad \det_{\mathscr{B}}(x_1,\dots,x_n)=\sum_{\sigma\in S_n}\varepsilon(\sigma)\prod_{j=1}^{n}x_{\sigma(j),j},$$

où, pour tout $j\in[\![1,n]\!]$, $(x_{i,j})_{i\in[1,n]}$ désigne les coordonnées de x_j dans la base $\mathscr{B}$.

Soit $\mathscr{B}$ une base de E, $(x_1,\dots,x_n)\in E^n$ et, pour tout $j\in[\![1,n]\!]$, $(x_{i,j})_i$ les coordonnées de x_j dans la base $\mathscr{B}$. Alors, on note

$$\det_{\mathscr{B}}(x_1,\dots,x_n)=\begin{vmatrix} x_{1,1} & x_{1,2} & \dots & x_{1,n}\\ x_{2,1} & x_{2,2} & \dots & x_{2,n}\\ \vdots & \vdots & & \vdots\\ x_{n,1} & x_{n,2} & \dots & x_{n,n}\end{vmatrix}.$$

Remarque

Utilisons cette formule explicite dans les cas $n=2$ et $n=3$.
▹ Le groupe symétrique S_2 contient deux éléments : l'identité de signature 1 et la transposition de signature -1. Par conséquent,

$$\begin{vmatrix} x_{1,1} & x_{1,2}\\ x_{2,1} & x_{2,2}\end{vmatrix}=1.x_{1,1}x_{2,2}+(-1)x_{1,2}x_{2,1}.$$

▹ Le groupe symétrique S_3 contient six éléments : 3 transpositions de signature -1, l'identité et 2 permutations circulaires de signature 1. Par conséquent,

$$\begin{vmatrix} x_{1,1} & x_{1,2} & x_{1,3}\\ x_{2,1} & x_{2,2} & x_{2,3}\\ x_{3,1} & x_{3,2} & x_{3,3}\end{vmatrix}=1.x_{1,1}x_{2,2}x_{3,3}+1.x_{2,1}x_{3,2}x_{1,3}+1.x_{3,1}x_{1,2}x_{2,3}$$
$$+(-1).x_{1,1}x_{2,3}x_{3,2}+(-1).x_{2,1}x_{3,3}x_{1,2}+(-1).x_{3,1}x_{1,3}x_{2,2}.$$

Ce résultat s'appelle règle de Sarrus. Comme chacun des termes correspond à une permutation, on peut se souvenir de cette formule avec la représentation suivante :

$\varepsilon(\sigma)=+1$

$$\begin{vmatrix} x_{1,1} & x_{1,2} & x_{1,3}\\ x_{2,1} & x_{2,2} & x_{2,3}\\ x_{3,1} & x_{3,2} & x_{3,3}\end{vmatrix}\begin{matrix} x_{1,1} & x_{1,2} & x_{1,3}\\ x_{2,1} & x_{2,2} & x_{2,3}\end{matrix}$$

$\varepsilon(\sigma)=-1$

$$\begin{vmatrix} x_{1,1} & x_{1,2} & x_{1,3}\\ x_{2,1} & x_{2,2} & x_{2,3}\\ x_{3,1} & x_{3,2} & x_{3,3}\end{vmatrix}\begin{matrix} x_{1,1} & x_{1,2} & x_{1,3}\\ x_{2,1} & x_{2,2} & x_{2,3}\end{matrix}$$

Exemple

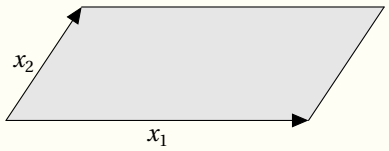
Avec les notations précédentes, si, pour tous $i>j$, $x_{i,j}=0$, alors

$$\det_{\mathscr{B}}(x_1,\dots,x_n)=\prod_{i=1}^{n}x_{i,i}.$$

Autrement dit, si le déterminant est de forme « triangulaire », il est égal au produit des éléments diagonaux.
Le résultat est également vrai si pour tout $i < j$, $x_{i,j} = 0$.

Remarque

Dans $\mathbb{R}^2$, le déterminant de deux vecteurs dans la base canonique s'interprète comme l'aire (au signe près) du parallélogramme construit sur ces deux vecteurs.



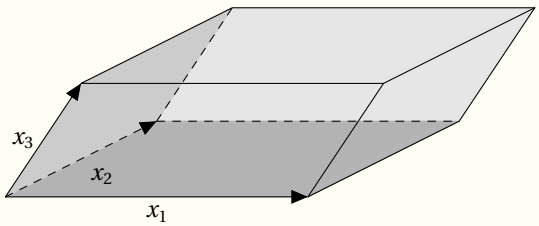
Commençons par remarquer qu'on peut retrancher à x_2 un multiple de x_1 (la projection orthogonale de x_2 sur la droite dirigée par x_1) de sorte à ce que le nouveau vecteur remplaçant x_2 soit orthogonal à x_1, ce que l'on suppose dorénavant. Alors, en notant (a, b) les coordonnées de x_1 et (c, d) les coordonnées du nouveau vecteur x_2 dans la base canonique, l'aire du parallélogramme (du rectangle en fait) est

$$\sqrt{a^2+b^2}\sqrt{c^2+d^2} = \sqrt{(ad-bc)^2+(ac+bd)^2} = |ad-bc|,$$

en utilisant l'orthogonalité à la dernière égalité. On retrouve bien la valeur absolue du déterminant

$$\begin{vmatrix} a & c \\ b & d \end{vmatrix} = ad - bc.$$

De même dans $\mathbb{R}^3$, le déterminant de trois vecteurs dans la base canonique s'interprète comme le volume (au signe près) du parallélépipède construit sur ces vecteurs.



1.3. Caractérisation des bases

Comme le déterminant dans une base donnée est n-linéaire et alterné, on connaît les modifications entraînées par les opérations élémentaires. Cela entraîne le résultat suivant de caractérisation des bases.

Proposition 19.9.

Soit $\mathscr{B}$ une base d'un $\mathbb{K}$-espace vectoriel E de dimension n. Alors, n vecteurs $x_1, \dots, x_n \in E$ forme une base de E si, et seulement si,

$$\det_{\mathscr{B}}(x_1, \dots, x_n) \neq 0.$$

Démonstration
Comme la famille a pour cardinal $n = \dim E$, il suffit de montrer que $(x_1, \dots, x_n) \in E^n$ est une famille libre si, et seulement si, $\det_{\mathscr{B}}(x_1, \dots, x_n) \neq 0$.
($\Leftarrow$) Si la famille est liée, alors l'un des vecteurs s'exprime comme combinaison linéaire des autres. En utilisant la linéarité en ce vecteur et le caractère alterné, on en déduit que le déterminant de cette famille est nul.
($\Rightarrow$) Si la famille $\mathscr{X} = (x_1, \dots, x_n)$ est une base alors les deux formes n-linéaires alternées $\det_{\mathscr{X}}$ et $\det_{\mathscr{B}}$ sont proportionnelles : il existe $\lambda \in \mathbb{K}$ tel que $\det_{\mathscr{X}} = \lambda \det_{\mathscr{B}}$; en évaluant en $(x_1, \dots, x_n)$ cette égalité, on obtient que $\lambda \det_{\mathscr{B}}(x_1, \dots, x_n) = 1 \neq 0$. ■

Exemple

Considérons $n > 2$ réels $(a_i)_{i \in [1,n]}$ et déterminons le déterminant dans la base canonique de $\mathbb{R}^n$ des vecteurs x_j pour $j \in [\![1, n]\!]$ dont les coordonnées sont $\cos(a_i + a_j)$ pour tout $i \in [\![1, n]\!]$.
Par exemple, pour $n = 3$, ce déterminant est

$$\begin{vmatrix} \cos(a_1+a_1) & \cos(a_1+a_2) & \cos(a_1+a_3) \\ \cos(a_2+a_1) & \cos(a_2+a_2) & \cos(a_2+a_3) \\ \cos(a_3+a_1) & \cos(a_3+a_2) & \cos(a_3+a_3) \end{vmatrix}.$$

En notant u le vecteur de coordonnées $\cos(a_i)$ et v le vecteur de coordonnées $\sin(a_i)$, on obtient que, pour tout $j \in [\![1, n]\!]$, $x_j = \cos(a_j)u - \sin(a_j)v$ donc que la famille $(x_1, \dots, x_n)$ est de rang au plus 2. Par conséquent, ce n'est pas une base de $\mathbb{R}^n$ et son déterminant est nul.

Comme le déterminant d'une base est non nul, on peut se poser la question de son signe.

Définition 19.10.

Un $\mathbb{R}$-espace vectoriel orienté de dimension finie est un $\mathbb{R}$-espace vectoriel E de dimension finie muni d'une base de référence $\mathscr{B}$.
Une base $\mathscr{B}' = (e'_1, \dots, e'_n)$ de E est directe si $\det_{\mathscr{B}}(e'_1, \dots, e'_n) > 0$; elle est indirecte sinon.

Ce concept devient important lorsque l'on dispose de la notion de base orthonormée.

Exemple

Considérons l'espace $\mathbb{R}^3$ orienté par la base canonique (e_1, e_2, e_3). Alors, les bases (e_2, e_3, e_1) et (e_3, e_1, e_2) sont directes, alors que la base (e_1, e_3, e_2) est indirecte.

2. Applications linéaires et matrices

2.1. Déterminant d'un endomorphisme

Proposition 19.11.

Soit $u \in \mathscr{L}(E)$ et $\mathscr{B} = (e_1, \dots, e_n)$ une base de E. Le scalaire $\det_{\mathscr{B}}(u(e_1), \dots, u(e_n))$ ne dépend pas de la base $\mathscr{B}$.

On en déduit la définition suivante.

Définition 19.12.

Le déterminant d'un endomorphisme $u \in \mathscr{L}(E)$ est le scalaire $\det_{\mathscr{B}}(u(e_1),\dots,u(e_n))$ où $\mathscr{B} = (e_1,\dots,e_n)$ est une base de E.

Remarque

L'application $(x_1,\dots,x_n) \mapsto \det_{\mathscr{B}}(u(x_1),\dots,u(x_n))$ est n-linéaire, alternée donc proportionnelle à $\det_{\mathscr{B}}$. Ainsi, il existe $\lambda \in \mathbb{K}$ tel que

$$\forall (x_1,\dots,x_n) \in E^n, \qquad \det_{\mathscr{B}}(u(x_1),\dots,u(x_n)) = \lambda.\det_{\mathscr{B}}(x_1,\dots,x_n).$$

En évaluant en $(e_1,\dots,e_n)$, on obtient que $\lambda = \det u$ donc

$$\forall (x_1,\dots,x_n) \in E^n, \qquad \det_{\mathscr{B}}(u(x_1),\dots,u(x_n)) = \det(u).\det_{\mathscr{B}}(x_1,\dots,x_n).$$

Démonstration

Considérons $\mathscr{B} = (e_1,\dots,e_n)$ et $\mathscr{B}' = (e'_1,\dots,e'_n)$ deux bases de E. Remarquons que les applications $\det_{\mathscr{B}'}$ et $(x_1,\dots,x_n) \mapsto \det_{\mathscr{B}}(u(x_1),\dots,u(x_n))$ sont n-linéaires, alternées donc proportionnelles à $\det_{\mathscr{B}}$. Il existe donc deux scalaires λ, μ tels que, pour tout $(x_1,\dots,x_n) \in E^n$,

$$\det_{\mathscr{B}}(u(x_1),\dots,u(x_n)) = \lambda \det_{\mathscr{B}}(x_1,\dots,x_n),$$
$$\det_{\mathscr{B}'}(x_1,\dots,x_n) = \mu \det_{\mathscr{B}}(x_1,\dots,x_n).$$

En évaluant la première relation en $(e_1,\dots,e_n)$, on obtient que

$$\lambda = \det_{\mathscr{B}}(u(e_1),\dots,u(e_n)).$$

En multipliant la première relation par μ, on obtient, en utilisant la seconde relation,

$$\det_{\mathscr{B}'}(u(x_1),\dots,u(x_n)) = \lambda \det_{\mathscr{B}'}(x_1,\dots,x_n),$$

et en évaluant en $(e'_1,\dots,e'_n)$, on obtient

$$\lambda = \det_{\mathscr{B}'}(u(e'_1),\dots,u(e'_n)).$$

Ce qui prouve le résultat annoncé. ■

Exemple

Soit $\lambda \in \mathbb{K}$ et $u = \lambda \mathrm{Id}$; alors, $\det(u) = \det_{\mathscr{B}}(\lambda e_1,\dots,\lambda e_n) = \lambda^n \det_{\mathscr{B}}(e_1,\dots,e_n) = \lambda^n$.

Exemple

Calculons le déterminant de l'endomorphisme u de $\mathscr{M}_n(\mathbb{R})$ défini par $u : M \mapsto M^{\mathrm{T}}$.
D'après la définition, il suffit de calculer le déterminant de l'une des matrices de u (donc de taille $n^2 \times n^2$).
▹ Une première idée consiste à considérer la base canonique de $\mathscr{M}_n(\mathbb{R})$:

$$(E_{1,1}, E_{1,2},\dots,E_{1,n}, E_{2,1}, E_{2,2},\dots,E_{2,n},\dots,E_{n,1}, E_{n,2},\dots,E_{n,n}).$$

Toutefois la matrice de u dans cette base n'est pas très « facile » à manipuler.
▹ Une meilleure idée consiste à obtenir une base de $\mathscr{M}_n(\mathbb{R})$ en mettant les $\frac{n(n+1)}{2}$ vecteurs d'une base de $S_n(\mathbb{R})$ puis les $\frac{n(n-1)}{2}$ vecteurs d'une base de $A_n(\mathbb{R})$ (ces deux sous-espaces sont bien supplémentaires). Dans cette base la matrice de u est diagonale et son déterminant se calcule faci-

lement :

$$\det u = (-1)^{\frac{n(n+1)}{2}} 1^{\frac{n(n-1)}{2}} = (-1)^{\frac{n(n+1)}{2}}.$$

Proposition 19.13.

Soit $u,\ v \in \mathscr{L}(E)$. Alors, $\det(u \circ v) = \det u . \det v$.

Démonstration

Pour tout $(x_1, \dots, x_n) \in E^n$,

$$\begin{aligned}\det_{\mathscr{B}}(u \circ v(x_1), \dots, u \circ v(x_n)) &= \det u . \det_{\mathscr{B}}(v(x_1), \dots, v(x_n)) \\ &= \det u . \det v . \det_{\mathscr{B}}(x_1, \dots, x_n).\end{aligned}$$

D'où $\det(u \circ v) = \det u . \det v$. ■

Proposition 19.14.

Soit $u \in \mathscr{L}(E)$. L'endomorphisme u est inversible si, et seulement si, $\det(u) \neq 0$. Dans ce cas,

$$\det(u^{-1}) = \det(u)^{-1}.$$

Démonstration

Par définition, $\det(u) \neq 0$ si, et seulement si, la famille $(u(e_1), \dots, u(e_n))$ est une base donc si, et seulement si, u est un automorphisme. La formule pour $\det(u^{-1})$ provient de la proposition précédente puisque $\det(u^{-1})\det(u) = \det(u^{-1} \circ u) = \det(\mathrm{Id}) = 1$. ■

2.2. Déterminant d'une matrice

Définition 19.15.

Soit $M = (m_{i,j})_{i,j} \in \mathscr{M}_n(\mathbb{K})$. Le déterminant d'une matrice M est le scalaire noté $\det M$ défini comme le déterminant des vecteurs colonnes de M dans la base canonique de $\mathscr{M}_{n,1}(\mathbb{K})$, c'est-à-dire

$$\det M = \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{j=1}^{n} m_{\sigma(j),j}.$$

Exemple

Montrons que le déterminant d'une matrice triangulaire supérieure $M = (m_{i,j})_{i,j}$ est le produit des éléments diagonaux.

Pour cela, montrons que, pour toute permutation $\sigma \in S_n$ distincte de l'identité, il existe $j \in [\![1, n]\!]$ tel que $\sigma(j) > j$. En effet, la somme des entiers de 1 à n est (car σ est une permutation de $[\![1, n]\!]$)

$$\sum_{j=1}^{n} \sigma(j).$$

De plus, si $\sigma(j) \leq j$ pour tout $j \in [\![1, n]\!]$, alors $\sum_{j=1}^{n} \sigma(j) < \sum_{j=1}^{n} j$, (l'inégalité est stricte car $\sigma \neq \mathrm{Id}$) d'où la contradiction. Ainsi, le déterminant de M est

$$\det M = \sum_{\sigma\in S_n} \varepsilon(\sigma)\prod_{j=1}^{n} m_{\sigma(j),j} = \prod_{j=1}^{n} m_{j,j} + \sum_{\substack{\sigma\in S_n\\ \sigma\neq \mathrm{Id}}} \varepsilon(\sigma)\underbrace{\prod_{j=1}^{n} m_{\sigma(j),j}}_{=0} = \prod_{j=1}^{n} m_{j,j}.$$

Établissons le lien entre déterminant d'un endomorphisme et d'une matrice.

Proposition 19.16.

Soit $u \in \mathscr{L}(E)$ et $\mathscr{B}$ une base de E. Alors, $\det u = \det(\mathrm{mat}_{\mathscr{B}}(u))$.

Démonstration

Notons $\mathscr{B} = (e_1, \dots, e_n)$ et $M = \mathrm{mat}_{\mathscr{B}}(u)$. Alors, par définition,

$$\det u = \det_{\mathscr{B}}(u(e_1), \dots, u(e_n)) = \det M,$$

car les vecteurs colonnes représentant les $u(e_j)$ dans la base $\mathscr{B}$ sont les colonnes de M. ■

On déduit de la sous-section précédente les propriétés suivantes sur le déterminant de matrices.

Proposition 19.17.

- ▹ Pour toutes matrices M, $N \in \mathscr{M}_n(\mathbb{K})$, $\det(MN) = \det M \det N$.
- ▹ Pour toute matrice $M \in \mathscr{M}_n(\mathbb{K})$, M est inversible si, et seulement si, $\det M \neq 0$.

Corollaire 19.18.

Pour toutes matrices $M \in \mathscr{M}_n(\mathbb{K})$ et $P \in \mathrm{GL}_n(\mathbb{K})$, $\det(PMP^{-1}) = \det P$. Autrement dit, deux matrices semblables ont même déterminant.

La propriété suivante permet de travailler autant avec les colonnes (voir la définition) qu'avec les lignes.

Proposition 19.19.

Pour toute matrice $M \in \mathscr{M}_n(\mathbb{K})$, $\det(M^{\mathrm{T}}) = \det(M)$.

Démonstration

Par définition, en notant $M = (m_{i,j})_{i,j\in[1,n]}$,

$$\begin{aligned}
\det(M^{\mathrm{T}}) &= \sum_{\sigma\in S_n} \varepsilon(\sigma)\prod_{k=1}^{n} m_{k,\sigma(k)} = \sum_{\sigma\in S_n} \varepsilon(\sigma)\prod_{j=1}^{n} m_{\sigma^{-1}(j),\sigma(\sigma^{-1}(j))}\\
&= \sum_{\sigma\in S_n} \varepsilon(\sigma)\prod_{j=1}^{n} m_{\sigma^{-1}(j),j} = \sum_{\sigma\in S_n} \varepsilon(\sigma^{-1})\prod_{j=1}^{n} m_{\sigma^{-1}(j),j}\\
&= \sum_{\sigma^{-1}\in S_n} \varepsilon(\sigma^{-1})\prod_{j=1}^{n} m_{\sigma^{-1}(j),j} = \det M
\end{aligned}$$

■

3. Quelques méthodes de calcul

3.1. Opérations sur les lignes/colonnes

La première idée pour calculer un déterminant consiste à utiliser les opérations élémentaires sur les lignes ou les colonnes et à utiliser les propriétés de forme n-linéaire alternée de la proposition 19.6. Il s'agit de se ramener à un déterminant simple que l'on sait calculer.

Exemple

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 3 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ 4 & -3 & -6 \\ 7 & -6 & -18 \end{vmatrix} \qquad \begin{matrix} C_2 \leftarrow C_2 - 2C_1 \\ C_3 \leftarrow C_3 - 3C_1 \end{matrix}$$

$$= \begin{vmatrix} 1 & 0 & 0 \\ 4 & -3 & 0 \\ 7 & -6 & -6 \end{vmatrix} \qquad C_3 \leftarrow C_3 - 2C_2$$

$$= 18.$$

Exemple

Soit $a \in \mathbb{C}$.

$$\begin{vmatrix} a & 1 & \cdots & 1 \\ 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \cdots & 1 & a \end{vmatrix} = \begin{vmatrix} a+n-1 & 1 & \cdots & \cdots & 1 \\ a+n-1 & a & \ddots & & \vdots \\ \vdots & 1 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 1 \\ a+n-1 & 1 & \cdots & 1 & a \end{vmatrix} \qquad C_1 \leftarrow \sum_{i=1}^{n} C_i$$

$$= (a+n-1) \begin{vmatrix} 1 & 1 & \cdots & \cdots & 1 \\ 1 & a & \ddots & & \vdots \\ \vdots & 1 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 1 \\ 1 & 1 & \cdots & 1 & a \end{vmatrix}$$

en utilisant la linéarité en la première colonne. Ainsi,

$$\begin{vmatrix} a & 1 & \cdots & 1 \\ 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \cdots & 1 & a \end{vmatrix} = (a+n-1) \begin{vmatrix} 1 & 0 & \cdots & \cdots & 0 \\ 1 & a-1 & \ddots & & \vdots \\ \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 1 & 0 & \cdots & 0 & a-1 \end{vmatrix} \qquad \begin{matrix} C_i \leftarrow C_i - C_1 \\ \text{pour } i \geq 2 \end{matrix}$$

$$= (a+n-1)(a-1)^{n-1}$$

Exemple

Soit $M = (m_{i,j})_{i,j\in[1,n]} \in \mathscr{M}_n(\mathbb{K})$. Considérons $\widetilde{M} = ((-1)^{i+j} m_{i,j})_{i,j\in[1,n]}$ et calculons $\det\widetilde{M}$. En effectuant successivement les opérations de dilatations $C_j \leftarrow (-1)^j C_j$ et $L_i \leftarrow (-1)^i L_i$ pour tout $i, j \in [\![1, n]\!]$, on obtient

$$\det\widetilde{M} = \Big[\prod_{i=1}^{n}(-1)^i\Big]\Big[\prod_{j=1}^{n}(-1)^j\Big]\det M,$$

d'où $\det\widetilde{M} = \det M$.

3.2. Développement par rapport à une ligne/colonne

Définition 19.20.

Soit $M \in \mathscr{M}_n(\mathbb{K})$ et $i, j \in [\![1, n]\!]$.
▹ Le mineur de position (i, j) de M est le scalaire, noté $\Delta_{i,j}$ défini comme le déterminant de la matrice $(m_{k,l})_{k\neq i, l\neq j}$ (c'est-à-dire de la matrice déduite de M en supprimant la ligne i et la colonne j).
▹ Le cofacteur de position (i, j) de M est le scalaire $(-1)^{i+j}\Delta_{i,j}$.
▹ La comatrice de M est la matrice notée $\mathrm{Com}(M)$ dont les coefficients sont les cofacteurs de M.

Proposition 19.21.

Soit $M \in \mathscr{M}_n(\mathbb{K})$.

- (Développement selon une colonne) $\forall j \leq n, \qquad \det M = \sum_{i=1}^{n}(-1)^{i+j} m_{i,j}\Delta_{i,j}$.
- (Développement selon une ligne) $\forall i \leq n, \qquad \det M = \sum_{j=1}^{n}(-1)^{i+j} m_{i,j}\Delta_{i,j}$.

Démonstration

D'après l'invariance du déterminant par transposition, il suffit de prouver l'une de ces deux formules. Commençons par fixer j et utiliser la linéarité du déterminant en sa $j^{\text{ième}}$ composante. On obtient alors, en notant $(E_i)_{i\in[1,n]}$ la base canonique de $\mathscr{M}_{n,1}(\mathbb{K})$ et $(C_i)_{i\in[1,n]}$ les colonnes de M,

$$\det M = \sum_{i=1}^{n} m_{i,j}\det(C_1, \dots, C_{j-1}, E_i, C_{j+1}, \dots, C_n).$$

En utilisant l'antisymétrie et la permutation $(j \dots n)$, on obtient alors

$$\det M = \sum_{i=1}^{n} m_{i,j}(-1)^{n-j}\det(C_1, \dots, C_{j-1}, C_{j+1}, \dots, C_n, E_i).$$

Notons, pour tout vecteur $X = (x_k)_{k\leq n}$, $\widetilde{X}$ le vecteur de coordonnées $(x_{\sigma(k)})_{k\leq n}$ où $\sigma = (i \ \dots \ n)$.
En utilisant la propriété d'antisymétrie sur les lignes, on obtient que

$$\det(C_1, \dots, C_{j-1}, C_{j+1}, \dots, C_n, E_i) = (-1)^{n-i}\det(\widetilde{C_1}, \dots, \widetilde{C_{j-1}}, \widetilde{C_{j+1}}, \dots, \widetilde{C_n}, E_n).$$

Il nous reste à calculer ce dernier déterminant par la formule explicite pour trouver qu'il vaut $\Delta_{i,j}$ (on remarque que pour qu'une permutation S_n contribue à ce déterminant, il faut qu'elle envoie n sur n et on peut ainsi l'identifier à une permutation de S_{n-1} de même signature). En combinant tous ces résultats, on obtient la formule annoncée. ■

Le développement par rapport à une ligne ou une colonne est très pratique pour calculer un déterminant avec de nombreux coefficients nuls mais peut aussi s'avérer utile pour raisonner par récurrence sur des déterminants de forme bien identifiée.

Exemple

Soit a, $b \in \mathbb{C}$. Calculons, pour tout $n \in \mathbb{N} \setminus \{0\}$, le déterminant

$$D_n = \begin{vmatrix} a+b & a & 0 & \cdots & 0 \\ b & a+b & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & a \\ 0 & \cdots & 0 & b & a+b \end{vmatrix}.$$

Remarquons que $D_1 = a+b$ et $D_2 = a^2+b^2+ab$ puis que, pour tout $n \geq 3$, on a, par développement par rapport à la première ligne,

$$D_n = (a+b)D_{n-1} - a \begin{vmatrix} b & a & 0 & \cdots & 0 \\ 0 & a+b & a & \ddots & \vdots \\ 0 & b & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & a \\ 0 & \cdots & 0 & b & a+b \end{vmatrix}.$$

En développant désormais le dernier déterminant par rapport à la première colonne, on obtient

$$D_n = (a+b)D_{n-1} - ab\,D_{n-2}.$$

On est ainsi ramené à l'étude d'une suite récurrente linéaire d'ordre 2 dont les racines de l'équation caractéristique sont a et b.

▹ Si $a \neq b$, alors il existe des scalaires λ, μ tels que

$$\forall n \geq 1, \qquad D_n = \lambda a^n + \mu b^n,$$

et les valeurs initiales entraînent $\lambda = \frac{-a}{b-a}$ et $\mu = \frac{b}{b-a}$ et donc

$$\forall n \geq 1, \qquad D_n = \frac{-a}{b-a}a^n + \frac{b}{b-a}b^n = \frac{b^{n+1}-a^{n+1}}{b-a}.$$

▹ Si $a = b$, alors il existe des scalaires λ, μ tels que

$$\forall n \geq 1, \qquad D_n = (\lambda + n\mu)a^n,$$

et les valeurs initiales entraînent $\lambda = \mu = 1$ et donc

$$\forall n \geq 1, \qquad D_n = (1+n)a^n.$$

On remarque que ce cas correspond au passage à la limite quand b tend vers a ce qui est logique car la fonction $b \mapsto D_n$ est continue car polynomiale.

Exemple Déterminant de Vandermonde

Soit $a_1, \dots, a_n$ des complexes. Considérons le déterminant défini par

$$V_n(a_1, \dots, a_n) = \begin{vmatrix} 1 & a_1 & \cdots & a_1^{n-1} \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ 1 & a_n & \cdots & a_n^{n-1} \end{vmatrix}.$$

Commençons par simplifier le déterminant en faisant les opérations élémentaires sur les colonnes $C_i \leftarrow C_i - a_1 C_{i-1}$ pour i allant de n à 2. Alors

$$V_n(a_1, \dots, a_n) = \begin{vmatrix} 1 & 0 & \cdots & 0 \\ \vdots & a_2 - a_1 & \cdots & a_2^{n-2}(a_2 - a_1) \\ \vdots & \vdots & & \vdots \\ 1 & a_n - a_1 & \cdots & a_n^{n-2}(a_n - a_1) \end{vmatrix}.$$

Ensuite, on développe par rapport à la première ligne :

$$\begin{aligned} V_n(a_1, \dots, a_n) &= \begin{vmatrix} a_2 - a_1 & \cdots & a_2^{n-2}(a_2 - a_1) \\ \vdots & & \vdots \\ a_n - a_1 & \cdots & a_n^{n-2}(a_n - a_1) \end{vmatrix} \\ &= \left(\prod_{k=2}^{n} (a_k - a_1) \right) V_{n-1}(a_2, \dots, a_n). \end{aligned}$$

Par récurrence, on obtient que

$$V_n(a_1, \dots, a_n) = \prod_{j>i} (a_j - a_i).$$

Résumons ces formules de développement de manière plus compacte.

Proposition 19.22.

Pour toute matrice $M \in \mathscr{M}_n(\mathbb{K})$, $\mathrm{Com}(M)^{\mathrm{T}} M = M \mathrm{Com}(M)^{\mathrm{T}} = \det(M) \mathrm{I}_n$.

Remarque

Cette formule (dite de la comatrice) s'avère inutile en pratique pour calculer le cas échéant l'inverse de M puisqu'elle nécessite de lourds calculs de déterminants alors que la méthode du pivot de Gauss est plus rapide sauf peut-être pour les matrices de $\mathrm{GL}_2(\mathbb{K})$:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Démonstration

Le terme en position (i, j) de $\mathrm{Com}(M)^{\mathrm{T}} M$ est $\sum_{k=1}^{n} (-1)^{i+k} \Delta_{k,i} m_{k,j}$ et il est égal à, après développement par rapport à la $i^{\text{ème}}$ colonne,

$$\det(C_1(M), \dots, C_{i-1}(M), C_j(M), C_{i+1}(M), \dots, C_n(M)),$$

où $(C_j(M))_{j\leq n}$ désignent les colonnes de M. Ainsi, ce coefficient est égal à $\delta_{i,j}\det M$ ce qui établit la formule de la comatrice. ■

3.3. Utilisation de la formule explicite

La formule explicite définissant le déterminant est peu maniable en pratique sauf pour des matrices de petite taille (pour $n=3$, c'est la règle de Sarrus). Toutefois, on peut l'exploiter et utiliser le fait que le déterminant est une fonction polynomiale en chacun de ses coefficients. Illustrons cette propriété sur différents exemples classiques.

Exemple Déterminant de Vandermonde

Soit $a_1,\dots,a_n$ des complexes. Considérons le déterminant défini par

$$V_n(a_1,\dots,a_n)=\begin{vmatrix} 1 & a_1 & \cdots & a_1^{n-1} \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ 1 & a_n & \cdots & a_n^{n-1} \end{vmatrix}.$$

L'application $a_n \mapsto V_n(a_1,\dots,a_n)$ est une fonction polynomiale de degré au plus $n-1$.
▹ Si deux des complexes $a_1,\dots,a_n$ sont égaux, alors le déterminant est nul car les lignes sont liées (deux sont égales).
▹ Sinon, le polynôme admet $n-1$ racines distinctes et son coefficient dominant est $V_{n-1}(a_1,\dots,a_{n-1})$ d'où

$$V_n(a_1,\dots,a_n)=V_{n-1}(a_1,\dots,a_{n-1})\prod_{k=1}^{n-1}(a_n-a_k).$$

Par récurrence, on obtient que

$$V_n(a_1,\dots,a_n)=\prod_{j>i}(a_j-a_i).$$

Remarque

Complétons l'exemple précédent : dans le cas où les complexes $a_1,\dots,a_n$ sont deux à deux disjoints, la matrice considérée est la matrice de passage (donc inversible) entre la base canonique et la base des polynômes interpolateurs de Lagrange aux points $a_1,\dots,a_n$: le déterminant est donc non nul.

Exemple

Soit $a,b,c\in\mathbb{C}$. Calculons le déterminant suivant

$$D=\begin{vmatrix} a & c & \cdots & c \\ b & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & c \\ b & \cdots & b & a \end{vmatrix}.$$

Pour cela, introduisons la fonction polynomiale

$$f : x \mapsto \begin{vmatrix} a+x & c+x & \cdots & c+x \\ b+x & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & c+x \\ b+x & \cdots & b+x & a+x \end{vmatrix}.$$

En effectuant les opérations sur les colonnes $C_j \leftarrow C_j - C_1$ pour $j \geq 2$, on vérifie que f est de degré au plus 1.
▹ Si $b \neq c$, alors $f(-b) = (a-b)^n$ et $f(-c) = (a-c)^n$ (ce sont des déterminants triangulaires). On en déduit ainsi l'expression de f et donc la valeur de $D = f(0)$

$$D = \frac{(a-c)^n b - (a-b)^n c}{b-c}.$$

▹ Pour calculer D lorsque $b = c$, il suffit de faire tendre c vers b puisque D est une fonction polynomiale donc continue en la variable b. Ainsi

$$D = (a-b)^{n-1}(a+(n-1)b).$$

3.4. Calculs par blocs

Exploitons un instant le fait que le produit de matrices par blocs s'effectue comme un produit matriciel coefficient par coefficient.

Exemple

Soit $A \in \mathscr{M}_r(\mathbb{K})$, $B \in \mathscr{M}_{r,n-r}(\mathbb{K})$. Alors,

$$\begin{vmatrix} A & B \\ 0_{n-r,r} & \mathrm{I}_{n-r} \end{vmatrix} = \det(A).$$

En effet, le résultat s'obtient en développant successivement par rapport aux $n-r$ dernières lignes.

Proposition 19.23.

Soit $A \in \mathscr{M}_r(\mathbb{K})$, $B \in \mathscr{M}_{r,n-r}(\mathbb{K})$ et $C \in \mathscr{M}_{n-r}(\mathbb{K})$. Alors

$$\begin{vmatrix} A & B \\ 0_{n-r,r} & C \end{vmatrix} = \det(A)\det(C).$$

Démonstration

Commençons par décomposer la matrice étudiée pour nous ramener à l'exemple précédent

$$\begin{bmatrix} A & B \\ 0_{n-r,r} & C \end{bmatrix} = \begin{bmatrix} I_r & 0_{n-r,r} \\ 0_{n-r,r} & C \end{bmatrix} \begin{bmatrix} A & B \\ 0_{n-r,r} & I_{n-r} \end{bmatrix}.$$

Ensuite, chacun de ces deux déterminants se calcule par développant par rapport aux r premières lignes ou aux $n-r$ dernières lignes d'où le résultat. ■

Exemple

Pour calculer le déterminant de la matrice

$$A = \begin{bmatrix} 1 & 0 & 3 & 5 & 6 \\ 2 & 2 & 7 & 0 & 8 \\ 0 & 0 & 2 & 0 & 9 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix}$$

on peut remarquer qu'elle est triangulaire par blocs avec les blocs diagonaux

$$\begin{bmatrix} 1 & 0 \\ 2 & 2 \end{bmatrix}, \qquad \begin{bmatrix} 2 & 0 & 9 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

de déterminant 2 et 8 respectivement. Le déterminant de A vaut donc 16.

Remarque

La proposition s'étend évidemment à une matrice triangulaire par blocs avec plus de deux blocs sur la diagonale. En revanche, le résultat devient faux sans hypothèse additionnelle pour des matrices non-triangulaires par blocs.

4. Applications

4.1. Formules de Cramer

Pour résoudre un système linéaire, on sait utiliser la méthode du pivot de Gauss qui est une succession d'opérations élémentaires. Toutefois, il existe, lorsque la solution d'un système linéaire est unique, une formule close pour exprimer celle-ci. L'intérêt pratique du résultat est quasi-nul.

Proposition 19.24. Formules de Cramer

Soit $AX = B$ un système linéaire de Cramer. L'unique solution $(x_i)_{i\in[1,n]}$ est donnée par

$$\forall i \in [\![1, n]\!], \qquad x_i = \frac{\det(C_1(A), \dots, C_{i-1}(A), B, C_{i+1}(A), \dots, C_n(A))}{\det(A)}.$$

Démonstration

Comme $AX = B$, on obtient que $B = \sum_{j=1}^{n} x_j C_j(A)$. Remplaçons cette expression dans le second membre

$$\begin{aligned} \det(C_1(A), \dots, C_{i-1}(A), B, C_{i+1}(A), \dots, C_n(A)) &= \det(C_1(A), \dots, C_{i-1}(A), \sum_{j=1}^{n} x_j C_j(A), C_{i+1}(A), \dots, C_n(A)) \\ &= \sum_{j=1}^{n} x_j \det(C_1(A), \dots, C_{i-1}(A), C_j(A), C_{i+1}(A), \dots, C_n(A)) \\ &= x_i \det(C_1(A), \dots, C_{i-1}(A), C_i(A), C_{i+1}(A), \dots, C_n(A)), \end{aligned}$$

en utilisant successivement la linéarité en la $i^{\text{ième}}$ variable et le caractère alterné. D'où le résultat. ■

Écrivons ces formules dans l'un des seuls cas où elles sont pratiques : en dimension 2.

Exemple

L'unique solution du système

$$\begin{cases} ax + by = \alpha \\ cx + dy = \beta \end{cases}$$

avec $ad - bc \neq 0$ vérifie

$$x = \frac{\alpha d - \beta b}{ad - bc}, \qquad y = \frac{a\beta - c\alpha}{ad - bc}.$$

4.2. Caractérisation du rang

Définition 19.25.

Soit $M = (m_{i,j})_{i\in[1,n], j\in[1,p]} \in \mathscr{M}_{n,p}(\mathbb{K})$ et $r \leq \min(n,p)$. Un déterminant extrait d'ordre r de M est un scalaire de la forme $\det(m_{i,j})_{i\in I_r, j\in J_r}$ où I_r et J_r sont des parties de $[\![1,n]\!]$ et $[\![1,p]\!]$ de cardinal r.

Remarque

Par définition, les déterminants extraits d'ordre $n-1$ sont exactement les mineurs.

Proposition 19.26.

Soit $M = (m_{i,j})_{i\in[1,n], j\in[1,p]} \in \mathscr{M}_{n,p}(\mathbb{K})$ et $r \leq \min(n,p)$. Le rang de M est supérieur ou égal à r si, et seulement s'il existe un déterminant extrait d'ordre r de M non nul.

Corollaire 19.27.

Soit $M = (m_{i,j})_{i\in[1,n], j\in[1,p]} \in \mathscr{M}_{n,p}(\mathbb{K})$ et $r \leq \min(n,p)$. Le rang de M est égal à r si, et seulement s'il existe un déterminant extrait d'ordre r de M non nul et tous les déterminants extraits d'ordre $r+1$ de M sont nuls.

Démonstration

Séparons les deux implications.

($\Rightarrow$) Si le rang de M est supérieur ou égal à r, alors il existe r colonnes de M linéairement indépendantes. Notons I_r l'ensemble des indices de telles colonnes et considérons la matrice $\widetilde{M}$ obtenue avec ces r colonnes. La matrice $\widetilde{M}$ est de rang r donc il existe r lignes linéairement indépendantes ; notons J_r l'ensemble des indices de telles lignes. Alors, la matrice $(m_{i,j})_{i\in I_r, j\in J_r}$ est de taille $r \times r$ et de rang r donc inversible : son déterminant est non nul.

($\Leftarrow$) Supposons qu'il existe I_r et J_r des parties de $[\![1,n]\!]$ et $[\![1,p]\!]$ de cardinal r telles que

$$\det(m_{i,j})_{i\in I_r, j\in J_r} \neq 0$$

et montrons que les colonnes $(C_i(M))_{i\in I_r}$ forment une famille libre. Supposons qu'il existe $(\lambda_i)_{i\in I_r} \in \mathbb{C}^r$ tel que $\sum_{i\in I_r} \lambda_i C_i(M) = 0_{n,1}$. Alors, le vecteur des $(\lambda_i)_{i\in I_r}$ appartient au noyau de la matrice inversible $(m_{i,j})_{i\in I_r, j\in J_r}$ donc est nul. ■

Exemple

Une matrice de la forme suivante

$$\begin{bmatrix} \star & \cdots & \cdots & \cdots & \star \\ 1 & \ddots & & & \vdots \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & \star \end{bmatrix}$$

est de rang supérieur ou égal à $n-1$ puisque le mineur $\Delta_{1,n}$ est égale à 1 (il est triangulaire) donc non nul.

Exemple

Déterminons le rang de la comatrice $\mathrm{Com}(M)$ d'une matrice $M \in \mathscr{M}_n(\mathbb{C})$. Distinguons trois cas.

▹ Si $\mathrm{rg}(M) = n$, c'est-à-dire si M est inversible, alors $\mathrm{Com}(M)$ est aussi inversible d'après la formule de la comatrice (d'inverse $\frac{1}{\det(M)} M^{\mathrm{T}}$) donc $\mathrm{rg}(\mathrm{Com}(M)) = n$.

▹ Si $\mathrm{rg}(M) = n-1$, alors, d'après la formule de la comatrice, $M\,\mathrm{Com}(M)^{\mathrm{T}} = 0_n$ donc toutes les transposées des lignes de $\mathrm{Com}(M)$ sont dans $\mathrm{Ker}\, M$ qui est de dimension 1. Par conséquent, l'espace engendré par les lignes de $\mathrm{Com}\, M$ est de dimension au plus 1 et $\mathrm{rg}(\mathrm{Com}(M)) \leq 1$. Par ailleurs, $\mathrm{Com}(M)$ est non nulle car il existe un déterminant extrait d'ordre $n-1$ de M qui est non nul (caractérisation du rang précédente). D'où $\mathrm{rg}(\mathrm{Com}(M)) = 1$.

▹ Si $\mathrm{rg}(M) \leq n-2$, alors, d'après la caractérisation précédente, tous les mineurs (qui sont des déterminants extraits d'ordre $n-1$) sont nuls donc $\mathrm{Com}(M) = 0_n$ et $\mathrm{rg}(\mathrm{Com}(M)) = 0$.

En résumé, pour $n \geq 2$,

$\mathrm{rg}(M)$	$\mathrm{rg}(\mathrm{Com}(M))$
n	n
$n-1$	1
$\leq n-2$	0

FICHE

SYNTHÈSE

Les résultats sur le groupe symétrique établis au chapitre précédent permettent d'étudier les formes n-linéaires alternées et de définir le déterminant dans une base.

Déterminant d'une famille de vecteurs

Soit E un espace vectoriel de dimension n dont une base est $\mathscr{B} = (e_1, \dots, e_n)$. Le déterminant dans la base $\mathscr{B}$ est la forme n-linéaire alternée définie par

$$\forall (x_1, \dots, x_n) \in E^n, \qquad f(x_1, \dots, x_n) = \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{j=1}^{n} x_{\sigma(j),j},$$

où, pour tout $j \in [\![1, n]\!]$, $x_j = \sum_{i=1}^{n} x_{i,j} e_i$.

On en déduit le déterminant d'une matrice ou d'un endomorphisme.

Déterminant d'une matrice, d'un endomorphisme

▹ Le déterminant d'une matrice de $\mathscr{M}_n(\mathbb{K})$ est le déterminant de ses colonnes dans la base canonique de $\mathscr{M}_{n,1}(\mathbb{K})$.
▹ Pour toutes matrices M, $N \in \mathscr{M}_n(\mathbb{K})$, $\det(MN) = \det(M)\det(N)$.

Déterminant d'un endomorphisme

▹ Le déterminant d'un endomorphisme de E est le déterminant de l'image d'une base de E dans cette base. Cette quantité ne dépend pas de la base choisie et est égale au déterminant d'une matrice de cet endomorphisme.
▹ Pour tous endomorphismes u, $v \in \mathscr{L}(E)$, $\det(u \circ v) = \det(u)\det(v)$.

Caractérisations

▹ Une famille de n vecteurs d'un espace de dimension n est une base si, et seulement si, son déterminant dans une base est non nul.
▹ Soit $M \in \mathscr{M}_n(\mathbb{K})$. Alors, M inversible si, et seulement si, $\det M \neq 0$.
▹ Soit $u \in \mathscr{L}(E)$. Alors, u inversible si, et seulement si, $\det u \neq 0$.

Une fois le déterminant défini, il faut savoir le calculer. Une méthode importante consiste à utiliser les opérations sur les lignes ou les colonnes.

Développement

▹ Soit $M \in \mathscr{M}_n(\mathbb{K})$.

- (Développement selon une colonne)

$$\forall j \leq n, \qquad \det M = \sum_{i=1}^{n} (-1)^{i+j} m_{i,j} \Delta_{i,j}.$$

- (Développement selon une ligne)

$$\forall i \leq n, \qquad \det M = \sum_{j=1}^{n} (-1)^{i+j} m_{i,j} \Delta_{i,j}.$$

▹ Soit $M \in \mathscr{M}_n(\mathbb{K})$.

$$\mathrm{Com}(M)^{\mathrm{T}} M = M \mathrm{Com}(M)^{\mathrm{T}} = \det(M) \mathrm{I}_n.$$

Il est important de connaître quelques déterminants classiques et de savoir utiliser le déterminant pour des applications simples.

EXERCICES

Vrai ou faux ?

		Vrai	Faux
a)	$\forall \lambda \in \mathbb{R},\ \det(\lambda \mathrm{I}_n) = \lambda$.	☐	☐
b)	Le déterminant d'une matrice triangulaire est le produit des éléments diagonaux.	☐	☐
c)	$\det(a_i^{j-1})_{i,j} = \prod_{j>i}(a_j - a_i)$.	☐	☐
d)	Pour $\sigma \in S_n$, $\det(\delta_{i,\sigma(j)}) = \varepsilon(\sigma)$.	☐	☐
e)	$\forall A, B, C, D \in \mathscr{M}_n(\mathbb{R})$, $\det\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \det A \det D - \det B \det C$.	☐	☐
f)	$\forall A,\ B,\ C \in \mathscr{M}_n(\mathbb{R})$, $\det\begin{bmatrix} A & B \\ 0_n & C \end{bmatrix} = \det A \det C$.	☐	☐
g)	$\forall A, B \in \mathscr{M}_n(\mathbb{R}),\ \det(AB) = \det(A)\det(B)$.	☐	☐
h)	$\forall A, B \in \mathscr{M}_n(\mathbb{R}),\ \det(A+B) = \det(A) + \det(B)$.	☐	☐
i)	$\forall A \in \mathscr{M}_n(\mathbb{R}),\ \det(A^{\mathrm{T}}) = \det(A)$.	☐	☐
j)	$\forall A \in \mathscr{M}_n(\mathbb{R}),\ \det(A^{\mathrm{T}}A) \geq 0$.	☐	☐
k)	$\det(\mathrm{GL}_n(\mathbb{R})) = \mathbb{R}^*$.	☐	☐
l)	Deux matrices semblables ont même déterminant.	☐	☐
m)	Deux matrices équivalentes ont même déterminant.	☐	☐
n)	Une matrice est inversible si, et seulement si, son déterminant est ± 1.	☐	☐
o)	La comatrice d'une matrice triangulaire est triangulaire.	☐	☐
p)	Pour tout $A \in \mathscr{M}_n(\mathbb{R})$, la fonction $t \mapsto \det(A + t\mathrm{I}_n)$ est polynomiale de degré n.	☐	☐

Exercices d'application du chapitre 19

●○○ **Exercice 1**

Soit $n \geq 3$ et $a_1, a_2, \ldots, a_n, b_1, b_2, \ldots, b_n \in \mathbb{C}$. Calculer $\det(a_i - b_j)$.

●○○ **Exercice 2**

Soit a, $b \in \mathbb{C}$. Donner une forme factorisée au déterminant suivant $\begin{vmatrix} 1 & a & b \\ a & 1 & b \\ b & a & 1 \end{vmatrix}$.

●○○ **Exercice 3**

Notons, pour tout $k \in [\![1, n]\!]$, $S_k = \sum_{i=1}^{k} i$. Calculer le déterminant

$$\begin{vmatrix} S_1 & S_1 & S_1 & \cdots & S_1 \\ S_1 & S_2 & S_2 & \cdots & S_2 \\ S_1 & S_2 & S_3 & \cdots & S_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ S_1 & S_2 & S_3 & \cdots & S_n \end{vmatrix}.$$

●○○ **Exercice 4**

Calculer le déterminant de la matrice $M_n \in \mathscr{M}_n(\mathbb{R})$ de coefficients $m_{i,j}$ égaux à 1 si $i = j$, $i = 1$ ou $j = 1$, nuls sinon.

$$\begin{vmatrix} 1 & \cdots & \cdots & 1 \\ \vdots & \ddots & & \\ \vdots & & \ddots & \\ 1 & & & 1 \end{vmatrix}.$$

●○○ **Exercice 5**

Calculer les déterminants suivants

1. $\det\left(\binom{i}{j-1}\right)$.
2. $\det\left(\binom{i+j-2}{j-1}\right)$

●○○ **Exercice 6**

Soit $M \in \mathscr{M}_n(\mathbb{R})$ et $M' \in \mathscr{M}_n(\mathbb{R})$ telle que, pour tout $j \leq n$, $C_j(M') = \sum_{k \neq j} C_k(M)$.

Calculer $\det M'$ en fonction de $\det M$.

●●○ **Exercice 7**

Soit $n \geq 2$ et $P \in \mathbb{R}_{n-2}[X]$. Calculer le déterminant

$$\begin{vmatrix} P(1) & P(2) & \cdots & P(n) \\ P(2) & P(3) & \cdots & P(n+1) \\ \vdots & \vdots & & \vdots \\ P(n) & P(n+1) & \cdots & P(2n-1) \end{vmatrix}.$$

●●● **Exercice 8**

Soit $P \in \mathbb{R}[X]$ de degré $n-1$. Calculer $\det(P(i+j))_{i,j \leq n}$.

●○○ **Exercice 9**

Calculer le déterminant de l'endomorphisme u de $\mathscr{M}_n(\mathbb{R})$ défini par $u : M \mapsto M + \operatorname{tr}(M)\mathrm{I}_n$.

EXERCICES

●○○ **Exercice 10**

Soit $E = \{x \mapsto P(x)e^x,\ P \in \mathbb{R}_n[X]\}$. Déterminer le déterminant de l'endomorphisme de dérivation sur E.

●○○ **Exercice 11**

Soit $M \in \mathscr{M}_n(\mathbb{R})$ à coefficients dans $\{\pm 1\}$. Montrer que l'entier $\det M$ est divisible par 2^{n-1}.

●●○ **Exercice 12**

Déterminer selon les valeurs de $a \in \mathbb{C}$ le rang de la matrice M_a définie par

$$M_a = \begin{bmatrix} 1 & a & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & & \ddots & \ddots & a \\ a & 0 & \dots & 0 & 1 \end{bmatrix}.$$

●●● **Exercice 13**

Soit $M = (m_{i,j})_{i,j}$ une matrice. Notons, pour toute permutation $\sigma \in S_n$,

$$C_\sigma = \prod_{i=1}^{n} m_{i,\sigma(i)}.$$

Montrer que si C_σ est une quantité non nulle indépendante de $\sigma \in S_n$, alors $\operatorname{rg} M = 1$.

●●● **Exercice 14**

Déterminer les matrices $A \in \mathscr{M}_n(\mathbb{R})$ telles que

$$\forall M \in \mathscr{M}_n(\mathbb{R}), \qquad \det(A+M) = \det A + \det M.$$

●●● **Exercice 15**

Soit $A \in \mathscr{M}_{p,q}(\mathbb{K})$ et $B \in \mathscr{M}_{q,p}(\mathbb{K})$. Montrer que $\det(\mathrm{I}_p - AB) = \det(\mathrm{I}_q - BA)$.

●●● **Exercice 16**

Soit $M \in \mathscr{M}_n(\mathbb{R})$.

1. Montrer que, si M est inversible, alors il n'existe pas de matrice inversible P telle que

$$\forall z \in \mathbb{C}, \qquad \det(P - zM) \neq 0.$$

2. Montrer que ce résultat est faux si M est de rang $r < n$.

●●● **Exercice 17**

Soit $A, B, C \in \mathscr{M}_3(\mathbb{R})$ telles que $\det A = \det B = \det C$ et $\det(A + iB) = \det(C + iA)$.
Montrer que $\det(A+B) = \det(C+A)$.

●●○ **Exercice 18**

Soit A et B deux matrices réelles carrées d'ordre n et semblables sur $\mathbb{C}$. Montrer que A et B sont semblables sur $\mathbb{R}$.

Exercices d'approfondissement du chapitre 19

●●○ **Exercice A**

1. Montrer l'inversibilité de la matrice $M \in \mathscr{M}_{2n}(\mathbb{C})$ dont les coefficients sont
$$m_{i,j} = \begin{cases} 1 & \text{si } i \neq j, \\ 0 & \text{sinon.} \end{cases}$$
2. En déduire que toute matrice de $\mathscr{M}_{2n}(\mathbb{C})$ dont les coefficients sont nuls sur la diagonale et ± 1 en dehors est inversible.

●●○ **Exercice B**

Soit $A,\ B \in \mathscr{M}_2(\mathbb{Z})$ telles que $A + kB$ soit inversible avec un inverse dans $\mathscr{M}_2(\mathbb{Z})$ pour tout $k \in \{0, \dots, 4\}$. Montrer que $A + 5B$ est inversible avec un inverse dans $\mathscr{M}_2(\mathbb{Z})$.

●●● **Exercice C**

Soit $M = (m_{i,j})_{i,j} \in \mathscr{M}_n(\mathbb{R})$ telle que $m_{1,1} \neq 0$. Posons, pour tous $i, j \geq 2$,
$$d_{i,j} = \begin{vmatrix} m_{1,1} & m_{1,j} \\ m_{i,1} & m_{i,j} \end{vmatrix}.$$
Montrer que
$$\det M = \frac{1}{m_{1,1}^{n-2}} \det(d_{i,j})_{i,j \geq 2}.$$

●●○ **Exercice D Déterminant anti-circulant**

Soit $b_1, \dots, b_n \in \mathbb{C}$; considérons la matrice $A \in \mathscr{M}_n(\mathbb{C})$ dont les coefficients sont $a_{i,j} = b_k$ si $i + j - 1 = k[n]$. Notons $\omega = e^{\frac{2i\pi}{n}}$, $U \in \mathscr{M}_n(\mathbb{R})$ la matrice de coefficients $u_{i,j} = \omega^{(i-1)(j-1)}$ pour tous $i, j \leq n$ et φ l'application
$$\varphi : \begin{cases} \mathbb{C} & \to & \mathbb{C} \\ z & \mapsto & \displaystyle\sum_{k=1}^{n} b_k z^{k-1} \end{cases}$$

1. Calculer les coefficients de AU à l'aide de la fonction φ.
2. En déduire que $\det A = \epsilon \prod_{k=1}^{n} \varphi(\omega^{k-1})$ avec ϵ un signe à déterminer en fonction de la parité de n.
3. Calculer φ puis $\det A$ dans les cas suivants :
   a. $b_k = a^{k-1}$
   b. $b_k = \binom{n-1}{k-1}$

●●○ **Exercice E Déterminant de Casorati**

Soit E un ensemble non-vide. Pour toute famille $(f_1, \dots, f_n)$ de $\mathbb{R}^E$, on appelle déterminant de Casorati de $(f_1, \dots, f_n)$ l'application
$$\mathscr{C}_{f_1,\dots,f_n} : \begin{cases} E^n & \to & \mathbb{R} \\ (x_1, \dots, x_n) & \mapsto & \det(f_i(x_j))_{i,j \leq n} \end{cases}$$

1. a. Montrer que si la famille $(f_1, \dots, f_n)$ est liée, alors $\mathscr{C}_{f_1,\dots,f_n}$ est l'application nulle.
   b. Réciproquement, montrer que, si $\mathscr{C}_{f_1,\dots,f_n} = 0$ et $\mathscr{C}_{f_1,\dots,f_{n-1}} \neq 0$, alors $f_n \in \mathrm{Vect}(f_1, \dots, f_{n-1})$.
   Indication : on pourra développer le déterminant par rapport à la dernière colonne.
   c. Conclure que si $\mathscr{C}_{f_1,\dots,f_n} = 0$, alors la famille $(f_1, \dots, f_n)$ est liée.

EXERCICES

2. On suppose dorénavant que $(f_1,\dots,f_n)$ est une famille libre de $\mathbb{R}^E$ et on note F le sous-espace de $\mathbb{R}^E$ engendré par ces fonctions.
   a. Justifier l'existence de $(x_1,\dots,x_n)$ tel que $\mathscr{C}_{f_1,\dots,f_n}(x_1,\dots,x_n)\neq 0$.
   Notons $(\widetilde{f_1},\dots,\widetilde{f_n})$ les applications de $\mathbb{R}^E$ définies, pour tout $j\leq n$, par
   $$\widetilde{f_j} : x \mapsto \mathscr{C}_{f_1,\dots,f_n}(x_1,\dots,x_{j-1},x,x_{j+1},\dots,x_n).$$
   b. Exprimer les fonctions $(\widetilde{f_1},\dots,\widetilde{f_n})$ dans la base $(f_1,\dots,f_n)$.
   En déduire que $(\widetilde{f_1},\dots,\widetilde{f_n})$ est une base de F.
   c. En déduire que $(g_1,\dots,g_n)$ soit une base de F si, et seulement si, $\mathscr{C}_{g_1,\dots,g_n}$ est une application non-nulle, colinéaire à $\mathscr{C}_{f_1,\dots,f_n}$.

●●● **Exercice F Permanent d'une matrice**

Le permanent d'une matrice $A=(a_{i,j})_{i,j}\in\mathscr{M}_n(\mathbb{R})$ est le réel $\operatorname{per}(A)$ défini par
$$\operatorname{per}(A)=\sum_{\sigma\in S_n}\prod_{i=1}^{n} a_{\sigma(i),i}.$$

1. Montrer que l'application permanent per, qui prend n vecteurs colonnes et renvoie le permanent de la matrice ainsi formée, est n-linéaire et symétrique.
2. Considérons une matrice $A\in\mathscr{M}_n(\mathbb{R})$ dont tous les coefficients sont positifs ou nuls. Montrer que si $\operatorname{per}(A)=0$, alors on peut extraire de A une matrice nulle de taille $r\times(n+1-r)$ avec $r\in[\![1,n]\!]$.
3. Soit $F=[\![1,n]\!]$ et $G=[\![1,p]\!]$ avec $p\leq n$ et $\Phi : G\to\mathscr{P}(F)$.
   Montrer que les propriétés suivantes sont équivalentes :
   - il existe une injection $\varphi : G\to F$ telle que, pour tout $k\in G$, $\varphi(k)\in\Phi(k)$;
   - pour toute partie G' de G,
   $$\operatorname{card}\Big(\bigcup_{k\in G'}\Phi(k)\Big)\geq \operatorname{card} G'.$$

   Indication : pour l'implication « délicate », on pourra introduire la matrice $A=(a_{i,j})\in\mathscr{M}_n(\mathbb{R})$ définie par
   $$a_{i,j}=\begin{cases}1 & \text{si } i>p,\\ 1 & \text{si } i\leq p \text{ et } j\in\Phi(i),\\ 0 & \text{sinon.}\end{cases}$$

Problèmes du chapitre 19

●●○ **Problème 1 Déterminants de Smith**

Pour toute fonction $f:\mathbb{N}^*\to\mathbb{R}$, $\sum\limits_{d|n} f(d)$ désigne la somme sur les diviseurs positifs d de n.

Partie A – Fonction de Mobius

1. a. Montrer qu'il existe une unique fonction $\mu:\mathbb{N}^*\to\mathbb{C}$ telle que $\mu(1)=1$ et
   $$\forall n\geq 2,\qquad \sum_{d|n}\mu(d)=0.$$

b. Soit $f : \mathbb{N}^* \to \mathbb{C}$ et F la fonction définie sur $\mathbb{N}^*$ par

$$\forall n \in \mathbb{N}^*, \qquad F(n) = \sum_{d|n} f(d).$$

Montrer que

$$\forall n \in \mathbb{N}^*, \qquad \sum_{d|n} \mu(d) F\Big(\frac{n}{d}\Big) = f(n).$$

2. a. Soit p un nombre premier. Montrer que $\mu(p) = -1$ puis calculer $\mu(p^k)$ pour $k \geq 2$.
   b. Comparer, pour $m, n \in \mathbb{N}^*$ premiers entre eux, les quantités $\mu(m)\mu(n)$ et $\mu(mn)$.
   c. En déduire l'ensemble des entiers $n \in \mathbb{N}^*$ tels que $\mu(n) = 1$.

Partie B – Calcul de déterminants

1. Soit $F : \mathbb{N}^* \to \mathbb{C}$. Considérons la matrice $A_F = (F(i \wedge j))_{i,j} \in \mathscr{M}_n(\mathbb{C})$.
   a. Soit $P \in \mathscr{M}_n(\mathbb{C})$ la matrice dont le coefficient en position i, j est 1 si i divise j, et 0 sinon. Calculer $\det P$.
   b. Trouver, à l'aide de la partie A, une fonction $f : \mathbb{N}^* \to \mathbb{C}$ telle que

$$A_F = P^{\mathrm{T}} \operatorname{diag}(f(1), \ldots, f(n)) P.$$

   c. En déduire $\det A_F$ en fonction de f.

Les trois questions suivantes sont indépendantes entres elles mais utilisent le résultat de la question précédente.

2. Calculer le « premier déterminant de Smith » $\det(i \wedge j)_{i,j}$.
3. Définissons $f(1) = 1$ et, pour tout $n \in \mathbb{N}^*$,

$$f(n) = \frac{1}{n} \prod_{p \in \mathscr{P}_n} (1 - p),$$

où $\mathscr{P}_n$ désigne l'ensemble des facteurs premiers positifs de n. Posons encore F définie par

$$\forall n \in \mathbb{N}^*, \qquad F(n) = \sum_{d|n} f(d).$$

   a. Montrer que, pour $m, n \in \mathbb{N}^*$ premiers entre eux, $f(mn) = f(m)f(n)$.
   Indication : on pourra exprimer $\mathscr{P}_{mn}$ en fonction de $\mathscr{P}_m$ et $\mathscr{P}_n$.
   b. En déduire que, pour $m, n \in \mathbb{N}^*$ premiers entre eux,

$$F(mn) = F(m)F(n).$$

   Indication : on pourra commencer par étudier l'ensemble des diviseurs positifs de mn.
   c. Prouver que, pour tout $n \in \mathbb{N}^*$, $F(n) = \frac{1}{n}$.
   d. En déduire le $\det\big(\frac{1}{i \wedge j}\big)_{i,j}$ puis le « second déterminant de Smith » $\det(i \vee j)_{i,j}$.
4. Calculer le déterminant de la matrice dont le coefficient en position i, j est la somme des diviseurs positifs communs à i et j.

PROBLÈMES

Problème 2 Théorème de Hankel-Sylvester

Définition 19.28.

Pour toute matrice $M = (m_{i,j}) \in \mathscr{M}_n(\mathbb{R})$, on définit

- $\Delta_{i_1}^{j_1}$ le mineur de M en position i_1, j_1, c'est-à-dire le déterminant de la matrice déduite de M en supprimant la ligne d'indice i_1 et la colonne d'indice j_1 ;
- $\Delta_{i_1,i_2}^{j_1,j_2}$ le déterminant de la matrice déduite de M en supprimant les lignes d'indice i_1 et i_2 et les colonnes d'indice j_1 et j_2.

Partie A – Formule de condensation de Desnanot-Jacobi

Dans cette partie, on cherche à montrer que, pour toute matrice $M \in \mathscr{M}_n(\mathbb{R})$ et tous $i \neq j$,

$$\Delta_{i,j}^{i,j} \det M = \Delta_i^i \Delta_j^j - \Delta_i^j \Delta_j^i. \tag{19.1}$$

1. Rappeler l'expression du produit matriciel $M\mathrm{Com}(M)^{\mathrm{T}}$.
2. Pour tout $i \neq j$, notons $M_{i,j}$ la matrice déduite de I_n où les colonnes d'indices i et j ont été remplacées par les colonnes correspondantes de $\mathrm{Com}(M)^{\mathrm{T}}$.
   a. Exprimer le déterminant de $M_{i,j}$ en fonction de Δ_i^i, Δ_j^j, Δ_i^j et Δ_j^i.
   b. Calculer le produit matriciel $MM_{i,j}$.
   c. Calculer $\det(MM_{i,j})$.
   d. En déduire la formule (19.1) dans le cas où M est inversible.
3. Déduire la formule (19.1) dans le cas général en considérant la matrice $M + x\mathrm{I}_n$ et en faisant tendre x vers 0.

Partie B – Déterminants de Hankel-Sylvester

Définition 19.29.

Soit $f : I \to \mathbb{C}$ une fonction de classe $\mathscr{C}^{2n}$. Le $n^{\text{ième}}$ déterminant de Sylvester $S_n(f)$ est la fonction qui associe, à tout $x \in I$, le déterminant de la matrice $(f^{(i+j-2)})_{i,j \leq n+1} \in \mathscr{M}_{n+1}(\mathbb{C})$.

1. Calculer les déterminants de Sylvester pour $f = \exp$.
2. Calculer le $n^{\text{ième}}$ déterminant de Sylvester pour la fonction $f : x \mapsto \sum_{k=1}^{n+1} a_k e^{\alpha_k x}$ et les complexes $a_1, \ldots, a_{n+1}, \alpha_1, \ldots, \alpha_{n+1}$.
   Indication : on pourra reconnaître un produit matriciel et exprimer le résultat à l'aide du déterminant de Vandermonde de $\alpha_1, \ldots, \alpha_{n+1}$.
3. Soit $f : I \to \mathbb{C}$ de classe $\mathscr{C}^{2n+2}$. Posons $M = (f^{(i+j-2)})_{i,j \leq n+2} \in \mathscr{M}_{n+2}(\mathbb{C})$, Δ_i^j les mineurs de M.
   a. Montrer que $(S_n f)' = \Delta_{n+1}^{n+2} = \Delta_{n+2}^{n+1}$ et que $(S_n f)'' = \Delta_{n+1}^{n+1}$.
   b. Déduire de la Partie A que
   $$S_{n+1} f S_{n-1} f = S_n f (S_n f)'' - (S_n f)'^2.$$
4. Montrer qu'une fonction f de classe $\mathscr{C}^{2n}$ est solution d'une équation différentielle linéaire d'ordre n, homogène à coefficients constants si, et seulement si, la fonction $S_n f$ est identiquement nulle.

Problème 3 Résultant

Définition 19.30.

Le résultant $\mathrm{Res}(P,Q)$ de deux polynômes $P=\sum_{k=0}^{p} a_k X^k$ et $Q=\sum_{k=0}^{q} b_k X^k$ avec $a_p \neq 0$ et $b_q \neq 0$ est le déterminant de la matrice suivante de $\mathscr{M}_{p+q}(\mathbb{C})$

$$\begin{bmatrix} a_0 & 0 & \cdots & 0 & b_0 & 0 & \cdots & 0 \\ a_1 & \ddots & \ddots & \vdots & b_1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots & \ddots & \ddots & 0 \\ \vdots & & \ddots & a_0 & \vdots & & \ddots & b_0 \\ a_p & & & a_1 & b_q & & & b_1 \\ 0 & \ddots & & \vdots & 0 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_p & 0 & \cdots & 0 & b_q \end{bmatrix}.$$

De manière informelle, pour obtenir le déterminant résultant, on écrit les coefficients du polynôme P en colonne en les décalant d'un cran vers le bas sur les $q=\deg Q$ premières colonnes. Ensuite, on procède de même avec les coefficients de Q, qui sont eux décalés sur les $p=\deg P$ dernières colonnes.

Partie A – Primalité relative

Les deux questions de cette partie montrent indépendamment le même résultat.

1. Soit P et Q deux polynômes de $\mathbb{C}[X]$.
   a. Montrer que P et Q admettent un diviseur commun si, et seulement si, la famille
   $$\{X^j P, X^k Q \text{ avec } j \leq q-1,\ k \leq p-1\}$$
   est liée.
   b. Montrer que P et Q admettent un diviseur commun si, et seulement si, leur résultant $\mathrm{Res}(P,Q)$ est nul.
2. Soit P et Q deux polynômes de $\mathbb{C}[X]$ et $z \in \mathbb{C}$.
   a. Montrer que $\mathrm{Res}(P,Q)$ est un déterminant à coefficients dans $\mathbb{C}$ dont tous les coefficients ne dépendent pas de z sauf ceux de la dernière ligne qui sont de la forme $z^i P(z)$ ou $z^j Q(z)$ avec $0 \leq i < q$ et $0 \leq j < p$, multiplié par une puissance de z.
   b. Montrer qu'il existe $(U,V) \in \mathbb{C}_{q-1}[X] \times \mathbb{C}_{p-1}[X]$ tel que
   $$\mathrm{Res}(P,Q) = U(z)P(z) + V(z)Q(z).$$
   c. Retrouver que des polynômes P et Q admettent un diviseur commun si, et seulement si, leur résultant $\mathrm{Res}(P,Q)$ est nul.

Partie B – Nombres algébriques

Définition 19.31.

Un complexe z est un nombre algébrique s'il est racine d'un polynôme non nul à coefficients rationnels.

PROBLÈMES

1. Soit x et y deux nombres algébriques, P et Q des polynômes non nuls à coefficients rationnels qui les annulent et $z = x + y$.
   a. Justifier que les polynômes $P(X)$ et $Q(z - X)$ ne sont pas premiers entre eux.
   b. En déduire que z est algébrique.
2. Montrer que le produit de deux nombres algébriques est algébrique.
3. Conclure que l'ensemble des nombres algébriques est un sous-anneau de $\mathbb{C}$.

Partie C – Discriminant

Définition 19.32.

Soit P un polynôme non-constant. On appelle discriminant de P la quantité

$$\Delta(P) = (-1)^{\frac{p(p-1)}{2}} a_p^{-1} \mathrm{Res}(P, P')$$

où p est le degré de P et a_p le coefficient dominant de P.

1. Déterminer l'expression du discriminant d'un polynôme de degré 2.
2. Montrer qu'un polynôme $P \in \mathbb{C}[X]$ est à racines simples si, et seulement si, $\Delta(P) \neq 0$.

CORRIGÉS

Corrigés des Vrai/Faux

a) Faux, c'est λ^n.

b) Vrai.

c) Vrai, c'est le déterminant de Vandermonde.

d) Vrai.

e) Faux.

f) Vrai.

g) Vrai.

h) Faux, par exemple $A = B = \mathrm{I}_n$ puisque $2^n \neq 2$ pour $n \geq 2$.

i) Vrai.

j) Vrai.

k) Vrai en considérant $\mathrm{diag}(\lambda, 1, \ldots, 1)$ pour $\lambda \in \mathbb{R}^*$.

l) Vrai.

m) Faux avec par exemple I_n et $2\mathrm{I}_n$.

n) Faux, $2\mathrm{I}_n$ est inversible de déterminant 2^n.

o) Vrai.

p) Vrai.

Corrigés des exercices d'application du chapitre 19

Exercice 1

Commençons par effectuer $C_k \leftarrow C_k - C_1$ pour tout $k \geq 2$. Dans le déterminant obtenu, les dernières colonnes sont proportionnelles entre elles (et proportionnelles à la colonne composée uniquement de coefficients égaux à 1) : le rang de la famille des colonnes est inférieur à $n-1$: ce déterminant est nul.

Exercice 2

L'objectif n'est pas de développer le déterminant puis de factoriser mais d'utiliser les opérations élémentaires afin d'obtenir directement une forme factorisée.

$$\begin{aligned}
\begin{vmatrix} 1 & a & b \\ a & 1 & b \\ b & a & 1 \end{vmatrix} &= \begin{vmatrix} 1+a+b & a & b \\ 1+a+b & 1 & b \\ 1+a+b & a & 1 \end{vmatrix} && C_1 \leftarrow C_1 + C_2 + C_3 \\
&= (1+a+b)\begin{vmatrix} 1 & a & b \\ 1 & 1 & b \\ 1 & a & 1 \end{vmatrix} \\
&= (1+a+b)\begin{vmatrix} 1 & 0 & 0 \\ 1 & 1-a & 0 \\ 1 & 0 & 1-b \end{vmatrix} && C_2 \leftarrow C_2 - aC_1,\ C_3 \leftarrow C_3 - bC_1 \\
&= (1+a+b)(1-a)(1-b).
\end{aligned}$$

Exercice 3

Commençons par effectuer $C_k \leftarrow C_k - C_{k-1}$ pour k allant de n à 2. Le déterminant obtenu est triangulaire inférieur avec sur la diagonale les coefficients 1, 2, …, n : son déterminant est le produit des éléments diagonaux donc $n!$.

Exercice 4

Commençons par développer cette matrice par rapport à la première ligne

$$\det M = 1 + \sum_{k=2}^{n} (-1)^{k+1} \det(X, E_1, \dots, E_{k-2}, E_k, \dots, E_{n-1}),$$

où E_1, …, E_{n-1} est la base canonique de $\mathscr{M}_{n-1}(\mathbb{R})$ et $X \in \mathscr{M}_{n-1}(\mathbb{R})$ est le vecteur colonne dont tous les coefficients sont égaux à 1. Or, par antisymétrie,

$$\begin{aligned}
\det(X, E_1, \dots, E_{k-2}, E_k, \dots, E_{n-1}) &= (-1)^{k-2} \det(E_1, \dots, E_{k-2}, X, E_k, \dots, E_{n-1}) \\
&= (-1)^{k-2} \det \mathrm{I}_{n-1} = (-1)^k.
\end{aligned}$$

Par conséquent, $\det M = 1 + \sum_{k=2}^{n} (-1) = 2 - n$.

Exercice 5

1. Commençons par effectuer $L_k \leftarrow L_k - L_{k-1}$ pour k allant de n à 2 :

$$\begin{vmatrix} 1 & \star & \dots & \star \\ 0 & & & \\ \vdots & & \left(\binom{i-1}{j-2}\right)_{2 \le i,j} & \\ 0 & & & \end{vmatrix}$$

car $\binom{i}{j-1} - \binom{i-1}{j-1} = \binom{i-1}{j-2}$. Développons ensuite par rapport à la première colonne. On trouve alors que le déterminant au rang n est égal au déterminant analogue au rang $n-1$. Par suite, le déterminant est égal de la même forme au rang 1, c'est-à-dire 1.

2. Commençons par effectuer $L_k \leftarrow L_k - L_{k-1}$ pour k allant de n à 2 puis développons par rapport à la première colonne. On obtient alors le déterminant

$$\det\left(\binom{i+j-2}{j-1} - \binom{i+j-3}{j-1}\right)_{2\leq i,j\leq n} = \det\left(\binom{i+j-3}{j-2}\right)_{2\leq i,j\leq n}.$$

Ensuite, effectuons $C_k \leftarrow C_k - C_{k-1}$ pour k allant de $n-1$ à 2. On obtient

$$\det\left(\binom{i+j-3}{j-2} - \binom{i+j-4}{j-3}\right)_{2\leq i,j\leq n} = \det\left(\binom{i+j-4}{j-2}\right)_{2\leq i,j\leq n}.$$

Ainsi, le déterminant au rang n est égal au déterminant analogue au rang $n-1$. Par suite, le déterminant est égal de la même forme au rang 1, c'est-à-dire 1.

Exercice 6

Avec l'opération élémentaire $C_1 \leftarrow C_1 + C_2 + \ldots + C_n$, on obtient

$$\det M' = (n-1)\det(C_1(M) + C_2(M) + \ldots + C_n(M), C_2(M'), \ldots, C_n(M')).$$

Effectuons ensuite $C_k \leftarrow C_k - C_1$ pour tout $k \geq 2$,

$$\det M' = (n-1)\det(C_1(M) + C_2(M) + \ldots + C_n(M), -C_2(M), \ldots, -C_n(M))$$

puis à nouveau l'opération élémentaire $C_1 \leftarrow C_1 + C_2 + \ldots + C_n$,

$$\begin{aligned}\det M' &= (n-1)\det(C_1(M), -C_2(M), \ldots, -C_n(M))\\ &= (n-1)(-1)^{n-1}\det M.\end{aligned}$$

Exercice 7

La famille de polynômes $P(X+1), \ldots, P(X+n)$ est liée car de cardinal n dans un espace de dimension $n-1$. Ainsi, il existe des scalaires $a_0, \ldots, a_{n-1}$ non tous nuls tels que

$$\sum_{k=1}^{n} a_k P(X+k) = 0.$$

En notant $C_1, \ldots, C_n$ les colonnes du déterminant précédent, $\sum_{k=1}^{n} a_k C_k = 0_{n,1}$, donc les colonnes sont liées et le déterminant est nul.

Exercice 8

Commençons par écrire la formule de Taylor :

$$\forall i, j \leq n, \qquad P(i+j) = \sum_{k=0}^{n-1} \frac{i^k}{k!} P^{(k)}(j).$$

Ainsi,

$$\det(P(i+j)) = \det\left(\frac{i^{j-1}}{(j-1)!}\right)\det(P^{(i-1)}(j)).$$

Mais d'une part

$$\det\left(\frac{i^{j-1}}{(j-1)!}\right) = \frac{1}{\prod_{k=0}^{n-1} k!} V(1, 2, \ldots, n) = 1$$

CORRIGÉS

et d'autre part, en utilisant le caractère n-linéaire alterné et en notant a le coefficient dominant de P,

$$\det(P^{(i-1)}(j)) = \det(a\frac{(n-1)!}{(n-i)!}j^{n-i}) = (a(n-1)!)^n \frac{1}{\prod_{k=0}^{n-1} k!}\det(j^{n-i})$$
$$= (a(n-1)!)^n \frac{1}{\prod_{k=0}^{n-1} k!}(-1)^{\frac{n(n-1)}{2}} V(1,2,\dots,n) = (-1)^{\frac{n(n-1)}{2}}(a(n-1)!)^n.$$

Le déterminant recherché est donc $(-1)^{\frac{n(n-1)}{2}}(a(n-1)!)^n$.

Exercice 9

Considérons une base de $\mathscr{M}_n(\mathbb{R})$ « adaptée » à l'endomorphisme u. Proposons pour cela de remplacer le premier vecteur de la base canonique, à savoir $E_{1,1}$, par la matrice I_n : on obtient encore toujours une base de $\mathscr{M}_n(\mathbb{R})$ et la matrice de u dans cette base est triangulaire supérieure de la forme

$$\begin{bmatrix} n & \star & \star & \dots & \star \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 0 & 1 \end{bmatrix}.$$

Les coefficients $\star$ sur la première ligne sont égaux à 1 ou 0 selon que le vecteur correspondant de la base canonique est ou non diagonal. Ainsi, le déterminant de u est le produit des éléments diagonaux, donc n.

Exercice 10

Déterminons la matrice de la dérivation dans la base formée des $x \mapsto x^k e^x$ pour $k \leq n$. Les coefficients de cette matrice sont $(m_{i,j})_{i,j}$ avec

$$m_{i,j} = \begin{cases} 1 & \text{si } i = j, \\ j & \text{si } j = i+1, \\ 0 & \text{sinon.} \end{cases}$$

Cette matrice est triangulaire et son déterminant est le produit des éléments diagonaux donc 1.

Exercice 11

Avec les opérations élémentaires $L_k \leftarrow L_k + L_1$ pour $k \in [\![1,n]\!]$, on obtient que $\det M$ est aussi le déterminant d'une matrice à coefficients entiers dont les $n-1$ dernières lignes ont des coefficients parmi $\{-2,0,2\}$ donc multiples de 2. En mettant en facteur 2 dans ces $n-1$ dernières lignes, on obtient que 2^{n-1} divise $\det M$.

Exercice 12

▹ Remarquons tout d'abord que le déterminant extrait avec les $n-1$ premières lignes et colonnes vaut 1 (déterminant triangulaire) donc $\operatorname{rg} M_a \geq n-1$.

▹ Par ailleurs, le déterminant de M_a vaut $1+(-1)^{n+1}a^n$ (en développant par rapport à la première colonne). La matrice M_a est donc inversible si, et seulement si, $(-a)^n \neq 1$ c'est-à-dire $-a \notin \mathbb{U}_n$.

En conclusion, si a est l'opposé d'une racine $n^{\text{ième}}$ de l'unité, alors $\operatorname{rg} M_a = n-1$; sinon, $\operatorname{rg} M_a = n$.

Exercice 13

D'après la caractérisation du rang donnée dans ce chapitre, il suffit de montrer que tous les déterminants extraits de taille 2 sont nuls.

Fixons $i \neq i'$ deux indices de lignes, $j \neq j'$ deux indices de colonnes. Considérons une permutation $\sigma \in S_n$ telle que $\sigma(i) = j$ et $\sigma(i') = j'$. Écrivons l'égalité entre C_σ et $C_{(j\ j')\circ\sigma}$ sous la forme $C_\sigma - C_{(j\ j')\circ\sigma} = 0$, c'est-à-dire

$$(m_{i,j}m_{i',j'} - m_{i,j'}m_{i',j}) \prod_{k \notin \{i,i'\}} m_{k,\sigma(k)} = 0.$$

Comme les quantités C_σ sont non nulles, les coefficients de M sont tous non nuls donc $m_{i,j}m_{i',j'} - m_{i,j'}m_{i',j} = 0$ ce qui est la nullité du déterminant extrait pour les lignes d'indices dans $\{i, i'\}$ et les colonnes d'indices dans $\{j, j'\}$.

En conclusion, $\operatorname{rg} M \leq 1$ et même $\operatorname{rg} M = 1$ car M est non nulle.

Exercice 14

Écrivons $A = PJ_rQ$ avec $P, Q \in \mathrm{GL}_n(\mathbb{R})$. Alors, pour tout $M \in \mathscr{M}_n(\mathbb{R})$,

$$\det(A+M) = \det(PJ_rQ+M) = \det P \det Q \det(J_r + P^{-1}MQ^{-1}),$$
$$\det A + \det M = \det P \det Q(\det J_r + \det P^{-1}MQ^{-1}),$$

donc, pour tout $N \in \mathscr{M}_n(\mathbb{R})$, $\det(J_r + N) = \det J_r + \det N$.

- Si $r = n$, le choix de la matrice $N = -\mathrm{I}_n$ amène une contradiction.
- Si $0 < r < n$, le choix de la matrice $N = \mathrm{I}_n - J_r$ amène une contradiction.

Par conséquent, $r = 0$ et $A = 0_n$.

Exercice 15

Notons r le rang de A et considérons les matrices $P \in \mathrm{GL}_p(\mathbb{K})$ et $Q \in \mathrm{GL}_q(\mathbb{K})$ telles que $A = PJ_rQ$; posons enfin, $B = Q^{-1}CP^{-1}$. Alors, d'une part,

$$\begin{aligned}\det(\mathrm{I}_p - AB) &= \det(P)\det(\mathrm{I}_p - J_rC)\det(P^{-1})\\ &= \begin{vmatrix} \mathrm{I}_r - C_1 & -C_2 \\ 0 & \mathrm{I}_{p-r,q-r} \end{vmatrix} = \det(\mathrm{I}_r - C_1),\end{aligned}$$

en décomposant par blocs $C = \begin{bmatrix} C_1 & C_2 \\ C_3 & C_4 \end{bmatrix}$, avec $C_1 \in \mathscr{M}_r(\mathbb{K})$ et $C_4 \in \mathscr{M}_{q-r,p-r}(\mathbb{K})$. D'autre part,

$$\begin{aligned}\det(\mathrm{I}_q - BA) &= \det(Q^{-1})\det(\mathrm{I}_q - CJ_r)\det(Q)\\ &= \begin{vmatrix} \mathrm{I}_r - C_1 & 0 \\ -C_3 & \mathrm{I}_{q-r,p-r} \end{vmatrix} = \det(\mathrm{I}_r - C_1).\end{aligned}$$

En conclusion, $\det(\mathrm{I}_q - BA) = \det(\mathrm{I}_p - AB)$.

Exercice 16

1. Supposons qu'une telle matrice P existe. Alors, $z \mapsto \det(P - zM) = \det(M)\det(M^{-1}P - z\mathrm{I}_n)$ ne s'annule pas. Or elle est polynomiale, non constante (et même de degré n avec la formule explicite du déterminant) : contradiction avec le théorème de D'Alembert-Gauss.
2. Supposons M de rang $r < n$. Comme M est équivalente à toute matrice de même rang, il existe U, $V \in \mathrm{GL}_n(\mathbb{R})$ telles que

$$M = U \begin{bmatrix} 0_{r,n-r} & \mathrm{I}_r \\ 0_{n-r} & 0_{n-r,r} \end{bmatrix} V.$$

Posons $P = UV \in \mathrm{GL}_n(\mathbb{R})$. Alors, pour tout $z \in \mathbb{C}$,

$$\det(P - zM) = \det U \begin{vmatrix} 1 & & z & & \\ & \ddots & & \ddots & \\ & & \ddots & & z \\ & & & \ddots & \\ & & & & 1 \end{vmatrix} \det V = \det U . \det V \neq 0$$

Exercice 17

Considérons les deux polynômes de $\mathbb{R}_3[X]$ définis par les fonctions polynomiales $P : x \mapsto \det(A + xB)$ et $Q : x \mapsto \det(C + xA)$. Comme $\det B = \det A$, ces deux polynômes ont le même degré et le même coefficient dominant. Par conséquent, $P - Q$ est de degré au plus 2 est admet 0 et i pour racine mais aussi $-i$ car il est à coefficients réels. Ainsi, $P - Q = 0$ et donc en particulier

$$\det(A + B) = P(1) = Q(1) = \det(C + A).$$

Exercice 18

Notons $U, V \in \mathscr{M}_n(\mathbb{R})$ telles que $U + iV \in \mathrm{GL}_n(\mathbb{C})$ et $A = (U + iV)B(U + iV)^{-1}$. Alors, $A(U + iV) = (U + iV)B$ et donc en identifiant les parties réelles et imaginaires de chaque coefficient, $AU = UB$ et $AV = VB$ et donc, pour tout $x \in \mathbb{R}$, $A(U + xV) = (U + xV)B$.

Introduisons la fonction polynomiale $P : x \mapsto \det(U + xV)$; comme $P(i) \neq 0$ (car $U + iV$ est inversible), P admet un nombre fini de racines et il existe donc un réel x tel que $P(x) \neq 0$. Alors, $U + xV \in \mathrm{GL}_n(\mathbb{R})$ et $A = (U + xV)B(U + xV)^{-1}$.

Corrigés des exercices d'approfondissement du chapitre 19

Exercice A

1. Notons $E_1, \ldots, E_{2n}$ la base canonique de $\mathscr{M}_{2n,1}(\mathbb{R})$ et $X \in \mathscr{M}_{2n,1}(\mathbb{R})$ la colonne dont tous les coefficients sont égaux à 1. Le déterminant recherché est donc

$$\det M = \det(X - E_1, X - E_2, \ldots, X - E_{2n}).$$

Utilisons alors la n-linéarité et le caractère alterné (tous les termes où apparaît la colonne X à au moins deux reprises sont nuls)

$$\det M = \det(-E_1, -E_2, \ldots, -E_{2n}) + \sum_{k=1}^{2n} \det(-E_1, -E_2, \ldots, \underset{\substack{\uparrow \\ k}}{X}, \ldots - E_{2n}).$$

Le premier terme vaut $(-1)^{2n} \det \mathrm{I}_{2n} = 1$. Pour tout $k \in [\![1, 2n]\!]$, l'opération $C_k \leftarrow \sum_{j=1}^{n} C_j$ donne

$$\det(-E_1, -E_2, \ldots, \underset{\substack{\uparrow \\ k}}{X}, \ldots - E_{2n}) = \det(-E_1, -E_2, \ldots, \underset{\substack{\uparrow \\ k}}{E_k}, \ldots - E_{2n}) = (-1)^{2n-1} \det \mathrm{I}_{2n} = -1.$$

En conclusion, le déterminant est $1 - 2n$.

2. Modulo 2, on a $-1 = 1[2]$ donc le déterminant d'une telle matrice est par la formule explicite égal modulo 2 au déterminant de la matrice M de la première question. Par conséquent, il est égal à $1[2]$ donc ne peut être nul : une telle matrice est inversible.

Exercice B

▹ Montrons tout d'abord qu'une matrice $M \in \mathscr{M}_2(\mathbb{Z})$ est inversible avec un inverse dans $\mathscr{M}_2(\mathbb{Z})$ si, et seulement si, $\det M \in \{-1, 1\}$.

• Si M admet un inverse dans $\mathscr{M}_2(\mathbb{Z})$, alors $\det(M)\det(M^{-1}) = 1$, $\det(M)$ et $\det(M^{-1})$ sont des entiers donc $\det(M) \in \{-1, 1\}$.

• Si $\det(M) \in \{-1, 1\}$, alors la formule de la comatrice donne que l'inverse de M appartient à $\mathscr{M}_2(\mathbb{Z})$ (puisque les cofacteurs de M sont des entiers).

▹ L'application polynomiale $P : x \longmapsto \det(A + xB)$ est de degré au plus 2 ; par hypothèse, $P(k) \in \{-1, 1\}$ pour tout $k \in \{0, \dots, 4\}$ donc P prend au moins trois fois une valeur $\epsilon \in \{-1, 1\}$. Comme P est de degré au plus 2, $P = \epsilon$. En particulier, $P(5) = \epsilon \in \{-1, 1\}$ donc $A + 5B$ est inversible avec un inverse dans $\mathscr{M}_2(\mathbb{Z})$.

Exercice C

Notons $C'_k = \mathrm{mat}(m_{i,k})_{i\geq 2}$ et $D_k = \mathrm{mat}(d_{i,k})_{i\geq 2}$.

$$\det M = \begin{vmatrix} m_{1,1} & m_{1,2} & \cdots & m_{1,n} \\ C'_1 & C'_2 & \cdots & C'_n \end{vmatrix} = m_{1,1} \begin{vmatrix} 1 & \frac{m_{1,2}}{m_{1,1}} & \cdots & \frac{m_{1,n}}{m_{1,1}} \\ C'_1 & C'_2 & \cdots & C'_n \end{vmatrix}$$

$$= \frac{1}{m_{1,1}^{n-2}} \begin{vmatrix} 1 & m_{1,2} & \cdots & m_{1,n} \\ C'_1 & m_{1,1}C'_2 & \cdots & m_{1,1}C'_n \end{vmatrix}$$

$$= \frac{1}{m_{1,1}^{n-2}} \begin{vmatrix} 1 & 0 & \cdots & 0 \\ C'_1 & m_{1,1}C'_2 - m_{1,2}C'_1 & \cdots & m_{1,1}C'_n - m_{1,n}C'_1 \end{vmatrix} \qquad C_k \leftarrow C_k - m_{1,k}C_1$$

$$= \frac{1}{m_{1,1}^{n-2}} \begin{vmatrix} 1 & 0 & \cdots & 0 \\ C'_1 & D_2 & \cdots & D_n \end{vmatrix} = \frac{1}{m_{1,1}^{n-2}} \det(D_2, \dots, D_n).$$

Exercice D

1. Le coefficient de la matrice AU en position (i, j) est

$$\sum_{k=1}^{n} a_{i,k} u_{k,j} = \sum_{k=1}^{n-i+1} b_{k+i-1}\omega^{(k-1)(j-1)} + \sum_{k=n-i+2}^{n} b_{k+i-1-n}\omega^{(k-1)(j-1)}$$
$$= \sum_{l=i}^{n} b_l \omega^{(l-i)(j-1)} + \sum_{l=1}^{i-1} b_l \omega^{(l-i)(j-1)}$$
$$= \left(\sum_{l=1}^{n} b_l \omega^{(l-1)(j-1)}\right)\omega^{(j-1)(1-i)} = \omega^{(j-1)(1-i)}\varphi(\omega^{j-1}).$$

2. On obtient $\det(AU) = \det(A)\det(U)$. Or $\det(AU) = \left(\prod_{k=1}^{n} \varphi(\omega^{k-1})\right)\det(\omega^{(j-1)(1-i)})_{i,j}$. Remarquons que $\det(\omega^{(j-1)(1-i)})_{i,j}$ et $\det(U)$ ne diffèrent que par l'ordre des lignes (la première ligne est identique dans

CORRIGÉS

les deux déterminants, La $i^{\text{ième}}$ ligne (pour $i \geq 2$) est la $(n-i+2)^{\text{ième}}$ ligne de l'autre). La permutation associée à ce changement d'ordre des lignes est la composée des transpositions $(i, n-i+2)$ donc on obtient facilement sa signature. En conclusion,

$$\det(A) = \epsilon \prod_{k=1}^{n} \varphi(\omega^{k-1}) \text{ avec } \epsilon = (-1)^{\lfloor \frac{n-1}{2} \rfloor} = \begin{cases} (-1)^{\frac{n}{2}-1} & \text{si } n \text{ pair,} \\ (-1)^{\frac{n-1}{2}} & \text{si } n \text{ impair.} \end{cases}$$

3. a. Si $b_k = a^{k-1}$, alors $\varphi(z) = \frac{1-(az)^n}{1-az}$. Par conséquent,

$$\det(A) = \epsilon \prod_{j=0}^{n-1} \frac{1-(a\omega^j)^n}{1-a\omega^j} = \epsilon(1-a^n)^{n-1}.$$

b. Si $b_k = \binom{n-1}{k-1}$, alors $\varphi(z) = (1+z)^{n-1}$. Par conséquent,

$$\det(A) = \epsilon \prod_{j=0}^{n-1} (1+\omega^j)^{n-1} = \epsilon(1-(-1)^n)^{n-1} = \begin{cases} 0 & \text{si } n \text{ pair,} \\ (-1)^{\frac{n-1}{2}} 2^{n-1} & \text{sinon.} \end{cases}$$

Exercice E

1. a. Si la famille $(f_1, \dots, f_n)$ est liée, alors l'une des fonctions peut s'exprimer comme combinaison linéaire des autres donc une ligne du déterminant $\mathscr{C}_{f_1,\dots,f_n}$ s'exprime comme combinaison linéaire des autres ; par conséquent $\mathscr{C}_{f_1,\dots,f_n}$ est l'application nulle.
b. Soit $(x_1, \dots, x_{n-1})$ tel que $\mathscr{C}_{f_1,\dots,f_{n-1}}(x_1, \dots, x_{n-1}) \neq 0$. En développant le déterminant $\mathscr{C}_{f_1,\dots,f_n}$ par rapport à la dernière colonne, on obtient pour tout x_n, $f_n(x_n) = \sum_{i=1}^{n-1} a_i f_i(x_n)$ avec

$$a_i = (-1)^{i+n+1} \frac{\mathscr{C}_{f_1,\dots,f_{i-1},f_{i+1},\dots,f_n}(x_1, \dots, x_{n-1})}{\mathscr{C}_{f_1,\dots,f_{n-1}}(x_1, \dots, x_{n-1})}.$$

D'où $f_n \in \mathrm{Vect}(f_1, \dots, f_{n-1})$.
c. En considérant le plus petit déterminant $\mathscr{C}_{f_1,\dots,f_k}$ qui est nul et en utilisant la question précédente, on obtient que $(f_1, \dots, f_n)$ est une famille liée.
2. a. D'après la question précédente, $\mathscr{C}_{f_1,\dots,f_n}$ n'est pas l'application nulle car la famille $(f_1, \dots, f_n)$ est libre. Il existe donc $(x_1, \dots, x_n)$ tel que $\mathscr{C}_{f_1,\dots,f_n}(x_1, \dots, x_n) \neq 0$.
b. En développant par rapport à la $k^{\text{ième}}$ colonne, on trouve que $\widetilde{f_k}$ s'exprime dans la base $(f_1, \dots, f_n)$ et que les coordonnées sont les cofacteurs de la matrice $(f_i(x_j))_{i,j}$ associés à la $k^{\text{ième}}$ colonne.
La matrice de ces coordonnées est inversibles (car $(f_i(x_j))_{i,j}$, et donc sa comatrice, est inversible). Par conséquent, les fonctions $f_1, \dots, f_n$ s'expriment comme combinaisons linéaires des fonctions $(\widetilde{f_1}, \dots, \widetilde{f_n})$. On en déduit que $(\widetilde{f_1}, \dots, \widetilde{f_n})$ est une base de F.
c. • Si $(g_1, \dots, g_n)$ est une base de F, notons P la matrice de passage entre les deux bases. La multiplicativité du déterminant entraîne que $\mathscr{C}_{g_1,\dots,g_n} = \det(P).\mathscr{C}_{f_1,\dots,f_n}$.
• Si $\mathscr{C}_{g_1,\dots,g_n}$ est une application non nulle, colinéaire à $\mathscr{C}_{f_1,\dots,f_n}$, alors les applications $\widetilde{f_1}, \dots, \widetilde{f_n}$ et $\widetilde{g_1}, \dots, \widetilde{g_n}$ engendrent le même espace donc la famille libre $(g_1, \dots, g_n)$ est une base de F.

Exercice F

1. L'application per est clairement n-linéaire. Elle est symétrique car, pour tout $\rho \in S_n$, $\sigma \mapsto \rho \circ \sigma$ est une bijection de S_n dans S_n.
2. Procédons par récurrence forte sur n.
• Si $A \in \mathscr{M}_1(\mathbb{R})$ vérifie $a_{1,1} = 0$, alors $I = J = \{1\}$ convient.

• Soit $n \in \mathbb{N}^*$. Supposons le résultat établi pour les matrices de taille inférieure ou égale à n.
Soit $A \in \mathscr{M}_{n+1}(\mathbb{R})$ telle que $\prod_{j=1}^{n+1} a_{\sigma(j),j} = 0$ pour tout $\sigma \in S_{n+1}$.
• Si la dernière colonne de A est nulle, alors $I = [\![1, n+1]\!]$, $J = \{n+1\}$ convient.
• Sinon, il existe $i_0 \in [\![1, n+1]\!]$ tel que $a_{i_0,n+1} \neq 0$. En considérant les permutations $\sigma \in S_{n+1}$ telles que $\sigma(n+1) = i_0$, on constate que la matrice extraite $(a_{i,j})_{i \neq i_0, j \leq n}$ vérifie l'hypothèse donc il existe des parties $I \subset [\![1, n+1]\!] \setminus \{i_0\}$ et $J \subset [\![1, n]\!]$ telles que $\operatorname{card} I + \operatorname{card} J = n+1$ et la matrice $(a_{i,j})_{i \in I, j \in J}$ est nulle. La condition sur les cardinaux implique que $\operatorname{card} I = \operatorname{card} J^c$. Considérons les permutations σ qui réalisent une bijection entre J^c et I (et donc entre J et I^c). On trouve alors que l'une des deux matrices $(a_{i,j})_{i \in I, j \in J^c}$ ou $(a_{i,j})_{i \in I^c, j \in J}$ vérifie l'hypothèse : supposons que cela soit $(a_{i,j})_{i \in I^c, j \in J}$ (qui est de taille $\operatorname{card} J \leq n$). D'après l'hypothèse de récurrence, il existe $J' \subset J$ et $I' \subset I^c$ tel que $\operatorname{card} I' + \operatorname{card} J' = \operatorname{card} J + 1$ et $(a_{i,j})_{i \in I', j \in J'}$ est nulle. Considérons la matrice extraite $(a_{i,j})_{i \in I \cup I', j \in J'}$: elle est nulle par construction et

$$\begin{aligned}\operatorname{card}(I \cup I') + \operatorname{card} J' &= \operatorname{card} I + \operatorname{card} I' + \operatorname{card} J' \\ &= \operatorname{card} I + \operatorname{card} J + 1 = n+2.\end{aligned}$$

D'où $\mathscr{P}_{n+1}$.
3. • Pour l'implication « directe », remarquons que $\operatorname{card} G' = \operatorname{card} \varphi(G')$ et que $\varphi(G') \subset \bigcup_{k \in G'} \Phi(k)$. Le résultat suit facilement.
• Pour le sens retour, considérons la matrice $A \in M_n(\mathbb{R})$ définie par les coefficients suivants :

$$a_{i,j} = \begin{cases} 1 & \text{si } i > p \\ 1 & \text{si } i \leq p \text{ et } j \in \Phi(i) \\ 0 & \text{sinon} \end{cases}$$

Montrons que $\operatorname{per}(A) \neq 0$ en utilisant la question précédente. Si $\operatorname{per}(A) = 0$, alors il existe $I, J \subset [\![1, n]\!]$ tels que

$$\operatorname{card} I + \operatorname{card} J = n+1$$

et $(a_{i,j})_{i \in I, j \in J}$ est nulle.
Par construction de la matrice A, cette hypothèse de nullité implique que $\left(\bigcup_{i \in I} \Phi(i)\right) \subset F \setminus J$. En passant aux cardinaux, on obtient :

$$\operatorname{card} I \leq \operatorname{card}\left(\bigcup_{i \in I} \Phi(i)\right) \leq \operatorname{card}(F \setminus J) = n - \operatorname{card} J = \operatorname{card} I - 1$$

D'où la contradiction.
Par conséquent, $\operatorname{per}(A) \neq 0$ et il existe $\sigma \in S_n$ tel que $\prod_{i=1}^{n} a_{\sigma(i),i} \neq 0$. Par construction, pour tout $i \in [\![1, m]\!]$, $\sigma^{-1} \in \Phi(i)$. En conclusion, l'application

$$\varphi : \begin{cases} G & \to & F \\ i & \mapsto & \sigma^{-1}(i) \end{cases}$$

vérifie les conditions recherchées.

Remarque

Le résultat de cette dernière question est appelé théorème de Hall ou « lemme des mariages ». Il admet l'interprétation suivante : soit F un ensemble de filles, G un ensemble de garçons que l'on munit d'une application Φ indiquant les préférences de chaque fille (c'est-à-dire, $\Phi(f)$ est l'ensemble des garçons acceptables par la fille f). La seconde condition est nécessaire et suffisante pour qu'il existe une attribution à chaque fille d'un garçon qui lui convient sans qu'un garçon soit attribué simultanément à deux filles.

Corrigés des problèmes du chapitre 19

Problème 1

Partie A – Fonction de Mobius

1. a. Raisonnons par récurrence avec la propriété $\mathscr{P}_n$: il existe une unique fonction $\mu : [\![1,n]\!] \to \mathbb{C}$ telle que $\mu(1)=1$ et

$$\forall k \in [\![2,n]\!], \qquad \sum_{d|k} \mu(d) = 0.$$

• La propriété est vraie au rang 1 d'après la condition $\mu(1)=1$.
• Soit $n \in \mathbb{N}^*$ telle que la propriété $\mathscr{P}_n$ soit vérifiée. D'après l'hypothèse de récurrence, les valeurs sur $[\![1,n]\!]$ d'une fonction $\mu : [\![1,n+1]\!] \to \mathbb{C}$ vérifiant la propriété sont uniquement déterminées. Or, pour avoir $\sum_{d|n+1} \mu(d) = 0$, il faut définir $\mu(n+1)$ comme

$$-\sum_{\substack{d|n+1\\ d\neq n+1}} \mu(d).$$

Ainsi, μ est uniquement définie sur $[\![1,n+1]\!]$.
b. Soit $n \in \mathbb{N}^*$. Alors,

$$\sum_{d|n} \mu(d) F(\frac{n}{d}) = \sum_{d|n} \mu(d) \sum_{d'|\frac{n}{d}} f(d') = \sum_{d'|n} f(d') \sum_{d|\frac{n}{d'}} \mu(d) = f(n),$$

car $\sum_{d|\frac{n}{d'}} \mu(d) = 0$ si $n \neq d'$.
2. a. ▹ Soit p un nombre premier. Alors, $\qquad 0 = \sum_{d|p} \mu(d) = \mu(1) + \mu(p)$.
D'où $\mu(p) = -1$.
▹ Pour tout $k \geq 2$,

$$\mu(p^k) = \sum_{d|p^k} \mu(d) - \sum_{d|p^{k-1}} \mu(d) = 0.$$

b. Montrons par récurrence sur $n \in \mathbb{N}^*$ la propriété $\mathscr{P}_n$: pour tout $(a,b) \in (\mathbb{N}^*)^2$ tel que $ab \leq n$ et $a \wedge b = 1$, $\mu(ab) = \mu(a)\mu(b)$.
• $\mathscr{P}_1$ est vérifiée car $\mu(1.1) = 1 = \mu(1)\mu(1)$.
• Soit $n \in \mathbb{N}^*$ tel que $\mathscr{P}_n$ est vérifiée et $(a,b) \in (\mathbb{N}^*)^2$ tel que $ab = n+1$ et $a \wedge b = 1$. Écrivons alors la propriété définissant μ aux rangs a, b et ab. On obtient grâce à l'hypothèse de récurrence

$$\begin{aligned}
\mu(ab) &= -\sum_{\substack{d|ab\\ d\neq ab}} \mu(d) \\
&= -\sum_{\substack{d|a, d'|b\\ dd'\neq ab}} \mu(dd') \\
&= -\sum_{\substack{d|a, d'|b\\ dd'\neq ab}} \mu(d)\mu(d') \\
&= -\mu(a) \sum_{\substack{d'|b\\ d'\neq b}} \mu(d') - \mu(b) \sum_{\substack{d|a\\ d\neq a}} \mu(d) - \sum_{\substack{d'|b\\ d'\neq b}} \mu(d') \sum_{\substack{d|a\\ d\neq a}} \mu(d) \\
&= \mu(a)\mu(b) - \left(\mu(a) + \sum_{\substack{d|a\\ d\neq a}} \mu(d)\right)\left(\mu(b) + \sum_{\substack{d'|b\\ d'\neq b}} \mu(d')\right) \\
&= \mu(a)\mu(b).
\end{aligned}$$

c. Écrivons la décomposition en facteurs premiers de $n = \prod_{i=1}^{k} p_i^{\alpha_i}$. D'après les questions précédentes,

$$\mu(n) = \prod_{i=1}^{k} \mu(p_i^{\alpha_i})$$

est nulle sur l'un des α_i est supérieur ou égale à 2 et vaut $(-1)^k$ si tous les α_i sont égaux à 1. L'ensemble des entiers $n \in \mathbb{N}^*$ tels que $\mu(n) = 1$ est donc composé des entiers sans facteur carré qui admette un nombre pair de diviseurs premiers.

Partie B – Calcul de déterminants

1. a. La matrice P est triangulaire avec des 1 sur la diagonale donc son déterminant est égal à 1.
b. Le coefficient en position i, j de la matrice $P^{\mathrm{T}} \operatorname{diag}(f(1), \ldots, f(n))P$ est $\sum_{d|i\wedge j} f(d)$. Pour que cette matrice soit effectivement A, il faut que ce terme soit égal à $F(i \wedge j)$ pour tout (i, j). D'après la formule d'inversion de Mobius, il suffit de prendre $f(n) = \sum_{d|n} \mu(d) F(\frac{n}{d})$ pour tout $n \in \mathbb{N}^*$.
c. Par la multiplicativité du déterminant,

$$\det A_F = \det P^{\mathrm{T}}.\det P.\prod_{k=1}^{n} f(k) = \prod_{k=1}^{n} f(k).$$

2. Pour $F = \mathrm{Id}_{\mathbb{N}^*}$, nous avons établi que la fonction correspondante est $f = \varphi$. La formule précédente donne donc

$$\det(i \wedge j)_{i,j} = \prod_{k=1}^{n} \varphi(k).$$

3. a. Soit $m, n \in \mathbb{N}^*$ premiers entre eux. Alors, $\mathscr{P}_{mn}$ est la réunion disjointe de $\mathscr{P}_m$ et $\mathscr{P}_n$. La relation $f(mn) = f(m)f(n)$ n'est alors que l'associativité du produit.
b. Notons D_a l'ensemble des diviseurs positifs d'un entier a. Soit $m, n \in \mathbb{N}^*$ premiers entre eux. L'application

$$\left\{ \begin{array}{ccc} D_m \times D_n & \to & D_{mn} \\ (a, b) & \mapsto & ab \end{array} \right.$$

est une bijection. Par conséquent,

$$\begin{aligned} F(mn) &= \sum_{d|mn} f(d) = \sum_{a|m} \sum_{b|n} f(ab) \\ &= \sum_{a|m} \sum_{b|n} f(a)f(b) = F(m)F(n). \end{aligned}$$

c. D'après ce qui précède, il suffit de vérifier que, pour tout nombre premier p et tout entier k, $F(p^k) = \frac{1}{p^k}$. Procédons par récurrence sur k.
- $F(p^0) = f(1) = 1$
- Soit $k \in \mathbb{N}$ tel que $F(p^k) = \frac{1}{p^k}$. Alors

$$F(p^{k+1}) = F(p^k) + f(p^{k+1}) = \frac{1}{p^k} + \frac{1}{p^{k+1}}(1-p) = \frac{1}{p^{k+1}}.$$

d. D'après la première question,

$$\det\left(\frac{1}{i \wedge j}\right)_{i,j} = \prod_{k=1}^{n} f(k) = \frac{1}{n!} \prod_{k=1}^{n} \prod_{p \in \mathscr{P}_k} (1-p).$$

Comme pour tous entiers $a, b \in \mathbb{N}^*$, $ab = (a \wedge b)(a \vee b)$, on obtient la relation matricielle

$$(i \vee j)_{i,j} = \operatorname{diag}(1, 2, \ldots, n)\left(\frac{1}{i \wedge j}\right)_{i,j} \operatorname{diag}(1, 2, \ldots, n).$$

En conclusion, « second déterminant de Smith » est

$$n!\prod_{k=1}^{n}\prod_{p\in\mathscr{P}_k}(1-p).$$

4. Notons $F(n)$ la somme des diviseurs de n, c'est-à-dire $F(n)=\sum_{d|n} d$. Le déterminant recherché est $\det(F(i\wedge j))_{i,j}$ donc il suffit de remarquer que la fonction f associée à F par inversion de Mobius est l'identité. Par conséquent, le déterminant recherché est $n!$.

Problème 2

Partie A – Formule de condensation de Desnanot-Jacobi

1. Pour toute matrice $M\in\mathscr{M}_n(\mathbb{K})$, $M\operatorname{Com}(M)^{\mathrm{T}}=\operatorname{Com}(M)^{\mathrm{T}}M=\det M\,\mathrm{I}_n$.
2. a. En développant le déterminant de $M_{i,j}$ successivement par rapport à toutes les colonnes sauf celles d'indice i et j, on obtient

$$\det M_{i,j}=\begin{vmatrix}\Delta_i^i & (-1)^{i+j}\Delta_j^i\\ (-1)^{i+j}\Delta_i^j & \Delta_j^j\end{vmatrix}=\Delta_i^i\Delta_j^j-\Delta_i^j\Delta_j^i.$$

b. D'après la formule du produit matriciel, le coefficient en position i_0, j_0 de $MM_{i,j}$ est

$$\begin{cases} m_{i_0,j_0} & \text{si } j\notin\{i,j\},\\ \displaystyle\sum_{k=1}^{n} m_{i_0,k}(-1)^{k+j_0}\Delta_{j_0}^k & \text{sinon.}\end{cases}$$

En utilisant la formule de la comatrice rappelée précédemment, on vérifie

$$\sum_{k=1}^{n} m_{i_0,k}(-1)^{k+j_0}\Delta_{j_0}^k=\begin{cases}\det M & \text{si } i_0\neq j_0,\\ 0 & \text{sinon.}\end{cases}$$

c. En développant le déterminant par rapport aux colonnes d'indices i et j, on obtient

$$\det(MM_{i,j})=(\det M)^2\Delta_{i,j}^{i,j}.$$

d. En combinant les deux résultats, on en déduit

$$(\det M)^2\Delta_{i,j}^{i,j}=\det M(\Delta_i^i\Delta_j^j-\Delta_i^j\Delta_j^i).$$

Comme M est inversible, on simplifie par $\det M\neq 0$ et on obtient la formule recherchée.
3. Il existe un nombre fini de réels x tel que $M+x\mathrm{I}_n$ est non inversible. Pour tout x tel que $M+x\mathrm{I}_n$ est inversible, la formule est établie. Les fonctions qui a x associent le déterminant ou l'un des mineurs de $M+x\mathrm{I}_n$ sont continues car polynomiales. Faisons alors tendre x vers 0 (ce qui est possible car seul un nombre fini de valeurs rend $M+x\mathrm{I}_n$ non inversible) : d'où le résultat.

Partie B – Déterminants de Hankel-Sylvester

Remarquons que le $0^{\text{ième}}$ *déterminant de Hankel de* f *est* f.
1. Pour $f=\exp$ et $n\geq 1$, le $n^{\text{ième}}$ déterminant de Hankel est nul car les deux premières colonnes sont égales.
2. Pour tout (i,j),

$$f^{(i+j-2)}(x)=\sum_{k=1}^{n+1}\left(a_k\alpha_k^{i-1}e^{\alpha_k x}\right)\left(\alpha_k^{j-1}\right).$$

Par conséquent, le $n^{\text{ième}}$ déterminant de Sylvester est le produit des déterminants des matrices $(\alpha_i^{j-1})_{i,j\leq n+1}$ et $(a_j\alpha_j^{i-1}e^{\alpha_j x})_{i,j\leq n+1}$. Le premier est simplement le déterminant de Vandermonde de $\alpha_1,\dots,\alpha_{n+1}$; pour

le second, on met en facteur pour chaque colonne le coefficient $a_j e^{\alpha_j x}$ et on retrouve le même déterminant de Vandermonde. En conclusion, pour tout $x \in \mathbb{R}$,

$$S_n f(x) = V(\alpha_1, \dots, \alpha_{n+1})^2 e^{x \sum_{k=1}^{n+1} \alpha_k} \prod_{k=1}^{n+1} a_k.$$

3. a. Soit $C_1, \dots, C_{n+1}$ les colonnes du déterminant de $S_n f$. Pour tout $x \in I$,

$$(S_n f)'(x) = \sum_{k=1}^{n+1} \det(C_1, \dots, C_{k-1}, C_k', C_{k+1}, \dots, C_{n+1}).$$

Comme $C_k' = C_{k+1}$, on obtient avec le caractère alterné du déterminant :

$$\begin{aligned}(S_n f)'(x) &= \det(C_1, \dots, C_{n+1}') \\ &= \det(C_1, \dots, C_n, C_{n+2}) = \Delta_{n+2}^{n+1}.\end{aligned}$$

Comme M est symétrique, $\Delta_{n+1}^{n+2} = \Delta_{n+2}^{n+1}$.
Pour calculer $(S_n f)''$, on recommence les calculs en dérivant selon les lignes et on obtient que ce déterminant est le même que celui de $(S_n f)'$ en dérivant la dernière ligne : ce déterminant est Δ_{n+1}^{n+1}.
b. Remarquons que $S_{n+1}(f) = \det M$, $S_n(f) = \Delta_{n+2}^{n+2}$ et $S_{n-1} f = \Delta_{n+1,n+2}^{n+1,n+2}$. En appliquant la formule de Desnanot-Jacobi à la matrice M, on obtient que

$$S_{n+1} f . S_{n-1} f = S_n f . (S_n f)'' - (S_n f)'^2.$$

4. Le sens direct est évident car si f est solution d'une équation différentielle linéaire d'ordre n, homogène à coefficients constants, alors les colonnes du déterminant de $S_n f$ sont liées.
Montrons le sens retour par récurrence sur n.
• Si $S_1 f = 0$, alors $f f'' - f'^2 = 0$. Sur tout intervalle où f ne s'annule pas, cette condition est équivalente à $\left(\frac{f'}{f}\right)' = 0$ donc il existe α tel que $f' = \alpha f$.
• Soit $p \in \mathbb{N}^*$ tel que la propriété soit vraie jusqu'au rang p et f telle que $S_{p+1} f = 0$. D'après la question précédente, $S_1(S_p(f)) = 0$ donc il existe α et C tel que $S_p f = C e^{\alpha x}$.
– Si $S_p f = 0$, alors on utilise l'hypothèse de récurrence.
– Sinon, $S_p f$ ne s'annule pas. Ainsi, la dernière colonne du déterminant $S_{p+1} f$ s'exprime comme combinaison linéaire des autres ; il existe des fonctions $\lambda_0, \dots, \lambda_p$ telles que

$$\forall k \in [\![0, p+1]\!], \qquad f^{(k+p+1)} = \sum_{j=0}^{p} \lambda_j f^{(j+k)}.$$

En dérivant (les fonctions $\lambda_0, \dots, \lambda_p$ sont de classe $\mathscr{C}^1$ d'après les formules de Cramer), on obtient (en utilisant la relation pour $k+1$)

$$\forall k \in [\![0, p]\!], \qquad 0 = \sum_{j=0}^{p} \lambda_j' f^{(j+k)}.$$

Or, comme $S_p f \neq 0$, le système en $\lambda_0', \dots, \lambda_p'$ est de Cramer et donc

$$\lambda_0' = \dots = \lambda_p' = 0.$$

Les fonctions $\lambda_0, \dots, \lambda_p$ sont donc constantes et f est solution d'une équation différentielle linéaire d'ordre p, homogène à coefficients constants.

Problème 3

Partie A – Primalité relative

1. a. ($\Rightarrow$) Si $\deg P \wedge Q \geq 1$, alors il existe $\widetilde{P} \in \mathbb{C}_{p-1}[X]$ et $\widetilde{Q} \in \mathbb{C}_{q-1}[X]$ non tous les deux nuls tels que $P = \widetilde{P}.P \wedge Q$ et $Q = \widetilde{Q}.P \wedge Q$; ainsi

$$\widetilde{Q}P - \widetilde{P}Q = 0$$

et la famille de l'énoncé est liée.
($\Leftarrow$) S'il existe $A \in \mathbb{C}_{p-1}[X]$ et $B \in \mathbb{C}_{q-1}[X]$ tels que $AQ + BP = 0$ et $P \wedge Q = 1$, alors $P \mid AQ$ donc $P \mid A$ (lemme de Gauss) ce qui est impossible d'après les degrés. Donc $P \wedge Q \neq 1$.
b. $\mathrm{Res}(P, Q)$ est le déterminant dans la base canonique de la famille précédente. On conclut en rappelant que la famille est liée si, et seulement si, son déterminant est nul.
2. a. Après la transformation $L_{p+q} \leftarrow \sum_{k=1}^{p+q} z^{k-1} L_k$, la dernière ligne est alors

$$(P(z)\ zP(z)\ \dots\ z^{p-1}P(z)\ Q(z)\ zQ(z)\ \dots\ z^{q-1}Q(z)).$$

b. Il suffit de développer par rapport à la dernière ligne pour avoir l'expression recherchée.
c. ($\Rightarrow$) L'expression précédente nous indique qu'alors $\mathrm{Res}(P, Q)$ est à la fois une constante et un polynôme qui admet une racine donc il est nul.
($\Leftarrow$) On conclut comme précédemment.

Partie B – Nombres algébriques

1. a. Les polynômes $P(X)$ et $Q(z - X)$ admettent x comme racine commune donc ne sont pas premiers entre eux.
b. D'après la partie précédente, $\mathrm{Res}(P(X), Q(z - X)) = 0$; d'après les règles de calcul du déterminant (donc du résultant), la quantité du membre de gauche définit un polynôme en z à coefficients rationnels. Ainsi, z est algébrique.
2. Soit x et y algébriques de polynômes annulateurs P et Q et $z = xy$. On procède comme à la question précédente avec les polynômes $P(X)$ et $X^q Q(\frac{z}{X})$ où $q = \deg Q$.
3. L'ensemble des nombres algébriques contient 1 (et, de manière générale, tous les rationnels), est stable par passage à l'opposé (si $P(z) = 0$ alors $Q(-z) = 0$ avec $Q(X) = P(-X)$), par somme et produit (les questions précédentes) donc c'est un sous-anneau de $\mathbb{C}$. C'est en fait un sous-corps de $\mathbb{C}$.

Partie C – Discriminant

1. Si $P(X) = aX^2 + bX + c$ avec $a \neq 0$, alors

$$\Delta(P) = -\frac{1}{a}\begin{vmatrix} c & b & 0 \\ b & 2a & b \\ a & 0 & 2a \end{vmatrix} = b^2 - 4ac.$$

2. Un polynôme $P \in \mathbb{C}[X]$ est à racines simples si, et seulement si, P et P' n'ont pas de racine commune c'est-à-dire si, et seulement si, $\mathrm{Res}(P, P') \neq 0$ donc $\Delta(P) \neq 0$.

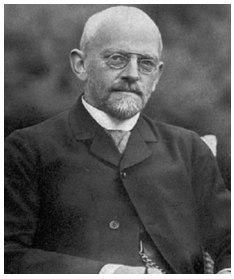
Chapitre 20

Espaces euclidiens

1. Produit scalaire — **2.** Orthogonalité — **3.** Isométries vectorielles

Objectifs et compétences du programme

- Revisiter les notions de géométrie euclidienne du plan.
- Effectuer des calculs en bases orthonormées.
- Calculer une projection orthogonale et calculer la distance à un sous-espace.
- Reprendre l'étude des isométries en petites dimensions.



David Hilbert, mathématicien allemand (1862-1943)
Il est considéré comme l'un des mathématiciens les plus influents du XXe siècle. Professeur à Göttingen, il a abordé de nombreux domaines mathématiques : on retient notamment son axiomatisation de la géométrie euclidienne, l'introduction des espaces de Hilbert pour l'analyse fonctionnelle (notamment pour l'étude des équations intégro-différentielles), son apport à la théorie algébrique des nombres ou encore sa promotion de la méthode axiomatique formelle...
Lors du Congrès International des Mathématiques de 1900 à Paris, Hilbert dresse, lors d'une conférence plénière restée célèbre, une liste de problèmes particulièrement féconds. Lors de la publication, cette liste comporte finalement 23 problèmes comme autant de défis pour le XXe siècle. Certains demeurent ouverts aujourd'hui.

COURS

Ce chapitre étend la notion d'orthogonalité aux $\mathbb{R}$-espaces vectoriels en définissant de manière générale la notion de produit scalaire. Une fois familiarisé avec cet outil, on peut s'intéresser aux vecteurs orthogonaux, aux bases orthonormées (aussi bien pour établir leur existence que pour les calculs qu'elles permettent). Des notions plus géométriques réapparaissent ensuite lors de l'étude des isométries vectorielles en petites dimensions.

1. Produit scalaire

1.1. Définitions

Définition 20.1.

Soit E un $\mathbb{R}$-espace vectoriel. Un produit scalaire sur E est une application

$$\left\{\begin{array}{ccc} E^2 & \longrightarrow & \mathbb{R} \\ (x,y) & \longmapsto & \langle x|y\rangle \end{array}\right.$$

- bilinéaire, c'est-à-dire
  - pour tout $y \in E$, l'application $x \longmapsto \langle x|y\rangle$ est une forme linéaire
  - pour tout $x \in E$, l'application $y \longmapsto \langle x|y\rangle$ est une forme linéaire
- symétrique
$$\forall x,\ y \in E, \qquad \langle x|y\rangle = \langle y|x\rangle$$
- positive
$$\forall x \in E, \qquad \langle x|x\rangle \geq 0$$
- définie positive
$$\forall x \in E \setminus \{0_E\}, \qquad \langle x|x\rangle > 0.$$

Un espace préhilbertien réel est un $\mathbb{R}$-espace vectoriel muni d'un produit scalaire ; un espace euclidien est un espace préhilbertien réel de dimension finie.
Le produit scalaire des vecteurs x et y peut aussi être noté $(x|y)$, $\langle x, y\rangle$ ou encore $x.y$.

Conseils méthodologiques

Pour établir qu'une application est bien un produit scalaire, il est astucieux de commencer par établir la symétrie afin de n'avoir que l'une des deux linéarités à écrire.

Définition 20.2.

La norme associée à un produit scalaire $\langle\ |\ \rangle$ est l'application

$$\left\{\begin{array}{ccc} E & \longrightarrow & \mathbb{R}_+ \\ x & \longmapsto & \|x\| = \sqrt{\langle x|x\rangle} \end{array}\right.$$

Remarque

Soit $\|\ \|$ la norme associée au produit scalaire $\langle\ |\ \rangle$ sur E. Alors, pour tout $x \in E$ et tout $\lambda \in \mathbb{R}$,

$$\|\lambda x\| = \sqrt{\langle \lambda x|\lambda x\rangle} = \sqrt{\lambda^2\langle x|x\rangle} = |\lambda|\|x\|.$$

Proposition 20.3. Identités remarquables

Soit $\langle\ |\ \rangle$ un produit scalaire sur E et $\|\ \|$ la norme associée. Alors, pour tous $x,\ y \in E$,

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2\langle x|y\rangle,$$
$$\|x-y\|^2 = \|x\|^2 + \|y\|^2 - 2\langle x|y\rangle,$$
$$\|x\|^2 - \|y\|^2 = \langle x+y|x-y\rangle.$$

Démonstration

▹ Montrons la première identité. Soit $x,\ y \in E$. Utilisons la bilinéarité du produit scalaire

$$\|x+y\|^2 = \langle x+y|x+y\rangle = \langle x|x\rangle + \langle x|y\rangle + \langle y|x\rangle + \langle y|y\rangle,$$

puis la définition de norme associée et la symétrie du produit scalaire

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2\langle x|y\rangle.$$

▹ La deuxième identité se déduit de la première en remplaçant y par $-y$.

▹ La dernière identité s'obtient comme la première en développant par bilinéarité la quantité $\langle x+y|x-y\rangle$ pour tous $x,\ y \in E$. ■

Remarque

Les identités remarquables sont souvent utilisées pour exprimer le produit scalaire en fonction de la norme associée. Elles sont alors appelées identités de polarisation ; pour tous $x,\ y \in E$,

$$\langle x|y\rangle = \frac{1}{2}(\|x+y\|^2 - \|x\|^2 - \|y\|^2),$$
$$\langle x|y\rangle = \frac{1}{2}(\|x\|^2 + \|y\|^2 - \|x-y\|^2),$$
$$\langle x|y\rangle = \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2).$$

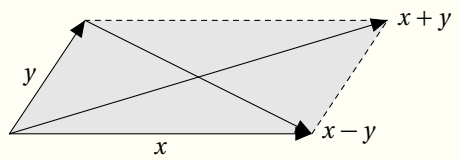
En ajoutant les deux premières identités remarquables, on obtient le résultat suivant.

Proposition 20.4. Identité du parallélogramme

Soit $\|\ \|$ la norme associée au produit scalaire $\langle\ |\ \rangle$ sur E. Alors, pour tous $x,\ y \in E$,

$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2).$$

L'appellation identité du parallélogramme provient de l'interprétation suivante en géométrie plane : la somme des carrés des longueurs des diagonales d'un parallélogramme est égale à la somme des carrés des longueurs des côtés.



1.2. Exemples

Donnons quelques exemples de produits scalaires usuels.

1.2.1. Sur $\mathbb{R}^n$

Proposition 20.5.

L'application qui, à deux vecteurs $(x_1,\dots,x_n)$ et $(y_1,\dots,y_n)\in\mathbb{R}^n$, associe le réel $\sum_{k=1}^n x_k y_k$ définit un produit scalaire sur $\mathbb{R}^n$, appelé produit scalaire canonique.

Démonstration

▹ L'application est clairement symétrique.

▹ Soit $(x_1,\dots,x_n)$, $(x'_1,\dots,x'_n)$, $(y_1,\dots,y_n)\in\mathbb{R}^n$ et $\lambda\in\mathbb{R}$. Alors,

$$\sum_{k=1}^n (x_k+\lambda x'_k)y_k = \sum_{k=1}^n x_k y_k + \lambda \sum_{k=1}^n x'_k y_k.$$

Par conséquent, l'application considérée est linéaire en son premier argument et par symétrie, bilinéaire.

▹ Soit $(x_1,\dots,x_n)\in\mathbb{R}^n$. Alors, $\sum_{k=1}^n x_k^2 \geq 0$, avec égalité si, et seulement si, tous ces carrés sont nuls, c'est-à-dire si $(x_1,\dots,x_n)=(0,\dots,0)$. L'application est définie positive. ■

Remarque

On peut en déduire d'autres produits scalaires en pondérant les différentes coordonnées avec des coefficients strictement positifs. Plus précisément, pour $a_1,\dots,a_n>0$, l'application qui à deux vecteurs $(x_1,\dots,x_n)$ et $(y_1,\dots,y_n)\in\mathbb{R}^n$ associe le réel $\sum_{k=1}^n a_k x_k y_k$ définit un produit scalaire sur $\mathbb{R}^n$.

1.2.2. Sur $\mathscr{M}_{n,1}(\mathbb{R})$

Proposition 20.6.

L'application qui, à deux matrices colonnes X et $Y\in\mathscr{M}_{n,1}(\mathbb{R})$, associe le réel $X^{\mathrm{T}}Y$ définit un produit scalaire sur $\mathscr{M}_{n,1}(\mathbb{R})$, appelé produit scalaire canonique.

Il s'agit en fait du « même » produit scalaire que précédemment en identifiant $\mathscr{M}_{n,1}(\mathbb{R})$ et $\mathbb{R}^n$. En effet, en notant $x_1,\dots,x_n$ les coefficients de X et $y_1,\dots,y_n$ les coefficients de Y, $X^{\mathrm{T}}Y=\sum_{k=1}^n x_k y_k$.

1.2.3. Sur $\mathscr{M}_n(\mathbb{R})$

Proposition 20.7.

L'application qui à deux matrices A et $B\in\mathscr{M}_n(\mathbb{R})$ associe le réel $\mathrm{tr}(A^{\mathrm{T}}B)$ définit un produit scalaire sur $\mathscr{M}_n(\mathbb{R})$, appelé produit scalaire canonique.

Démonstration

Notons $A=(a_{i,j})_{i,j}$ et $B=(b_{i,j})_{i,j}$. Alors,

$$\mathrm{tr}(A^{\mathrm{T}}B)=\sum_{i=1}^n [A^{\mathrm{T}}B]_{i,i} = \sum_{i=1}^n\sum_{k=1}^n a_{k,i}\,b_{k,i}.$$

On reconnaît encore une expression comme somme des produits de coordonnées et on conclut comme pour les produits scalaires précédents. ■

Proposition 20.8.

L'application qui, à deux polynômes réels $\sum_{k=0}^{+\infty} a_k X^k$ et $\sum_{k=0}^{+\infty} b_k X^k$, associe le réel $\sum_{k=0}^{+\infty} a_k b_k$ définit un produit scalaire sur $\mathbb{R}[X]$, appelé produit scalaire canonique.

Il s'agit une fois encore de considérer la somme des produits de coordonnées dans la base canonique.

Proposition 20.9.

Soit $a_0 < a_1 < \ldots < a_n$. L'application qui, à deux polynômes P et $Q \in \mathbb{R}_n[X]$, associe le réel $\sum_{k=0}^{n} P(a_k)Q(a_k)$ définit un produit scalaire sur $\mathbb{R}_n[X]$.

Démonstration

▹ L'application est clairement symétrique.

▹ Soit P, Q, $R \in \mathbb{R}_n[X]$ et $\lambda \in \mathbb{R}$. Alors,

$$\sum_{k=0}^{n}(P(a_k)+\lambda Q(a_k))R(a_k) = \sum_{k=0}^{n} P(a_k)R(a_k) + \lambda \sum_{k=0}^{n} Q(a_k)R(a_k).$$

Par conséquent, l'application considérée est linéaire en son premier argument et par symétrie, bilinéaire.

▹ Soit $P \in \mathbb{R}_n[X]$. Le réel $\sum_{k=0}^{n} P(a_k)^2$ est positif et n'est nul que si $P(a_0) = \ldots = P(a_n) = 0$ c'est-à-dire si $P = 0$ (car $\deg P \leq n$). L'application est donc définie positive. ■

1.2.4. Sur $\ell^2(\mathbb{R})$

Proposition 20.10.

L'ensemble $\ell^2(\mathbb{R})$ des suites $(u_n) \in \mathbb{R}^{\mathbb{N}}$ telles que la série de terme général u_n^2 converge est un $\mathbb{R}$-espace vectoriel et l'application $(u, v) \mapsto \sum_{k=0}^{+\infty} u_k v_k$ est un produit scalaire sur $\ell^2(\mathbb{R})$.

Démonstration

Pour montrer que $\ell^2(\mathbb{R})$ est stable par combinaison linéaire, il suffit essentiellement de montrer que si u, $v \in \ell^2(\mathbb{R})$, alors la série de terme général $u_n v_n$ converge. Pour cela, il suffit de remarquer avec l'inégalité arithmético-géométrique que

$$\forall n, \qquad |u_n v_n| \leq \frac{u_n^2 + v_n^2}{2}.$$

▹ Soit u, $v \in \ell^2(\mathbb{R})$ de $\lambda \in \mathbb{R}$. Alors, comme $(u_n + \lambda v_n)^2 = u_n^2 + 2\lambda u_n v_n + \lambda^2 v_n^2$, la suite $u + \lambda v \in \ell^2(\mathbb{R})$. Ainsi, la partie non vide $\ell^2(\mathbb{R})$ est un sous-espace de $\mathbb{R}^{\mathbb{N}}$.

▹ Ainsi, l'application est bien définie. Les propriétés de produit scalaire se vérifient immédiatement. ■

1.2.5. Sur $\mathscr{C}^0([a, b], \mathbb{R})$

Proposition 20.11.

L'application qui, à deux fonctions f et $g \in \mathscr{C}^0([a, b], \mathbb{R})$, associe le réel $\int_a^b f(t)g(t)\,\mathrm{d}t$ définit un produit scalaire sur $\mathscr{C}^0([a, b], \mathbb{R})$, appelé produit scalaire canonique.

Démonstration

▹ L'application est clairement symétrique.

▹ Soit $f, g, h \in \mathscr{C}^0([a,b],\mathbb{R})$ et $\lambda \in \mathbb{R}$. Alors, par linéarité de l'intégrale,

$$\int_a^b (f(t)+\lambda g(t))h(t)\,\mathrm{d}t = \int_a^b f(t)h(t)\,\mathrm{d}t + \lambda \int_a^b g(t)h(t)\,\mathrm{d}t\,.$$

Par conséquent, l'application considérée est linéaire en son premier argument et par symétrie, bilinéaire.

▹ Soit $f \in \mathscr{C}^0([a,b],\mathbb{R})$. Le réel $\int_a^b f(t)^2\,\mathrm{d}t$ est positif et n'est nul que si $f = 0$ (car il s'agit de l'intégrale nulle d'une fonction continue de signe constant). L'application est donc définie positive. ■

On peut généraliser ce produit scalaire en ajoutant une fonction de poids ω strictement positive dans l'intégrale.

1.3. Inégalité de Cauchy-Schwarz

Proposition 20.12. Inégalité de Cauchy-Schwarz

Soit $\langle\ |\ \rangle$ un produit scalaire sur E et $\|\ \|$ la norme associée. Alors, pour tous $x,\ y \in E$,

$$|\langle x|y\rangle| \leq \|x\|\|y\|.$$

Le cas d'égalité correspond à deux vecteurs colinéaires.

Démonstration

▹ Pour tous $x,\ y \in E$, et $\lambda \in \mathbb{K}$, $\|x+\lambda y\|^2 \geq 0$. En développant cette expression, on trouve que le polynôme (en la variable λ) $\|y\|^2\lambda^2 + 2\langle x|y\rangle\lambda + \|x\|^2$ est toujours positif.

- Si $\|y\|^2 = 0$, alors ce polynôme est de degré au plus 1 et pour être positif, il doit être égal à une constante positive d'où $\langle x|y\rangle = 0$ et l'inégalité est vérifiée.
- Sinon le polynôme est de degré 2 et son discriminant Δ doit être négatif. Or, ce discriminant vaut $4\langle x|y\rangle^2 - 4\|x\|^2\|y\|^2$, d'où l'inégalité.

▹ Soit $x,\ y \in E$ tels que $|\langle x|y\rangle| = \|x\|\|y\|$.

- Si $\|y\|^2 = 0$, alors $y = 0_E$ d'après le caractère défini positif du produit scalaire et y est colinéaire à x.
- Sinon, avec les notations précédentes, le discriminant Δ est nul. Le polynôme $\|y\|^2\lambda^2 + 2\langle x|y\rangle\lambda + \|x\|^2$ admet donc une unique racine λ_0 : $\|x+\lambda_0 y\| = 0$ et donc $x = -\lambda_0 y$.

■

Exemple

L'inégalité de Cauchy-Schwarz pour le produit canonique sur $\mathbb{R}^n$ indique que, pour tous vecteurs $(x_1,\dots,x_n)$ et $(y_1,\dots,y_n) \in \mathbb{R}^n$,

$$\left|\sum_{k=1}^n x_k y_k\right| \leq \sqrt{\sum_{k=1}^n x_k^2 \sum_{l=1}^n y_l^2}.$$

Exemple

L'inégalité de Cauchy-Schwarz pour le produit canonique sur $\mathscr{C}^0([a,b],\mathbb{R})$ indique que, pour tous

f et $g \in \mathscr{C}^0([a,b],\mathbb{R})$,

$$\left|\int_a^b f(t)g(t)\,\mathrm{d}t\right| \leq \sqrt{\int_a^b f(t)^2\,\mathrm{d}t \int_a^b g(t)^2\,\mathrm{d}t}.$$

Proposition 20.13. Inégalité triangulaire pour la norme

Soit $\|\ \|$ la norme associée au produit scalaire $\langle\ |\ \rangle$ sur E. Alors, pour tous $x,\ y \in E$,

$$\|x+y\| \leq \|x\| + \|y\|.$$

Démonstration

Soit $x,\ y \in E$. Alors,

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2\langle x|y\rangle \leq \|x\|^2 + \|y\|^2 + 2\|x\|\|y\| = (\|x\| + \|y\|)^2,$$

d'après l'inégalité de Cauchy-Schwarz. La démonstration se termine en appliquant la fonction $x \mapsto \sqrt{x}$ qui est croissante. ■

Comme pour la valeur absolue ou le module, on obtient immédiatement la seconde inégalité triangulaire.

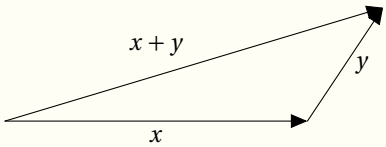
Corollaire 20.14. Inégalité triangulaire pour la norme

Soit $\|\ \|$ la norme associée au produit scalaire $\langle\ |\ \rangle$ sur E. Alors, pour tous $x,\ y \in E$,

$$\big|\|x\| - \|y\|\big| \leq \|x-y\|.$$

Remarque

L'appellation inégalité triangulaire se comprend aisément avec un dessin dans $\mathbb{R}^2$. La « longueur » du vecteur $x+y$ est inférieure à la somme des « longueurs » des vecteurs x et y.



1.4. Forme linéaire

Remarquons que, par définition du produit scalaire, l'application $\varphi_a : x \mapsto \langle x|a\rangle$ est une forme linéaire pour tout $a \in E$. La proposition suivante indique que cette constatation est générique sur un espace euclidien.

Proposition 20.15.

Soit E un espace euclidien. L'application

$$\Phi : \left\{ \begin{array}{ccc} E & \to & E^* \\ a & \mapsto & (x \mapsto \langle a|x \rangle) \end{array} \right.$$

est un isomorphisme.
En d'autres termes, pour toute forme linéaire φ sur E, il existe un unique $a \in E$ tel que

$$\forall x \in E, \qquad \varphi(x) = \langle a|x \rangle.$$

Démonstration
Remarquons tout d'abord que E et E^* sont de même dimension. Il suffit donc de démontrer que Φ est injective. Pour cela considérons $a \in E$ tel que $\Phi(a) = 0_{E^*}$, c'est-à-dire tel que $\langle a|x \rangle = 0$ pour tout $x \in E$; en prenant $x = a$, on obtient immédiatement que $a = 0_E$ donc que $\operatorname{Ker}\Phi = \{0_E\}$. ■

Remarque
▹ Soit H un hyperplan de $\mathbb{R}^n$. Considérons une forme linéaire φ sur $\mathbb{R}^n$ telle que $H = \operatorname{Ker}\varphi$.
D'après la proposition précédente, il existe un unique $a = (a_1, \ldots, a_n) \in \mathbb{R}^n$ non nul tel que

$$\forall x \in E, \qquad \varphi(x) = \langle a|x \rangle.$$

Alors, un vecteur $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ appartient à H si, et seulement si, $\langle a|x \rangle = 0$, c'est-à-dire si

$$a_1 x_1 + \ldots + a_n x_n = 0$$

on a ainsi trouvé une équation de l'hyperplan vectoriel H.
▹ Pour obtenir une équation d'un hyperplan affine $\mathscr{H}$ de direction H, il suffit de changer le second membre de sorte à ce qu'un point $m \in \mathscr{H}$ vérifie l'équation. Plus précisément, $x \in \mathscr{H}$ si, et seulement si, $x - m \in H$ donc

$$a_1 x_1 + \ldots + a_n x_n = a_1 m_1 + \ldots + a_n m_n.$$

2. Orthogonalité

2.1. Définitions

Définition 20.16.

Soit E un espace préhilbertien réel de produit scalaire $\langle \, | \, \rangle$.
▹ Deux vecteurs de E sont orthogonaux si leur produit scalaire est nul.
▹ Une famille de vecteurs de E est orthogonale si le produit de deux vecteurs distincts de cette famille est nul.
▹ Une famille de vecteurs de E est orthonormale ou orthonormée si elle est orthogonale et si tous les vecteurs de cette famille sont de norme 1 (pour la norme associée au produit scalaire).

Exemple

Soit $a_0, \dots, a_n$ des réels deux à deux distincts. Considérons, pour tout $i \in [\![0, n]\!]$, le polynôme interpolateur de Lagrange

$$L_i(X) = \prod_{\substack{j=0 \\ j \neq i}}^{n} \frac{X - a_j}{a_i - a_j}.$$

Alors, la famille $(L_i)_{i \in [0,n]}$ est orthonormale pour le produit scalaire sur $\mathbb{R}[X]$ défini par

$$\langle P|Q \rangle = \sum_{i=0}^{n} P(a_i)Q(a_i).$$

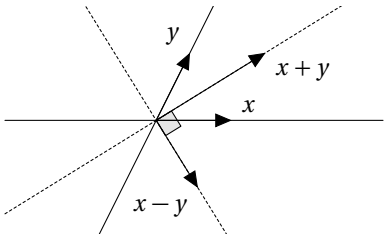
En effet, il suffit de remarquer que, pour tous $i, j \in [\![0, n]\!]$, $L_i(a_j) = \delta_{i,j}$ et les propriétés suivent directement.

Exemple

Soit x et y deux vecteurs de même norme. Alors, $x - y$ et $x + y$ sont orthogonaux car

$$\langle x - y | x + y \rangle = \|x\|^2 - \|y\|^2 = 0.$$

Ceci illustre un fait bien connu en géométrie plane. Considérons les droites dirigées par les vecteurs x et y (de même norme). Alors les vecteurs $x + y$ et $x - y$ dirigent les deux bissectrices de ces droites et le résultat ici établi indique que ces deux bissectrices sont orthogonales.



Proposition 20.17.

Soit E un espace préhilbertien réel. Toute famille orthogonale ne contenant pas le vecteur nul est libre. En particulier, toute famille orthonormale est libre.

Démonstration

Soit $(x_1, \dots, x_n)$ une famille orthogonale de vecteurs non nuls et $(\lambda_1, \dots, \lambda_n) \in \mathbb{R}^n$ tel que

$$\sum_{i=1}^{n} \lambda_i x_i = 0_{\mathbb{R}^n}.$$

En considérant le produit scalaire avec x_k, on trouve :

$$\begin{aligned} 0 = \langle \sum_{i=1}^{n} \lambda_i x_i | x_k \rangle &= \sum_{i=1}^{n} \lambda_i \langle x_i | x_k \rangle \\ &= \lambda_k \|x_k\|^2. \end{aligned}$$

Comme $x_k \neq 0_E$, $\lambda_k = 0$. En procédant pour tout $k \in [\![1, n]\!]$, on obtient que la famille est libre. ■

Les familles orthogonales sont intéressantes car elles permettent une simplification des identités remarquables. Avec une récurrence immédiate, on trouve donc le célèbre théorème de Pythagore.

Proposition 20.18. Théorème de Pythagore

Soit $x_1, \dots, x_n$ une famille orthogonale d'un espace préhilbertien réel. Alors

$$\left\|\sum_{k=1}^{n} x_i\right\|^2 = \sum_{k=1}^{n} \|x_i\|^2.$$

Remarque

Il est important de noter qu'avec plus de deux vecteurs, la réciproque est fausse ; considérons en effet les vecteurs $e_1 = (2,0)$, $e_2 = (2,0)$ et $e_3 = (-1,2)$ de $\mathbb{R}^2$ muni du produit scalaire canonique. Alors

$$\|e_1 + e_2 + e_3\|^2 = \|e_1\|^2 + \|e_2\|^2 + \|e_3\|^2,$$

et pourtant ces vecteurs ne sont pas orthogonaux entre eux.

2.2. Existence de bases orthonormales

Commençons par énoncer un premier résultat dont l'intérêt est également dans les aspects algorithmiques de sa preuve.

Théorème 20.19. Procédé d'orthonormalisation de Gram-Schmidt

Soit $(e_1, \dots, e_n)$ une famille libre d'un espace euclidien E. Il existe une unique famille orthonormalisée $(\varepsilon_1, \dots, \varepsilon_n)$ telle que, pour tout $p \in [\![1, n]\!]$,

- $(e_1, \dots, e_p)$ et $(\varepsilon_1, \dots, \varepsilon_p)$ engendrent le même sous-espace ;
- $\langle e_p | \varepsilon_p \rangle > 0$.

La première condition est essentielle pour la construction qui suit alors que la seconde condition ne sert intuitivement qu'à garantir l'unicité (en conservant « l'orientation » de la base).

Démonstration

On construit la famille $(\varepsilon_1, \dots, \varepsilon_n)$ par une récurrence finie :

▹ Le vecteur ε_1 doit être unitaire, colinéaire à e_1 et vérifier la condition $\langle e_1 | \varepsilon_1 \rangle > 0$: ε_1 est nécessairement égal à $\frac{1}{\|e_1\|} e_1$.

▹ Soit $p < n$. Supposons établies l'existence et l'unicité de la famille $(\varepsilon_1, \dots, \varepsilon_p)$ satisfaisant les conditions de l'énoncé.

Si un vecteur ε_{p+1} satisfait aux conditions du théorème alors

$$\begin{aligned} \varepsilon_{p+1} &\in \mathrm{Vect}(e_1, \dots, e_p) \oplus \mathrm{Vect}(e_{p+1}) \\ &\in \mathrm{Vect}(\varepsilon_1, \dots, \varepsilon_p) \oplus \mathrm{Vect}(e_{p+1}). \end{aligned}$$

Par conséquent, ε_{p+1} est de la forme $\sum_{i=1}^{p} \alpha_i \varepsilon_i + \alpha e_{p+1}$ avec $\alpha \neq 0$ car sinon, $\varepsilon_{p+1} \in \mathrm{Vect}(\varepsilon_1, \dots, \varepsilon_p)$ et la famille n'est pas orthonormale.

Puisque ε_{p+1} est orthogonal à ε_i pour $i \leq p$, alors $\langle \varepsilon_{p+1} | \varepsilon_i \rangle = 0$ et par conséquent $\alpha_i = -\alpha \langle e_{p+1} | \varepsilon_i \rangle$. De plus, comme ε_{p+1} est unitaire, on obtient $|\alpha| = \frac{1}{\|e_{p+1} - \sum_{i=1}^{p} \langle e_{p+1} | \varepsilon_i \rangle \varepsilon_i\|}$.

Enfin, la condition $\langle e_{p+1} | \varepsilon_{p+1} \rangle > 0$ se réécrit

$$\alpha \left(\|e_{p+1}\|^2 - \sum_{i=1}^{p} \langle e_{p+1} | \varepsilon_i \rangle^2 \right) > 0.$$

Remarquons alors que

$$\begin{aligned}\|e_{p+1}-\sum_{i=1}^{p}\langle e_{p+1}|\varepsilon_i\rangle\varepsilon_i\|^2 &= \|e_{p+1}\|^2+\|\sum_{i=1}^{p}\langle e_{p+1}|\varepsilon_i\rangle\varepsilon_i\|^2-2\langle e_{p+1}|\sum_{i=1}^{p}\langle e_{p+1}|\varepsilon_i\rangle\varepsilon_i\rangle\\ &= \|e_{p+1}\|^2+\sum_{i=1}^{p}\langle e_{p+1}|\varepsilon_i\rangle^2\|\varepsilon_i\|^2-2\sum_{i=1}^{p}\langle e_{p+1}|\varepsilon_i\rangle\langle e_{p+1}|\varepsilon_i\rangle\\ &= \|e_{p+1}\|^2-\sum_{i=1}^{p}\langle e_{p+1}|\varepsilon_i\rangle^2.\end{aligned}$$

Par conséquent,

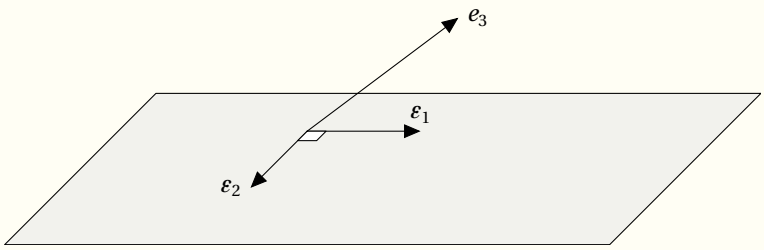
$$\|e_{p+1}\|^2-\sum_{i=1}^{p}\langle e_{p+1}|\varepsilon_i\rangle^2>0$$

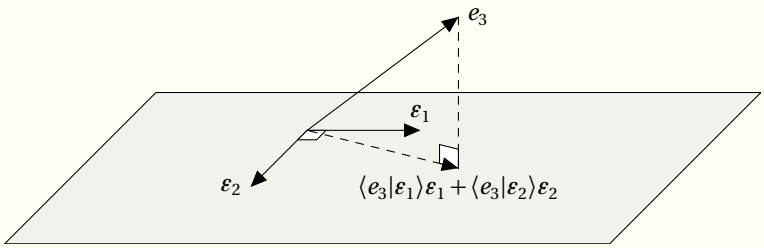
donc $\alpha>0$.
Réciproquement, le vecteur

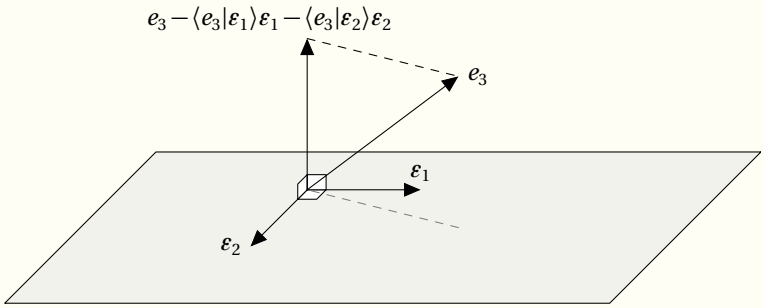
$$\varepsilon_{p+1}=\frac{e_{p+1}-\sum_{i=1}^{p}\langle e_{p+1}|\varepsilon_i\rangle\varepsilon_i}{\|e_{p+1}-\sum_{i=1}^{p}\langle e_{p+1}|\varepsilon_i\rangle\varepsilon_i\|}$$

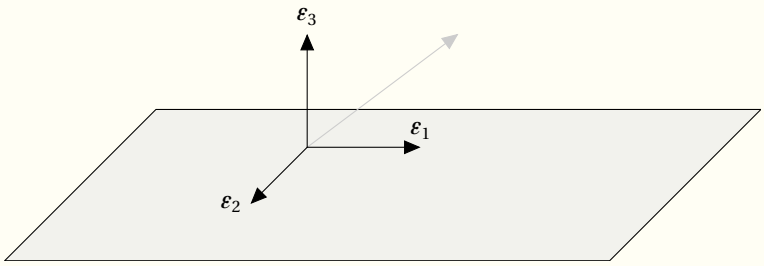
satisfait bien les conditions. ■

Voici une illustration de l'étape d'hérédité pour la construction de ε_3.









Comme tout espace de dimension finie admet une base (comme complétée d'une famille libre), l'application du procédé de Gram-Schmidt donne alors la proposition suivante.

Corollaire 20.20.

Tout espace euclidien admet une base orthonormée.

Démonstration

Un espace euclidien est de dimension finie donc admet une base. En orthonormalisant par le procédé de Gram-Schmidt, on obtient une famille orthonormée engendrant le même espace dont une base orthonormée. ■

Corollaire 20.21.

Toute famille orthonormale peut être complétée en une base orthonormale.

Démonstration

Une famille orthonormale $(e_1, \dots, e_p)$ est libre donc, d'après le théorème de la base incomplète, peut être complétée en une base $(e_1, \dots, e_n)$. Orthonormalisons cette famille par le procédé de Gram-Schmidt, on obtient une base orthonormée $(\varepsilon_1, \dots, \varepsilon_n)$. Par construction, pour tout $k \leq p$, $\varepsilon_k = e_k$ (car la famille de départ est déjà orthonormée). ■

Exemple

Soit $e_1 = (3,4,0)$, $e_2 = (5,0,0)$ et $e_3 = (0,0,1)$ des vecteurs qui forment une base de $\mathbb{R}^3$ (muni du produit scalaire canonique).
Déterminons la base orthonormalisée $(\varepsilon_1, \varepsilon_2, \varepsilon_3)$ associée par l'algorithme de Gram Schmidt.
▹ Commençons par normaliser e_1 pour obtenir ε_1 ; $\|e_1\| = 5$ donc $\varepsilon_1 = (\frac{3}{5}, \frac{4}{5}, 0)$.
▹ On sait qu'ensuite ε_2 est de la forme $\lambda(e_2 + \mu\varepsilon_1)$ avec $\mu = -\langle e_2|\varepsilon_1\rangle = -3$. Pour obtenir ε_2, il suffit alors de normaliser $e_2 - 3\varepsilon_1$. Comme $\|e_2 - 3\varepsilon_1\| = 4$, on obtient $\varepsilon_2 = (\frac{4}{5}, -\frac{3}{5}, 0)$.
▹ On peut recommencer la même procédure pour calculer ε_3 en l'écrivant comme combinaison

linéaire de e_3, ε_1 et ε_2. Toutefois, ce n'est pas nécessaire dans cet exemple car e_3 est orthogonal à ε_1 et ε_2 et est de norme 1.
La base recherchée est donc $(\frac{3}{5}, \frac{4}{5}, 0)$, $(\frac{4}{5}, -\frac{3}{5}, 0)$ et $(0, 0, 1)$.

2.3. Sous-espace orthogonal

Définition 20.22.

Soit E un espace euclidien. L'orthogonal d'une partie non vide $A \subset E$ est l'espace vectoriel

$$A^{\perp} = \{x \in E,\ \forall a \in A,\ \langle x|a\rangle = 0\}.$$

On fera bien attention à noter que $A^{\perp}$ est un sous-espace vectoriel de E pour toute partie A (c'est-à-dire même si A n'est pas un espace vectoriel). En effet, $0_E \in A^{\perp}$ et, pour tous x, $y \in A^{\perp}$ et tout $\lambda \in \mathbb{R}$,

$$\forall a \in A, \langle x + \lambda y|a\rangle = \langle x|a\rangle + \lambda\langle y|a\rangle = 0$$

donc $x + \lambda y \in A^{\perp}$.

Proposition 20.23.

Soit A et B deux parties de E telles que $A \subset B$. Alors, $B^{\perp} \subset A^{\perp}$.

Démonstration
Un vecteur $x \in B^{\perp}$ est orthogonal à tous les vecteurs de B donc à tous les vecteurs de A ; par conséquent, pour tout $x \in B^{\perp}$, $x \in A^{\perp}$. ■

Proposition 20.24.

L'orthogonal de l'espace euclidien E tout entier est réduit au vecteur nul.

Démonstration
Soit $x \in E^{\perp}$. Alors,

$$\forall y \in E, \quad \langle x|y\rangle = 0.$$

En particulier, avec $y = x$, on trouve $\|x\|^2 = 0$ et donc $x = 0_E$. ■

Corollaire 20.25.

Deux vecteurs x et y d'un espace euclidien E sont égaux si, et seulement s'ils vérifient $\langle x|z\rangle = \langle y|z\rangle$ pour tout $z \in E$.

Démonstration
Le sens direct est évident. Pour le sens retour, remarquons que cela entraîne (en ramenant tous les termes du même côté de l'égalité) que $x - y \in E^{\perp} = \{0_E\}$. ■

Conseils méthodologiques

Pour montrer l'égalité de deux termes dans un espace euclidien, il suffit de montrer qu'ils admettent le même produit scalaire avec n'importe quel autre vecteur. C'est pratique lorsque les termes sont essentiellement connus par leurs produits scalaires.

Proposition 20.26.

Soit F et G deux sous-espaces d'un espace euclidien E. Alors $F^\perp \cap G^\perp = (F+G)^\perp$.

Démonstration

($\subset$) Si $x \in F^\perp \cap G^\perp$, $f \in F$ et $g \in G$, alors

$$\langle x|f+g\rangle = \langle x|f\rangle + \langle x|g\rangle = 0.$$

Ainsi $F^\perp \cap G^\perp \subset (F+G)^\perp$.

($\supset$) Si $x \in (F+G)^\perp$, alors, pour tout $f \in F$ et tout $g \in G$,

$$\langle x|f\rangle = \langle x|f+0_E\rangle = 0,$$
$$\langle x|g\rangle = \langle x|0_E+g\rangle = 0.$$

D'où $(F+G)^\perp \subset F^\perp \cap G^\perp$. ■

Proposition 20.27.

Soit F un sous-espace d'un espace euclidien E.
▹ Les sous-espaces F et $F^\perp$ sont supplémentaires dans E.
▹ $(F^\perp)^\perp = F$.

Démonstration

▹ Remarquons que $F \cap F^\perp = \{0_E\}$ donc F et $F^\perp$ sont en somme directe.
De plus, si $(e_1, \dots, e_p)$ est une base orthonormale de F, nous pouvons la compléter en une base orthonormale de E notée $(e_1, \dots, e_n)$. Alors $(e_{p+1}, \dots, e_n)$ est une famille libre de $F^\perp$ d'où

$$\dim F^\perp \geq n-p = n - \dim F.$$

Par conséquent, F et $F^\perp$ sont supplémentaires dans E.
▹ D'après ce qui précède $\dim(F^\perp)^\perp = \dim F$. Or, pour tous $x \in F$ et $y \in F^\perp$, $\langle x|y\rangle = 0$. D'où $F \subset (F^\perp)^\perp$. Par conséquent, $F = (F^\perp)^\perp$. ■

Avec la proposition 20.26, on obtient le corollaire suivant.

Corollaire 20.28.

Soit F et G deux sous-espaces d'un espace euclidien E. Alors, $(F \cap G)^\perp = F^\perp + G^\perp$.

Exemple

Considérons le produit scalaire $\langle P|Q\rangle = \int_{-1}^{1} P(t)Q(t)\,\mathrm{d}t$ sur $\mathbb{R}[X]$ et définissons $\mathscr{P}$ (respectivement $\mathscr{I}$) l'espace vectoriel des polynômes pairs (respectivement impairs).

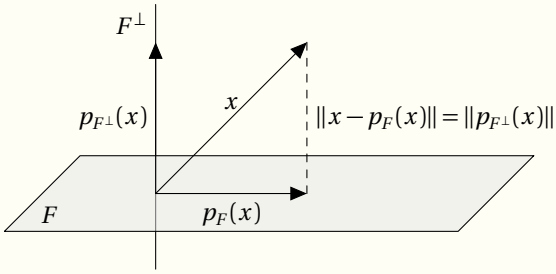
Montrons que $\mathscr{I} \subset \mathscr{P}^{\perp}$; en effet, pour tous $P \in \mathscr{P}$ et $Q \in \mathscr{I}$,

$$\begin{aligned}\langle P|Q\rangle &= \int_{-1}^{1} P(t)Q(t)\,\mathrm{d}t \\ &= \int_{1}^{-1} P(-t)Q(-t)(-\mathrm{d}t) = -\langle P|Q\rangle.\end{aligned}$$

D'où $\langle P|Q\rangle = 0$. Or, $\mathscr{P}$ et $\mathscr{I}$ sont supplémentaires donc $\mathscr{I} = \mathscr{P}^{\perp}$.

Définition 20.29.

Soit E un espace euclidien, et F un sous-espace de E. La projection orthogonale sur F est la projection sur F parallèlement au supplémentaire orthogonal $F^{\perp}$.



Proposition 20.30. Inégalité de Bessel

Soit E un espace euclidien et F un sous-espace de E. La projection orthogonale p_F sur F vérifie

$$\forall x \in E, \qquad \|p_F(x)\| \leq \|x\|.$$

Démonstration

Soit $x \in E$. Décomposons $x = p_F(x) + (x - p_F(x))$ et remarquons que $x - p_F(x) \in F^{\perp}$. Le théorème de Pythagore donne alors

$$\|x\|^2 = \|p_F(x)\|^2 + \|x - p_F(x)\|^2 \geq \|p_F(x)\|^2.$$

En passant à la racine, on obtient l'inégalité souhaitée. ■

Proposition 20.31.

Soit E un espace euclidien, $x \in E$ et F un sous-espace de E.
La distance de x à F définie par

$$d(x, F) = \inf\{\|x - y\|,\ y \in F\}$$

est donnée par $\|x - p_F(x)\|$, où p_F est la projection orthogonale de x sur le sous-espace F.

Démonstration
Pour tout $y \in F$, le théorème de Pythagore donne, avec $x - p_F(x) \in F^\perp$ et $p_F(x) - y \in F$,

$$\begin{aligned}\|x-y\|^2 &= \|x - p_F(x) + p_F(x) - y\|^2 \\ &= \|x - p_F(x)\|^2 + \|p_F(x) - y\|^2 \\ &\geq \|x - p_F(x)\|^2,\end{aligned}$$

car $x - p_F(x)$ et $p_F(x) - y$ sont orthogonaux. Ainsi, $\|x - p_F(x)\|$ est un minorant atteint de l'ensemble $\{\|x-y\|,\ y \in F\}$ donc c'est sa borne inférieure. ■

2.4. Calculs en base orthonormale

Proposition 20.32.

Soit E un espace euclidien de base orthonormale $(\varepsilon_1, \dots, \varepsilon_n)$ et $x \in E$.
Alors, les coordonnées de x dans cette base sont données par

$$x = \sum_{k=1}^{n} \langle x | \varepsilon_k \rangle \varepsilon_k.$$

Démonstration
Écrivons $x = \sum_{i=1}^{n} x_i \varepsilon_i$ puis calculons, pour chaque k, le produit scalaire

$$\langle x | \varepsilon_k \rangle = \Big\langle \sum_{i=1}^{n} x_i \varepsilon_i \Big| \varepsilon_k \Big\rangle = \sum_{i=1}^{n} x_i \langle \varepsilon_i | \varepsilon_k \rangle = x_k.$$

■

Remarque
Soit E un espace euclidien de base orthonormale $\mathscr{B} = (\varepsilon_1, \dots, \varepsilon_n)$ et $u \in \mathscr{L}(E)$.
D'après la proposition précédente, on obtient que le coefficient en position i, j de la matrice de u dans la base $\mathscr{B}$ est

$$[\mathrm{mat}_{\mathscr{B}}(u)]_{i,j} = \langle u(\varepsilon_j) | \varepsilon_i \rangle.$$

Proposition 20.33.

Soit E un espace euclidien de base orthonormale $(\varepsilon_1, \dots, \varepsilon_n)$ et $x,\ y \in E$ de décompositions

$$x = \sum_{i=1}^{n} x_i \varepsilon_i, \qquad y = \sum_{i=1}^{n} y_i \varepsilon_i.$$

Alors, le produit scalaire de x et y et la norme de x s'expriment à l'aide des coordonnées dans la base orthonormale de la façon suivante

$$\langle x | y \rangle = \sum_{i=1}^{n} x_i y_i, \qquad \|x\|^2 = \sum_{i=1}^{n} x_i^2.$$

Ce résultat justifie l'intérêt du produit scalaire canonique sur $\mathbb{R}^n$ puisque tout produit scalaire s'y ramène en considérant l'expression en base orthonormale.

Démonstration
Il suffit d'utiliser la bilinéarité du produit scalaire. En effet

$$\langle x|y\rangle = \Big\langle \sum_{i=1}^{n} x_i \varepsilon_i \Big| \sum_{j=1}^{n} y_j \varepsilon_j \Big\rangle = \sum_{i=1}^{n} x_i \sum_{j=1}^{n} y_j \langle \varepsilon_i | \varepsilon_j \rangle.$$

Or, $\langle \varepsilon_i|\varepsilon_j\rangle = \delta_{i,j}$ car la famille $(\varepsilon_1,\dots,\varepsilon_n)$ est orthonormale. D'où le résultat pour l'expression de $\langle x|y\rangle$ et celle de $\|x\|^2$ en prenant $y = x$. ■

Ce résultat se réécrit en utilisant les colonnes décrivant les vecteurs en base orthonormée.

Corollaire 20.34.

Soit E un espace euclidien de base orthonormée $\mathscr{B}$, x, $y \in E$. Alors, le produit scalaire de x et y s'exprime

$$\langle x|y\rangle = X^{\mathrm{T}} Y$$

avec $X = \mathrm{mat}_{\mathscr{B}}(x)$ et $Y = \mathrm{mat}_{\mathscr{B}}(y)$.

Proposition 20.35.

Soit F un sous-espace vectoriel, de base orthonormée $(e_1,\dots,e_p)$, de l'espace euclidien E. Pour tout $x \in E$, la projection orthogonale de x sur F est donnée par

$$p_F(x) = \sum_{k=1}^{p} \langle x|e_k\rangle e_k.$$

Démonstration
Complétons $(e_1,\dots,e_p)$ en une base orthonormale $(e_1,\dots,e_n)$ de E. Alors,

$$x = \sum_{k=1}^{n} \langle e_k|x\rangle e_k = \sum_{k=1}^{p} \langle e_k|x\rangle e_k + \sum_{k=p+1}^{n} \langle e_k|x\rangle e_k,$$

et $\sum_{k=1}^{p} \langle e_k|x\rangle e_k \in F$ et $\sum_{k=p+1}^{n} \langle e_k|x\rangle e_k \in F^{\perp}$.
Par conséquent, $p_F(x) = \sum_{k=1}^{p} \langle e_k|x\rangle e_k$. ■

Conseils méthodologiques

Soit F un sous-espace et x un vecteur de E. Pour déterminer la projection de x sur F (ou de manière équivalente la distance de x à F), il suffit de
- déterminer une base orthonormale de F ;
- calculer les produits scalaires des éléments de cette base avec x.

Toutefois, quand la dimension de F est grande (ou que $F^{\perp}$ est mieux connu), il est plus intéressant de calculer la projection de x sur $F^{\perp}$ puis d'utiliser que $p_F(x) = x - p_{F^{\perp}}(x)$.

Exemple

Calculons la projection orthogonale de $x = (1,1,1)$ sur le plan F d'équation $x + y - 2z = 0$. Pour cela, remarquons que $e_1 = (1,1,1)$ et $e_2 = (1,-1,0)$ est une base orthogonale de F. La base orthonormalisée de F est donc $\varepsilon_1 = \frac{1}{\sqrt{3}}(1,1,1)$ et $\varepsilon_2 = \frac{1}{\sqrt{2}}(1,-1,0)$. Or, $\langle \varepsilon_1|x\rangle = \frac{1}{\sqrt{3}}$ et $\langle \varepsilon_2|x\rangle = 0$ d'où $p_F(x) = \frac{1}{\sqrt{3}}\varepsilon_1 + 0\varepsilon_2 = \frac{1}{3}(1,1,1)$.

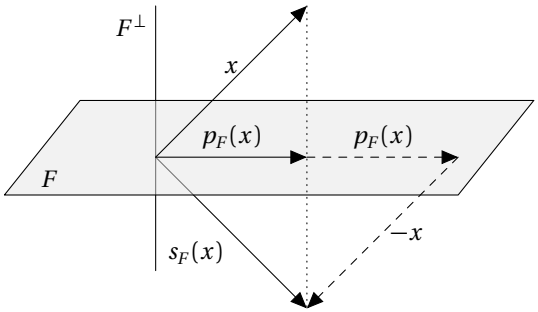
Exemple

Le symétrique du vecteur x par rapport au sous-espace vectoriel F de base orthonormale $(e_1,\dots,e_p)$ est

$$s_F(x)=p_F(x)-(x-p_F(x))=2p_F(x)-x,$$

et en remplaçant $p_F(x)$ par son expression sur la famille $(e_1,\dots,e_p)$, on obtient

$$s_F(x)=2\sum_{k=1}^{p}\langle x|e_k\rangle e_k-x.$$



Exemple

L'expression de la projection orthogonale du vecteur x sur le sous-espace vectoriel F de base orthogonale $(e_1,\dots,e_p)$ est

$$p_F(x)=\sum_{k=1}^{p}\frac{\langle x|e_k\rangle}{\|e_k\|^2}e_k.$$

En effet, la famille $(\frac{e_1}{\|e_1\|},\dots,\frac{e_p}{\|e_p\|})$ est une base orthonormale de F donc nous pouvons appliquer la proposition précédente et

$$p_F(x)=\sum_{k=1}^{p}\langle x\Big|\frac{e_k}{\|e_k\|}\rangle\frac{e_k}{\|e_k\|}.$$

D'où le résultat.

3. Isométries vectorielles

3.1. Généralités

Définition 20.36.

Une isométrie vectorielle (ou endomorphisme orthogonal) d'un espace euclidien E est un endomorphisme $u\in\mathscr{L}(E)$ tel que $\|u(x)\|=\|x\|$ pour tout $x\in E$.
L'ensemble des isométries vectorielles est noté $\mathscr{O}(E)$.

Proposition 20.37.

L'ensemble $\mathscr{O}(E)$ est un sous-groupe de $\mathrm{GL}(E)$.

Démonstration

▹ Soit $u \in \mathscr{O}(E)$ et $x \in \operatorname{Ker} u$. Alors, $\|x\| = \|u(x)\| = 0$ donc $x = 0_E$: par conséquent, $\operatorname{Ker} u = \{0_E\}$ et l'endomorphisme u est injectif donc inversible (car E est de dimension finie).

En conclusion, $\mathscr{O}(E) \subset \mathrm{GL}(E)$.

▹ Soit $u, v \in \mathscr{O}(E)$ et $x \in E$. Alors,

$$\|u \circ v(x)\| = \|u(v(x))\| = \|v(x)\| = \|x\|,$$

donc $u \circ v \in \mathscr{O}(E)$.

▹ Soit $u \in \mathscr{O}(E)$ et $x \in E$. Alors, $\|x\| = \|u(u^{-1}(x)\| = \|u^{-1}(x)\|$. Par conséquent, $u^{-1} \in \mathscr{O}(E)$.

D'après la caractérisation des sous-groupes, $\mathscr{O}(E)$ est un sous-groupe de $\mathrm{GL}(E)$. ■

Proposition 20.38.

Un endomorphisme u d'un espace euclidien E est une isométrie vectorielle si, et seulement si, pour tous $x, y \in E$,

$$\langle u(x) | u(y) \rangle = \langle x | y \rangle.$$

Démonstration

Le sens retour est évident en prenant $x = y$. Pour le sens direct, utilisons une identité de polarisation ; pour tous $x, y \in E$,

$$\begin{aligned}\langle u(x) | u(y) \rangle &= \frac{1}{2}\left(\|u(x)\|^2 + \|u(y)\|^2 - \|u(x) - u(y)\|^2\right) \\ &= \frac{1}{2}\left(\|x\|^2 + \|y\|^2 - \|x - y\|^2\right) = \langle x | y \rangle.\end{aligned}$$

■

Proposition 20.39.

Un endomorphisme u d'un espace euclidien E est une isométrie vectorielle si, et seulement si, l'image d'une base orthonormée de E par u est une base orthonormée de E.

Démonstration

Le sens direct est évident avec la proposition précédente. Pour le sens retour utilisons l'expression de la norme en base orthonormée. Soit $(e_1, \ldots, e_n)$ une base orthonormée de E telle que $(u(e_1), \ldots, u(e_n))$ soit orthonormée et $x = \sum_{k=1}^{n} x_i e_i$. Alors, en appliquant le théorème de Pythagore dans les deux bases orthonormées,

$$\|u(x)\|^2 = \left\|\sum_{k=1}^{n} x_i\, u(e_i)\right\|^2 = \sum_{k=1}^{n} x_i^2 = \left\|\sum_{k=1}^{n} x_i e_i\right\|^2 = \|x\|^2,$$

donc $u \in \mathscr{O}(E)$. ■

Exemple

Soit s la symétrie orthogonale par rapport à F. Soit $x \in E$, $x_F \in F$ et $x_{F^\perp} \in F^\perp$ tels que $x = x_F + x_{F^\perp}$.

Alors,
$$\|x\|^2 = \|x_F\|^2 + \|x_{F^\perp}\|^2 = \|x_F - x_{F^\perp}\|^2 = \|s(x)\|^2$$
donc $s \in \mathscr{O}(E)$.
En conclusion, une symétrie orthogonale est une isométrie vectorielle.

Définition 20.40.

Une réflexion est une symétrie orthogonale par rapport à un hyperplan.

3.2. Matrices orthogonales

Définition 20.41.

Une matrice $A \in \mathscr{M}_n(\mathbb{R})$ est orthogonale si $A^{\mathrm{T}}A = \mathrm{I}_n$.
L'ensemble des matrices orthogonales est noté $\mathscr{O}_n(\mathbb{R})$.

Proposition 20.42.

Une matrice est orthogonale si, et seulement si, ses vecteurs-colonnes forment une base orthonormale de l'espace $\mathscr{M}_{n,1}(\mathbb{R})$ muni du produit scalaire canonique.

Démonstration
Soit $A \in \mathscr{M}_n(\mathbb{R})$. Il suffit de remarquer que $[A^{\mathrm{T}}A]_{i,j} = C_i(A)^{\mathrm{T}} C_j(A) = \langle C_i(A) | C_j(A) \rangle$.
Donc $A^{\mathrm{T}}A = \mathrm{I}_n$ si, et seulement si, pour tous i, j, $\langle C_i(A) | C_j(A) \rangle = \delta_{i,j}$. ■

Exemple

La matrice de passage entre deux bases orthonormées est orthogonale.

Exemple

Trouvons $a, b, c, d, e \in \mathbb{R}$ tels que la matrice
$$\frac{1}{\sqrt{6}} \begin{bmatrix} 1 & 1 & c \\ 1 & a & d \\ 2 & b & e \end{bmatrix}$$
soit orthogonale. Remarquons tout d'abord que le premier vecteur colonne est bien de norme 1.
Pour que le second vecteur colonne forme avec le premier une famille orthonormée, il faut que
$$\begin{cases} 1 + a + 2b = 0 \\ 1 + a^2 + b^2 = 6 \end{cases}$$
ce qui entraîne $b = \frac{-2 \pm 2\sqrt{6}}{5}$ et donc $a = \frac{-1 \mp 4\sqrt{6}}{5}$. Afin de trouver le dernier vecteur colonne, il suffit de considérer le produit vectoriel (ou son opposé) des deux premiers vecteurs colonnes. Les calculs sont laissés au lecteur.

Proposition 20.43.

Soit E un espace euclidien et $\mathscr{B}$ une base orthonormée de E. Alors, un endomorphisme u de E est une isométrie vectorielle si, et seulement si, la matrice $\mathrm{mat}_{\mathscr{B}}(u)$ est orthogonale.

Démonstration

Notons $\mathscr{B} = (e_1, \dots, e_n)$. La matrice $\mathrm{mat}_{\mathscr{B}}(u)$ est orthogonale si, et seulement si, ses colonnes forment une base orthonormée c'est-à-dire que les vecteurs $u(e_1), \dots, u(e_n)$ forment une base orthonormée, c'est-à-dire encore si u est une isométrie vectorielle. ■

Remarque

Avec cette proposition et le fait que $\mathscr{O}(E)$ est un sous-groupe de $\mathrm{GL}(E)$, on obtient immédiatement que $\mathscr{O}_n(\mathbb{R})$ est sous groupe de $\mathrm{GL}_n(\mathbb{R})$.

Proposition 20.44.

Le déterminant d'une matrice orthogonale est 1 ou -1.

Démonstration

Soit $A \in \mathscr{O}_n(\mathbb{R})$. Comme $A^{\mathrm{T}} A = \mathrm{I}_n$, $(\det A)^2 = 1$ donc $\det A \in \{\pm 1\}$. ■

Définition 20.45.

Une matrice orthogonale ou une isométrie vectorielle est positive si son déterminant vaut 1 ; elle est négative sinon.
L'ensemble des matrices orthogonales positives est noté $\mathscr{SO}_n(\mathbb{R})$, l'ensemble des isométries vectorielles positives de E est noté $\mathscr{SO}(E)$.

Exemple

La matrice $-\mathrm{I}_n$ est orthogonale positive si n est pair, négative sinon.

Exemple

La matrice de passage entre deux bases orthonormées directes est orthogonale positive.

Exemple

Considérons quelques exemples importants pour la suite. Soit $\theta \in \mathbb{R}$.

- La matrice $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ est celle d'une isométrie positive.
- La matrice $\begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}$ est celle d'une isométrie négative.

Proposition 20.46.

L'ensemble $\mathscr{SO}_n(\mathbb{R})$ est un sous-groupe de $\mathscr{O}_n(\mathbb{R})$.

Démonstration
Il suffit de remarquer que le produit et l'inverse de matrices de déterminant 1 sont bien de déterminant 1. ■

3.3. Cas de la dimension 2

Proposition 20.47.

- $\forall M \in \mathscr{SO}_2(\mathbb{R}),\ \exists \theta \in \mathbb{R}, \qquad M = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$
- $\forall M \in \mathscr{O}_2(\mathbb{R}) \setminus \mathscr{SO}_2(\mathbb{R}),\ \exists \theta \in \mathbb{R}, \qquad M = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}.$

Démonstration
Soit $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ une matrice orthogonale.
La condition d'orthonormalité des colonnes implique que $a^2+c^2 = b^2+d^2 = 1$ et $ab+cd = 0$; en posant $a = \cos\theta$ et $d = \cos\theta'$, on trouve $c = \sin\theta$, $b = \sin\theta'$ (aux signes près ce qui revient à changer θ ou θ' en leurs opposés) et la dernière condition implique $\sin(\theta+\theta') = 0$ d'où $\theta' = -\theta[\pi]$ d'où le résultat selon le déterminant. ■

Proposition 20.48.

Soit $\rho \in \mathscr{SO}(E)$. Il existe un unique réel $\theta \in [0, 2\pi[$ (ou, de manière équivalente, un unique réel défini modulo 2π) tel que la matrice de ρ dans n'importe quelle base orthonormée directe soit

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

Démonstration
La matrice de ρ dans la base canonique appartient à $\mathscr{SO}_2(\mathbb{R})$ donc il existe θ tel que cette matrice soit

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

Pour toute autre base orthonormée directe, la matrice de passage est encore une matrice de $\mathscr{SO}_2(\mathbb{R})$ donc il existe α tel que cette matrice soit

$$\begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}.$$

Pour conclure, il suffit de remarquer que

$$\begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}^{-1} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix},$$

donc que la matrice dans la nouvelle base orthonormée directe est la même que celle dans la base canonique. ■

Définition 20.49.

La rotation d'angle $\theta \in \mathbb{R}$ est l'application $\rho_\theta \in \mathscr{SO}(E)$ dont la matrice dans toute base orthonormée

directe est
$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

En effectuant le calcul du produit matriciel de matrices de $\mathscr{SO}_2(\mathbb{R})$, on obtient immédiatement :

Proposition 20.50.

Pour tous $\theta_1, \theta_2 \in \mathbb{R}$,
$$\rho_{\theta_1} \circ \rho_{\theta_2} = \rho_{\theta_2} \circ \rho_{\theta_1} = \rho_{\theta_1+\theta_2}, \qquad \rho_{\theta_1}^{-1} = \rho_{-\theta_1}.$$

Proposition 20.51.

Soit x, y deux vecteurs du plan de norme 1. Il existe une unique rotation ρ telle que $\rho(x) = y$.

Démonstration

Considérons une base orthonormée directe du plan (e_1, e_2) telle que $e_1 = x$ et notons a, b les coordonnées de y dans cette base. S'il existe une rotation ρ envoyant x sur y alors sa première colonne est nécessairement $\begin{bmatrix} a \\ b \end{bmatrix}$.

Par conséquent, la matrice de cette unique rotation possible dans la base (e_1, e_2) est (d'après la proposition 20.47)
$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix}.$$

Pour conclure, il suffit alors de vérifier que cette matrice définit effectivement une rotation qui envoie x sur y. ■

Définition 20.52.

L'angle (orienté) entre deux vecteurs du plan non nuls x et y est la mesure modulo 2π de l'angle de l'unique rotation ρ telle que $\rho\left(\frac{x}{\|x\|}\right) = \frac{y}{\|y\|}$. Il est noté $(\widehat{x, y})$.

Proposition 20.53.

Pour tous x, y, z vecteurs du plan non nuls et $\lambda \in \mathbb{R}_+^*$, on a les relations de Chasles suivantes :
$$(\widehat{x, z}) = (\widehat{x, y}) + (\widehat{y, z})[2\pi], \qquad (\widehat{y, x}) = -(\widehat{x, y})[2\pi].$$

On retrouve les expressions connues pour le produit scalaire canonique et le déterminant en base orthonormée directe.

Proposition 20.54.

Pour tous x, y vecteurs du plan non nuls,
$$\langle x | y \rangle = \|x\|.\|y\|.\cos(\widehat{x, y}), \qquad \det(x, y) = \|x\|.\|y\|.\sin(\widehat{x, y}).$$

Proposition 20.55.

Pour tous x, y vecteurs du plan non nuls et toute rotation $\rho \in \mathscr{SO}(E)$,

$$(\widehat{\rho(x), \rho(y)}) = (\widehat{x, y})[2\pi].$$

Démonstration

Soit r la rotation qui envoie $\frac{x}{\|x\|}$ sur $\frac{y}{\|y\|}$. Comme r et ρ commutent (car le groupe $\mathscr{SO}(E)$ est commutatif), r envoie également $\frac{\rho(x)}{\|x\|}$ sur $\frac{\rho(y)}{\|y\|}$. Par définition de l'angle (orienté) entre deux vecteurs non nuls, $(\widehat{\rho(x), \rho(y)}) = (\widehat{x, y})[2\pi]$. ■

Proposition 20.56.

Les isométries négatives du plan sont les réflexions.

Démonstration

Dans une base orthonormée, une isométrie négative admet une matrice de la forme

$$\begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}.$$

On montre immédiatement que le carré de cette matrice est I_2 : c'est une symétrie orthogonale. Le sous-espace par rapport auquel elle symétrise est une droite (car sinon, c'est un sous-espace trivial et les applications Id et $-$Id sont des isométries positives). ■

Remarque

Remarquons toutefois que lorsque l'on change de base orthonormée directe la matrice d'une réflexion n'est pas conservée puisque le produit

$$\begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}^{-1} \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix} \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}$$

est égal à

$$\begin{bmatrix} \cos(\theta-2\alpha) & \sin(\theta-2\alpha) \\ \sin(\theta-2\alpha) & -\cos(\theta-2\alpha) \end{bmatrix}.$$

Ainsi, on voit qu'en appliquant une rotation d'angle $\frac{\theta}{2}$ au repère initial, on est ramené à la matrice

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Pour calculer la matrice d'une réflexion σ dont on connaît un vecteur normal u de la droite par rapport à laquelle on symétrise, on peut revenir à l'expression explicite de cette réflexion

$$\forall x \in E, \qquad \sigma(x) = x - 2\frac{1}{\|u\|^2}\langle x|u\rangle u.$$

Par exemple, la réflexion σ relative à $\mathrm{Vect}(7e_1 + 12e_2)$ admet pour matrice

$$\frac{1}{193}\begin{bmatrix} -97 & -168 \\ -168 & 97 \end{bmatrix},$$

car $\sigma(e_1) = \begin{bmatrix} 1 \\ 0 \end{bmatrix} - \frac{2.7}{193}\begin{bmatrix} 7 \\ 12 \end{bmatrix} = \frac{1}{193}\begin{bmatrix} -97 \\ -168 \end{bmatrix}$.

Proposition 20.57.

Soit x, y deux vecteurs du plan de norme 1. Il existe une unique réflexion σ telle que $\sigma(x) = y$.

Démonstration

▹ Si $x = y$, alors la réflexion par rapport à la droite $\mathrm{Vect}(x)$ est la seule qui convient.

▹ Sinon, $x - y$ est orthogonal à la droite d'une réflexion envoyant x sur y. La seule réflexion possible correspond donc à la droite $\mathrm{Vect}(x-y)^{\perp} = \mathrm{Vect}(x+y)$ et on vérifie immédiatement qu'elle convient. ■

Proposition 20.58.

Soit σ une réflexion et ρ une rotation. Alors, $\sigma \circ \rho \circ \sigma = \rho^{-1}$.

Démonstration

Il suffit de remarquer que $\sigma \circ \rho$ est une réflexion car c'est une isométrie vectorielle de déterminant $\det\sigma \det\rho = -1$. ■

Proposition 20.59.

Pour tous x, y vecteurs du plan non nuls et toute réflexion $\sigma \in \mathscr{O}(E) \setminus \mathscr{SO}(E)$,

$$(\widehat{\sigma(x), \sigma(y)}) = -(\widehat{x, y})[2\pi].$$

Démonstration

Soit ρ la rotation qui envoie $\frac{x}{\|x\|}$ sur $\frac{y}{\|y\|}$. Alors, ρ^{-1} envoie $\frac{\sigma(x)}{\|x\|}$ sur $\frac{\sigma(y)}{\|y\|}$ car

$$\begin{aligned}\rho^{-1}\Big(\frac{\sigma(x)}{\|x\|}\Big) &= \sigma \circ \rho \circ \sigma\Big(\frac{\sigma(x)}{\|x\|}\Big)\\ &= \sigma \circ \rho\Big(\frac{x}{\|x\|}\Big)\\ &= \sigma\Big(\frac{y}{\|y\|}\Big) = \frac{\sigma(y)}{\|y\|}.\end{aligned}$$

Ainsi, $(\widehat{\sigma(x), \sigma(y)}) = -(\widehat{x, y})[2\pi]$. ■

FICHE

SYNTHÈSE

Ce chapitre reprend et approfondit la notion d'orthogonalité dans les $\mathbb{R}$-espaces vectoriels. La première définition est celle de produit scalaire qui sert de fondement à la notion de vecteurs orthogonaux.

Produit scalaire

▹ Un produit scalaire est une forme bilinéaire, symétrique, définie positive. Un espace euclidien est un $\mathbb{R}$-espace vectoriel de dimension finie muni d'un produit scalaire.
▹ La norme associée au produit scalaire $\langle\ |\ \rangle$ est l'application $x \longmapsto \sqrt{\langle x|x\rangle} = \|x\|$.
▹ Les identités remarquables sont, pour tous $x,\ y \in E$,

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2\langle x|y\rangle,$$
$$\|x-y\|^2 = \|x\|^2 + \|y\|^2 - 2\langle x|y\rangle,$$
$$\|x\|^2 - \|y\|^2 = \langle x+y|x-y\rangle.$$

▹ Pour tous $x,\ y \in E$, $|\langle x|y\rangle| \leq \|x\|\|y\|$. Le cas d'égalité correspond à deux vecteurs colinéaires.

Outre l'orthogonalité des vecteurs, on peut définir l'orthogonal d'une partie (et le supplémentaire orthogonal d'un sous-espace vectoriel, ce qui permet de parler de projection ou de symétrie orthogonale).

Orthogonal d'une partie

▹ L'orthogonal d'une partie A d'un espace euclidien E est le sous-espace

$$A^\perp = \{x \in E,\ \forall a \in A,\ \langle a|x\rangle = 0\}.$$

▹ Pour tout sous-espace F d'un espace euclidien $E = F \oplus F^\perp$ et $(F^\perp)^\perp = F$.
▹ Pour tous sous-espaces F et G d'un espace euclidien E,

$$(F+G)^\perp = F^\perp \cap G^\perp, \qquad (F \cap G)^\perp = F^\perp + G^\perp.$$

Pour un produit scalaire fixé, on peut chercher des familles voire des bases orthonormales. Le procédé de Gram-Schmidt permet d'en obtenir dans les espaces euclidiens (avec un algorithme explicite).

Calcul en base orthonormée

Soit $\mathscr{B} = (e_1, \ldots, e_n)$ une base orthonormée d'un espace euclidien E.
Alors, pour tous $x = \sum_{k=1}^n x_k e_k$ et $y = \sum_{k=1}^n y_k e_k$,
▹ $\langle x|y\rangle = \displaystyle\sum_{k=1}^n x_k y_k$, et $\|x\|^2 = \displaystyle\sum_{k=1}^n x_k^2$,
▹ la projection orthogonale de x sur $F = \mathrm{Vect}(e_1, \ldots, e_p)$ est $p_F(x) = \displaystyle\sum_{k=1}^p x_k e_k$.

Comme le produit scalaire (*via* la norme associée) fournit une notion de distance, on peut étudier les endomorphismes qui conservent les distances.

Isométries et matrices orthogonales

Un endomorphisme u d'un espace euclidien E est une isométrie vectorielle s'il vérifie l'une des propriétés équivalentes suivantes :

- pour tout $x \in E$, $\|u(x)\| = \|x\|$,
- pour tous x, $y \in E$, $\langle u(x) | u(y) \rangle = \langle x | y \rangle$,
- l'image d'une base orthonormée de E par u est une base orthonormée de E,
- sa matrice dans une base orthonormée est orthogonale.

On traite enfin le cas particulier des matrices orthogonales de taille 2 (avec seulement des rotations et des réflexions).

Matrices orthogonales de taille 2

- $\forall M \in \mathscr{SO}_2(\mathbb{R}),\ \exists \theta \in \mathbb{R}, \qquad M = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$.
- $\forall M \in \mathscr{O}_2(\mathbb{R}) \setminus \mathscr{SO}_2(\mathbb{R}),\ \exists \theta \in \mathbb{R}, \qquad M = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}$.

EXERCICES

Vrai ou faux ?

		Vrai	Faux
a)	$(A,B)\mapsto \operatorname{tr}(BA^{\mathrm{T}})$ est un produit scalaire sur $\mathscr{M}_n(\mathbb{R})$.	☐	☐
b)	Soit $S\in\mathscr{S}_n(\mathbb{R})$. $(X,Y)\mapsto X^{\mathrm{T}}SY$ est un produit scalaire sur $\mathscr{M}_{n,1}(\mathbb{R})$.	☐	☐
c)	$(f,g)\mapsto \int_0^1 f(t)g(t)\,\mathrm{d}t$ est un produit scalaire sur $\mathscr{C}^0([-1,1],\mathbb{R})$.	☐	☐
d)	Soit $a_1<a_2<\ldots<a_n$. $(P,Q)\mapsto \sum_{k=1}^{n} P(a_k)Q(a_k)$ est un produit scalaire sur $\mathbb{R}_n[X]$.	☐	☐
e)	Soit $(x_1,x_2,\ldots,x_n)\in\mathbb{R}^n$. Alors $$\left(\sum_{k=1}^{n} x_k\right)^2 \leqslant n\sum_{k=1}^{n} x_k^2.$$	☐	☐
f)	Toute famille orthogonale d'un espace euclidien est libre.	☐	☐
g)	Tout espace euclidien admet une base orthonormée.	☐	☐
h)	Pour tout produit scalaire, $\mathbb{R}[X]$ admet une base orthonormée échelonnée en degré.	☐	☐
i)	Pour toute partie A d'un espace euclidien, $A^{\perp}$ est un sous-espace vectoriel.	☐	☐
j)	Pour tous sous-espaces vectoriels F et G d'un espace euclidien, $F^{\perp}\cap G^{\perp}=(F+G)^{\perp}$.	☐	☐
k)	Dans un espace euclidien E, tout sous-espace vectoriel admet un supplémentaire orthogonal.	☐	☐
l)	Si $(e_1,\ldots,e_p)$ est une base orthogonale de F, la projection orthogonale de x sur F est $$\sum_{k=1}^{p}\langle x\mid e_k\rangle e_k.$$	☐	☐
m)	Un endomorphisme est une isométrie vectorielle si, et seulement s'il transforme toute paire de vecteurs orthogonaux en une paire de vecteurs orthogonaux.	☐	☐
n)	Une matrice M est orthogonale si, et seulement si, $MM^{\mathrm{T}}=\mathrm{I}_n$.	☐	☐
o)	Une matrice triangulaire orthogonale est diagonale.	☐	☐
p)	Les projections orthogonales sont des isométries vectorielles.	☐	☐

q)	Les symétries orthogonales sont des isométries vectorielles.	□	□
r)	Toute matrice de $\mathscr{O}_2(\mathbb{R})\setminus\mathscr{SO}_2(\mathbb{R})$ s'écrit de la forme $\begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}$.	□	□

Exercices d'application du chapitre 20

●○○ **Exercice 1**

Soit $x_1,\dots,x_n \in \mathbb{R}_+^*$ tels que $x_1+\dots+x_n=1$. Montrer que $n^2 \leq \sum_{k=1}^{n} \frac{1}{x_k}$.

●○○ **Exercice 2**

Soit $f:[0,1]\to\mathbb{R}_+$ continue. Posons, pour tout $n\in\mathbb{N}$,

$$I_n = \int_0^1 t^n f(t)\,\mathrm{d}t\,.$$

Montrer que, pour tous $n,\,p\in\mathbb{N}$, $I_{n+p}^2 \leq I_{2n}I_{2p}$.

●○○ **Exercice 3**

Montrer que, pour toute matrice $A\in\mathscr{M}_n(\mathbb{R})$, $|\operatorname{tr} A| \leq \sqrt{n}\|A\|$.

●○○ **Exercice 4**

Soit f et g deux fonctions continues de $\mathbb{R}$ dans $\mathbb{R}$ telles que $f(x)g(x)\geq 4$ pour tout $x\in\mathbb{R}$. Montrer que

$$\left(\int_{-1}^{2} f(x)\,\mathrm{d}x\right)\left(\int_{-1}^{2} g(x)\,\mathrm{d}x\right) \geq 36.$$

●○○ **Exercice 5**

Montrer que deux vecteurs x, y d'un espace euclidien sont orthogonaux si, et seulement si,

$$\forall\lambda\in\mathbb{R},\qquad \|x+\lambda y\| \geq \|x\|.$$

●●○ **Exercice 6**

Soit n vecteurs $x_1,\dots,x_n$ d'un espace euclidien E. La matrice de Gram de la famille $(x_1,\dots,x_n)$ est la matrice $G(x_1,\dots,x_n)=(\langle x_i|x_j\rangle)_{i,j\in[1,n]}$.

1. Montrer que la famille $(x_1,\dots,x_n)$ est libre si, et seulement si, $G(x_1,\dots,x_n)$ est inversible.
2. Soit n vecteurs $x_1,\dots,x_n$ et $\mathscr{B}=(e_1,\dots,e_n)$ une base orthonormale. Montrer que
$$\det G(x_1,\dots,x_n) = [\det_{\mathscr{B}}(x_1,\dots,x_n)]^2.$$
3. Soit $x_1,\dots,x_p\in E$, $F=\operatorname{Vect}(x_1,\dots,x_p)$ et $x\in E$. Montrer que
$$d(x,F)^2\det G(x_1,\dots,x_p) = \det G(x,x_1,\dots,x_p).$$

EXERCICES

●●○ **Exercice 7**

Considérons $\mathbb{R}_n[X]$ muni du produit scalaire

$$\langle P|Q\rangle = \int_0^1 P(x)Q(x)\,\mathrm{d}x\,.$$

Posons L_k le polynôme égal à la dérivée $k^{\text{ième}}$ de $[X(X-1)]^k$ pour $k \in [\![0,n]\!]$.

1. Montrer que la famille $(L_k)_{k\in[0,n]}$ est orthogonale.
2. Calculer la norme euclidienne de L_k pour $k \in [\![0,n]\!]$.

●●○ **Exercice 8**

Considérons $\mathbb{R}_n[X]$ muni de $\langle P|Q\rangle = \int_0^1 P(t)Q(t)\,\mathrm{d}t$ et P un polynôme non-nul de $\mathbb{R}_{n-1}[X]^\perp$.

1. Préciser le degré de P.
2. Montrer que la fonction $R : x \mapsto \int_0^1 P(t)t^x\,\mathrm{d}t$ est rationnelle.
3. Trouver R à une constante multiplicative près.
4. En déduire les coefficients de P.
5. En déduire une base orthogonale de $\mathbb{R}_n[X]$.

●●○ **Exercice 9**

Considérons $\mathscr{C}([0,1],\mathbb{R})$ muni du produit scalaire $\langle f|g\rangle = \int_0^1 f(t)g(t)\,\mathrm{d}t$, et l'hyperplan

$$H = \{f \in \mathscr{C}([0,1],\mathbb{R}),\ f(0)=0\}.$$

Déterminer $H^\perp$.

●○○ **Exercice 10**

Considérons $\mathbb{R}^n$ muni du produit scalaire canonique et l'hyperplan

$$H = \{(x_1,\dots,x_n) \in \mathbb{R}^n,\ a_1x_1 + \cdots + a_nx_n = 0\}$$

où $a_1,\dots,a_n$ sont des réels fixés non tous nuls. Déterminer la matrice dans la base canonique de la projection orthogonale sur H.

●○○ **Exercice 11**

Soit $a_0,\dots,a_n \in \mathbb{R}$ deux à deux distincts. Munissons $\mathbb{R}_n[X]$ du produit scalaire

$$(P,Q) \mapsto \langle P|Q\rangle = \sum_{k=0}^{n} P(a_k)Q(a_k).$$

Calculer la distance de X^n au sous-espace

$$F = \{P \in \mathbb{R}_n[X],\ \sum_{k=0}^{n} P(a_k) = 0\}.$$

●○○ **Exercice 12**

Considérons $E = \mathscr{M}_2(\mathbb{R})$ muni du produit scalaire $(A,B) \mapsto \mathrm{tr}(A^{\mathrm{T}}B)$ et le sous-espace

$$F = \left\{\begin{bmatrix} a & b \\ -b & a \end{bmatrix},\ a,b \in \mathbb{R}\right\}.$$

1. Déterminer une base orthonormée de F.
2. En déduire la distance à F de la matrice

$$J = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

●●○ **Exercice 13**

Considérons $\mathbb{R}_n[X]$ avec le produit scalaire canonique. Posons $H = \{P \in E,\ P(1) = 0\}$.

1. Trouver une base orthonormale de H.
2. Montrer que $d(X, H) = \frac{1}{\sqrt{n+1}}$.

●●○ **Exercice 14**

Soit E un espace euclidien et $e_1, \dots, e_n \in E$ tels que

$$\forall x \in E, \qquad \|x\|^2 = \sum_{i=1}^{n} \langle x | e_i \rangle^2.$$

1. Montrer que la famille $(e_1, \dots, e_n)$ est génératrice.
2. Supposons, dans cette question, que les vecteurs $e_1, \dots, e_n$ sont unitaires. Montrer que $(e_1, \dots, e_n)$ est une base orthonormale de E.
3. Supposons, dans cette question, que $\dim E = n$.
   a. Montrer que $(e_1, \dots, e_n)$ est une base de E.
   b. Montrer que
   $$\forall (x, y) \in E^2, \qquad \langle x | y \rangle = \sum_{i=1}^{n} \langle x | e_i \rangle \langle y | e_i \rangle.$$
   c. Soit $M \in \mathscr{M}_n(\mathbb{R})$ la matrice dont le coefficient en position i, j est $\langle e_i | e_j \rangle$. Montrer que $M^2 = M$ et conclure.

●●○ **Exercice 15**

Soit p un projecteur orthogonal d'un espace euclidien E.

1. Montrer que $\|p(x)\|^2 = \langle p(x) | x \rangle$ pour tout $x \in E$.
2. Montrer que, pour toute base orthonormée $(e_1, \dots, e_n)$ de E,
$$\sum_{k=1}^{n} \|p(e_k)\|^2 = \mathrm{rg}(p).$$

●●○ **Exercice 16**

Soit p un projecteur d'un espace euclidien E. Montrer que p est un projecteur orthogonal si, et seulement si, $\|p(x)\| \leq \|x\|$ pour tout $x \in E$.

●○○ **Exercice 17**

Déterminer a, b, c tels que la matrice

$$\begin{bmatrix} -\frac{\sqrt{6}}{3} & a & \frac{\sqrt{3}}{3} \\ \frac{\sqrt{6}}{6} & -\frac{\sqrt{2}}{2} & b \\ c & \frac{\sqrt{2}}{2} & \frac{\sqrt{3}}{3} \end{bmatrix}$$

soit une matrice orthogonale.

●○○ **Exercice 18**

Soit $a,\ b \in \mathbb{R}$. Montrer qu'il existe une matrice $R \in \mathscr{SO}_2(\mathbb{R})$ tel que

$$R \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} c \\ 0 \end{bmatrix}$$

où c est un réel à préciser.

EXERCICES

●●○ **Exercice 19**

Soit $A=(a_{i,j})_{1\leq i,j\leq n}\in\mathscr{O}_n(\mathbb{R})$. Montrer les inégalités suivantes

1. $\displaystyle\sum_{1\leq i,j\leq n}\left|a_{i,j}\right|\leq n^{\frac{3}{2}}$.
2. $\displaystyle\left|\sum_{1\leq i,j\leq n}a_{i,j}\right|\leq n$.

●●● **Exercice 20**

Soit F et G deux sous-espaces d'un espace euclidien E tels que $\dim F=\dim G$. Montrer qu'il existe $u\in\mathscr{O}(E)$ tel que $u(F)=G$.

●●● **Exercice 21**

Soit E un espace euclidien, $p\geq 1$, $x_1,\ldots,x_p$, $y_1,\ldots,y_p\in E$. Montrer qu'il existe $u\in\mathscr{O}(E)$ telle que, pour tout i de $[\![1,p]\!]$, $u(x_i)=y_i$ si, et seulement si,

$$\forall(i,j)\in[\![1,p]\!]^2,\qquad\langle x_i|x_j\rangle=\langle y_i|y_j\rangle.$$

Exercices d'approfondissement du chapitre 20

●○○ **Exercice A**

Soit E un espace euclidien, une suite $(x_n)\in E^{\mathbb{N}}$ et $x\in E$. Montrer l'équivalence entre les propriétés

$$\|x_n-x\|\to 0,$$

et

$$\forall y\in E,\qquad\langle x_n-x|y\rangle\to 0.$$

●●○ **Exercice B**

Soit E un espace euclidien, F, G deux sous-espaces. Montrer que $F=G^{\perp}$ si, et seulement si, pour tout $x\in E$,

$$\|x\|^2=d(x,F)^2+d(x,G)^2.$$

●●○ **Exercice C**

Soit E un espace euclidien et $u\in\mathscr{L}(E)$ tel que, pour tout $x\in E$, $\langle u(x)|x\rangle=0$. Montrer que

$$E=\operatorname{Ker}u\oplus\operatorname{Im}u.$$

●●○ **Exercice D**

Soit x, y deux vecteurs d'un espace euclidien de dimension au moins 2 tels que $\langle x|y\rangle=\|y\|^2$. Montrer qu'il existe un unique hyperplan H tel que la projection orthogonale de x sur H est y.

●●● **Exercice E**

Soit F, G deux sous-espaces d'un espace euclidien E tels que $F^{\perp}$ et $G^{\perp}$ sont orthogonaux, p_F et p_G les projections orthogonales sur F et sur G. Montrer que $p_F+p_G-p_{F\cap G}=\mathrm{Id}_E$ et $p_F\circ p_G=p_G\circ p_F=p_{F\cap G}$.

●●○ **Exercice F**

Soit E un espace euclidien. Déterminer les endomorphismes $u\in\mathscr{L}(E)$ tels que, pour tout sous-espace F de E, $u(F^{\perp})=u(F)^{\perp}$.

Problème du chapitre 20

Problème 1

Définition 20.60.

Une fonction de poids sur un intervalle I est une fonction $\omega : I \to \mathbb{R}_+^*$ continue par morceaux.

L'objectif de ce problème est d'étudier une famille de polynômes orthogonaux pour un produit scalaire relatif à une fonction de poids puis de s'en servir pour fournir une méthode de calcul approché d'intégrales.

Partie A – Base ω-orthogonale échelonnée

Soit ω une fonction de poids sur un intervalle I.

1. Vérifier que $(P,Q) \mapsto (P|Q)_\omega = \int_I P(t)Q(t)\omega(t)\,\mathrm{d}t$ est un produit scalaire sur $\mathbb{R}[X]$.

Définition 20.61.

Une base ω-orthogonale, respectivement ω-orthonormale, échelonnée (en degrés) est une famille orthogonale, respectivement orthonormale, (pour $(.|.)_\omega$) $(P_n)_n$ de polynômes tels que

$$\forall n \in \mathbb{N}, \qquad \deg(P_n) = n.$$

2. a. Montrer qu'il existe une base ω-orthogonale échelonnée.
   b. Considérons deux bases ω-orthogonales échelonnées $(P_n)_n$ et $(Q_n)_n$. Montrer que pour tout $n \in \mathbb{N}$, les polynômes P_n et Q_n sont colinéaires.
3. Montrer que, pour tout $n \geq 1$, P_n admet n racines dans l'intervalle I.
   Indication : on pourra introduire un polynôme dont les racines sont celles de P_n d'ordre impair dans l'intervalle I.
4. Considérons une base ω-orthogonale échelonnée $(P_n)_n$ et définissons,
   - $a_0 = P_0(0)$ et $b_0 = 0$,
   - a_1 et b_1 les réels tels que $P_1(X) = a_1 X + b_1$,
   - pour tout $n \geq 2$, a_n et b_n les réels tels que $P_n - a_n X^n - b_n X^{n-1} \in \mathbb{R}_{n-2}[X]$.

   a. Vérifier qu'il existe un unique couple $(\alpha_0, \beta_0) \in \mathbb{R}^2$ tel que
   $$XP_0 = \alpha_0 P_1 + \beta_0 P_0.$$
   b. Montrer que, pour tout $n \geq 2$ et tout $Q \in \mathbb{R}_{n-2}[X]$, $(XP_n|Q)_\omega = 0$.
   c. En déduire que, pour tout $n \geq 1$, il existe un unique triplet $(\alpha_n, \beta_n, \gamma_n) \in \mathbb{R}^3$ tel que
   $$XP_n = \alpha_n P_{n+1} + \beta_n P_n + \gamma_n P_{n-1}.$$
   d. Montrer que, pour tout $n \geq 1$,
   $$\alpha_n = \frac{a_n}{a_{n+1}}, \qquad \beta_n = \frac{b_n}{a_n} - \frac{b_{n+1}}{a_{n+1}}, \qquad \gamma_n = \alpha_{n-1}\frac{\|P_n\|_\omega^2}{\|P_{n-1}\|_\omega^2}$$
   où $\|.\|_\omega$ désigne la norme associée au produit scalaire $(.|.)_\omega$.

Partie B – Formule de quadrature

Soit ω une fonction de poids sur I, $(P_n)_n$ une base ω-orthonormale échelonnée. Fixons $n \in \mathbb{N}^*$ et notons $\lambda_1^{(n)} < \ldots < \lambda_n^{(n)}$ les n racines de P_n dans l'ordre croissant.

PROBLÈME

1. Montrer qu'il existe $c_1, \dots, c_n \in \mathbb{R}$ tels que
$$\forall Q \in \mathbb{R}_{n-1}[X], \qquad \int_I Q(t)\omega(t)\,\mathrm{d}t = \sum_{k=1}^{n} c_k Q\big(\lambda_k^{(n)}\big).$$
2. Montrer que
$$\forall Q \in \mathbb{R}_{2n-1}[X], \qquad \int_I Q(t)\omega(t)\,\mathrm{d}t = \sum_{k=1}^{n} c_k Q\big(\lambda_k^{(n)}\big).$$
Indication : on pourra effectuer la division euclidienne de Q par P_n.
3. Vérifier que les coefficients $c_1, \dots, c_n$ sont positifs.
4. Définissons, pour toute fonction $f : I \to \mathbb{R}$ de classe $\mathscr{C}^{2n}$, l'erreur de quadrature par
$$\mathscr{E}(f) = \int_I f(t)\omega(t)\,\mathrm{d}t - \sum_{k=1}^{n} c_k f\big(\lambda_k^{(n)}\big).$$
   a. Montrer qu'il existe un unique polynôme $H \in \mathbb{R}_{2n-1}[X]$ tel que, pour tout $k \leq n$,
$$H(\lambda_k^{(n)}) = f(\lambda_k^{(n)}) \qquad H'(\lambda_k^{(n)}) = f'(\lambda_k^{(n)}).$$
   b. Montrer que $\mathscr{E}(f) = \mathscr{E}(f - H)$.
   c. Soit $x \in I$ qui n'est pas racine de P_n et $\lambda \in \mathbb{R}$ tel que la fonction $f - H - \lambda P_n^2$ s'annule en x. Montrer qu'il existe $c_x \in \mathbb{R}$ tel que
$$\lambda = \frac{1}{d_n^2 (2n)!} f^{(2n)}(c_x),$$
   où d_n désigne le coefficient dominant de P_n.
   Indication : on pourra appliquer le théorème de Rolle.
   d. En déduire la majoration suivante de l'erreur
$$|\mathscr{E}(f)| \leq \frac{1}{d_n^2 (2n)!} \|f^{(2n)}\|_\infty \int_I P_n(x)^2 \omega(x)\,\mathrm{d}x$$
   avec $\|f^{(2n)}\|_\infty = \sup\{|f^{(2n)}(t)|,\ t \in I\}$.

CORRIGÉS

Corrigés des Vrai/Faux

a) Vrai, c'est le produit scalaire canonique.

b) Faux, il manque *a priori* le caractère défini positif.

c) Faux.

d) Faux, le polynôme $(X-a_1)\dots(X-a_n)$ serait de « norme » nulle.

e) Vrai.

f) Faux si elle contient le vecteur nul.

g) Vrai.

h) Vrai avec l'algorithme de Gram-Schmidt appliqué à la base canonique.

i) Vrai.

j) Vrai.

k) Vrai.

l) Faux, la base n'est a priori pas orthonormée.

m) Faux avec par exemple, 2Id.

n) Vrai, l'inverse de M est M^{T}.

o) Vrai, en utilisant l'orthogonalité des colonnes.

p) Faux, à part l'identité, aucun n'est inversible.

q) Vrai.

r) Vrai.

Corrigés des exercices d'application du chapitre 20

Exercice 1

Appliquons l'inégalité de Cauchy-Schwarz dans $\mathbb{R}^n$ muni de son produit scalaire canonique aux vecteurs $(\sqrt{x_1},\dots,\sqrt{x_n})$ et $(\frac{1}{\sqrt{x_1}},\dots,\frac{1}{\sqrt{x_n}})$:

$$n^2=\left(\sum_{k=1}^{n}\sqrt{x_k}\frac{1}{\sqrt{x_k}}\right)^2\leq\left(\sum_{k=1}^{n}\sqrt{x_k}^2\right)\left(\sum_{k=1}^{n}\frac{1}{\sqrt{x_k}^2}\right)=\sum_{k=1}^{n}\frac{1}{x_k}.$$

Exercice 2

Considérons $E = \mathscr{C}^0([0,1],\mathbb{R})$ muni du produit scalaire défini par $\langle f|g\rangle = \int_0^1 f(t)g(t)\,\mathrm{d}t$. Appliquons l'inégalité de Cauchy-Schwarz aux fonctions $f_n : t \mapsto t^n\sqrt{f(t)}$ et $f_p : t \mapsto t^p\sqrt{f(t)}$.

$$I_{n+p}^2 = \langle f_n|f_p\rangle^2 \leq \|f_n\|^2\|f_p\|^2 = I_{2n}I_{2p}.$$

Exercice 3

Considérons l'espace euclidien $\mathscr{M}_n(\mathbb{R})$ muni de son produit scalaire canonique. Appliquons l'inégalité de Cauchy-Schwarz à A et I_n

$$|\operatorname{tr} A| = |\langle \mathrm{I}_n|A\rangle| \leq \|I_n\|\|A\| = \sqrt{n}\|A\|.$$

Exercice 4

Par hypothèse, les fonctions continues f et g ne s'annulent pas et sont de même signe. Supposons sans perte de généralité ces fonctions positives.

Appliquons l'inégalité de Cauchy-Schwarz dans l'espace préhilbertien $\mathscr{C}^0([-1,2],\mathbb{R})$ aux fonctions $\sqrt{f}$ et $\sqrt{g}$. Alors,

$$\left(\int_{-1}^2 f(x)\,\mathrm{d}x\right)\left(\int_{-1}^2 g(x)\,\mathrm{d}x\right) \geq \left(\int_{-1}^2 \sqrt{f(x)g(x)}\,\mathrm{d}x\right)^2 \geq \left(\int_{-1}^2 2\,\mathrm{d}x\right)^2 = 36.$$

Exercice 5

On peut supposer $y \neq 0_E$ car pour $y = 0_E$ le résultat est évident.

($\Rightarrow$) On remarque que, pour tout $\lambda \in \mathbb{R}$, $\|x+\lambda y\|^2 = \|x\|^2 + \lambda^2\|y\|^2$ (c'est le théorème de Pythagore) donc $\|x+\lambda y\|^2 \geq \|x\|^2$.

($\Leftarrow$) Pour tout $\lambda \in \mathbb{R}$, $\|x+\lambda y\|^2 = \|x\|^2 + 2\lambda\langle x|y\rangle + \lambda^2\|y\|^2$. L'inégalité supposée entraîne donc

$$\forall \lambda \in \mathbb{R}, \qquad \lambda\left(2\langle x|y\rangle + \lambda\|y\|^2\right) \geq 0.$$

On obtient donc un polynôme de degré 2 en λ qui toujours positif : il admet donc au plus une racine, qui est 0. Par conséquent, $\langle x|y\rangle = 0$ et x et y sont orthogonaux.

Exercice 6

1. Si la matrice $G(x_1,\dots,x_n)$ n'est pas inversible, alors il existe une combinaison linéaire nulle non-triviale de ses vecteurs colonnes $C_1,\dots,C_n$

$$\sum_{k=1}^n \lambda_k C_k = 0_{n,1}.$$

Alors, pour tout $j \in [\![1,n]\!]$, $\langle \sum_{k=1}^n \lambda_k x_k|x_j\rangle = 0$. En particulier $\langle \sum_{k=1}^n \lambda_k x_k|\sum_{k=1}^n \lambda_k x_k\rangle = 0$ et il existe une combinaison linéaire nulle non-triviale de la famille $(x_1,\dots,x_n)$ qui est donc liée.

Réciproquement, si $\sum_{k=1}^n \lambda_k x_k = 0_E$, alors $\sum_{k=1}^n \lambda_k C_k = 0_{n,1}$.

2. Si X désigne la matrice des coordonnées de la famille $(x_1,\dots,x_n)$ dans la base $\mathscr{B}$ (c'est-à-dire la matrice dont le coefficient en position i,j est $\langle x_j|e_i\rangle$), alors, l'expression du produit scalaire en base orthonormée nous donne

$$\forall i,j \in [\![1,n]\!], \qquad \langle x_i|x_j\rangle = \sum_{k=1}^n \langle x_i|e_k\rangle\langle x_j|e_k\rangle.$$

Donc $G(x_1,\dots,x_n) = X^{\mathrm{T}}X$. On conclut en passant au déterminant $\det(G(x_1,\dots,x_n)) = (\det X)^2$.

Remarque
Le résultat de cette question permet de répondre d'une manière différente à la première question : la matrice $G(x_1,\dots,x_n)$ est inversible si, et seulement si, $\det_{\mathscr{B}}(x_1,\dots,x_n)\neq 0$ c'est-à-dire si la famille $(x_1,\dots,x_n)$ est libre.

3. Écrivons $x=p_F(x)+(x-p_F(x))$ dans l'expression de $\det G(x,x_1,\dots,x_p)$ puis utilisons que

$$\forall k\in[\![1,p]\!],\qquad \langle x-p_F(x)|x_k\rangle=0.$$

Ainsi

$$\det G(x,x_1,\dots,x_p)=\begin{vmatrix} \|p_F(x)\|^2+\|x-p_F(x)\|^2 & \langle p_F(x)|x_1\rangle & \dots & \langle p_F(x)|x_p\rangle \\ \langle p_F(x)|x_1\rangle & \langle x_1|x_1\rangle & \dots & \langle x_1|x_p\rangle \\ \vdots & \vdots & & \vdots \\ \langle p_F(x)|x_p\rangle & \langle x_1|x_p\rangle & \dots & \langle x_p|x_p\rangle \end{vmatrix}.$$

Utilisons ensuite la linéarité en la première colonne pour obtenir l'expression

$$\det G(p_F(x),x_1,\dots,x_p)+\begin{vmatrix} \|x-p_F(x)\|^2 & \langle p_F(x)|x_1\rangle & \dots & \langle p_F(x)|x_p\rangle \\ 0 & \langle x_1|x_1\rangle & \dots & \langle x_1|x_p\rangle \\ \vdots & \vdots & & \vdots \\ 0 & \langle x_1|x_p\rangle & \dots & \langle x_p|x_p\rangle \end{vmatrix}.$$

Or, la famille $(p_F(x),x_1,\dots,x_p)$ est liée donc le premier terme est nul et on calcule le second par développement par rapport à la première colonne

$$\det G(x,x_1,\dots,x_p)=\|x-p_F(x)\|^2\det G(x_1,\dots,x_p).$$

Exercice 7
1. Soit $k>j$. Calculons $\langle L_k|L_j\rangle$ en effectuant $l\leq k$ intégrations par parties successives et en remarquant la nullité des crochets

$$\begin{aligned}\langle L_k|L_j\rangle&=-\int_0^1\left[(x(x-1))^k\right]^{(k-1)}\left[(x(x-1))^j\right]^{(j+1)}\mathrm{d}x\\ &=(-1)^l\int_0^1\left[(x(x-1))^k\right]^{(k-l)}\left[(x(x-1))^j\right]^{(j+l)}\mathrm{d}x\,.\end{aligned}$$

Avec $l=j+1$, on obtient $\left[(x(x-1))^j\right]^{(2j+1)}=0$ et donc $\langle L_k|L_j\rangle=0$.
2. Reprenons le calcul précédent avec $l=j=k\in[\![0,n]\!]$

$$\|L_k\|^2=(-1)^k\int_0^1\left[(x(x-1))^k\right]^{(0)}\left[(x(x-1))^k\right]^{(2k)}\mathrm{d}x=(-1)^k(2k)!\int_0^1(x(x-1))^k\,\mathrm{d}x\,.$$

Calculons encore cette dernière intégrale par intégrations par parties (en intégrant le terme x^k et en dérivant le terme $(x-1)^k$)

$$\|L_k\|^2=(-1)^k(-1)^k\int_0^1\frac{k!}{(2k)!}x^{2k}.k!\,\mathrm{d}x=\frac{k!^2}{2k+1}.$$

CORRIGÉS

Exercice 8

1. Remarquons que P est de degré n car sinon P serait orthogonal à P donc nul, ce qui est impossible par définition.
2. Notons $P = \sum_{k=0}^{n} a_k X^k$. Alors, pour tout $x \in \mathbb{R}$,

$$R(x) = \sum_{k=0}^{n} a_k \int_0^1 t^{k+x}\,\mathrm{d}t = \sum_{k=0}^{n} \frac{a_k}{k+1+x}.$$

D'après cette expression, R est rationnelle.
3. D'après la question précédente, R s'écrit $\frac{N(x)}{\prod_{k=0}^{n}(k+1+x)}$ avec N de degré au plus n.

Or, l'hypothèse $P \in \mathbb{R}_{n-1}[X]^{\perp}$ entraîne que les entiers $0, 1, \ldots, (n-1)$ sont racines de R donc de N. En conclusion, $N(x) = \lambda \prod_{k=0}^{n-1}(X-k)$.
4. On dispose de deux expressions de R : l'une avec les coefficients de P,

$$R(x) = \sum_{k=0}^{n} \frac{a_k}{k+1+x},$$

l'autre à partir de l'expression de la question précédente en développant en éléments simples

$$R(x) = \sum_{k=0}^{n} \frac{\lambda(-1)^{n+k}\frac{(n+k)!}{k!^2(n-k)!}}{k+1+x}.$$

En identifiant, on obtient

$$P(X) = \lambda \sum_{k=0}^{n} (-1)^{n+k} \binom{2k}{k}\binom{n+k}{2k} X^k.$$

5. En notant P_n le polynôme obtenu à la question précédente avec $\lambda = 1$, on a immédiatement que $(P_0, \ldots, P_n)$ est une base orthogonale de $\mathbb{R}_n[X]$.

Exercice 9

Pour tout $f \in H^{\perp}$, $g : x \mapsto f(x)x^2$ appartient à H donc

$$0 = \langle f|g\rangle = \int_0^1 f(t)^2 t^2\,\mathrm{d}t\,.$$

En particulier, f est nulle sur $]0,1]$ et par continuité sur $[0,1]$. En conclusion, $H^{\perp} = \{0\}$.

Exercice 10

Commençons par remarquer que le vecteur $a = (a_1, \ldots, a_n)$ dirige la droite $H^{\perp}$ car $x = (x_1, \ldots, x_n)$ appartient à H si, et seulement si, $\langle a|x\rangle = 0$. Ainsi, le vecteur $\frac{a}{\|a\|}$ constitue une base orthonormée de $H^{\perp}$. D'après la formule du cours, la projection orthogonale sur $H^{\perp}$ est donc

$$p_{H^{\perp}} : x \mapsto \frac{\langle a|x\rangle}{\|a\|^2} a.$$

Comme la projection orthogonale sur H se déduit de $p_{H^{\perp}}$ par la relation $p_H = \mathrm{Id}_E - p_{H^{\perp}}$, on obtient que les images de la base canonique $(e_1, \ldots, e_n)$ sont données par $p_H(e_i) = e_i - \frac{a_i}{\|a\|^2} a$. En conclusion, le coefficient de position (i, j) de la matrice de p_H dans la base canonique est $\delta_{i,j} - \frac{a_i a_j}{\sum_{k=1}^{n} a_k^2}$.

Exercice 11

Remarquons que par définition du produit scalaire, F est l'hyperplan orthogonal au vecteur $P = \sum_{k=0}^{n} X^k$. Alors,

$$d(X^n, F) = \|p_{F^\perp}(X^n)\| = \|\langle X^n | P\rangle P\|.$$

Après calculs, $\langle X^n | P\rangle = 1$ et $\|P\| = \sqrt{n+1}$ donc $d(X^n, F) = \sqrt{n+1}$.

Exercice 12

1. Remarquons tout d'abord que, par définition, une base de F est composée de I_2 et $E_{2,1} - E_{1,2}$. Pour obtenir une base orthonormée, il suffit de lui appliquer le procédé de Gram-Schmidt.

▹ La matrice I_2 est de norme $\sqrt{2}$ donc on pose $M_1 = \frac{1}{\sqrt{2}}\mathrm{I}_2$.

▹ La matrice $E_{2,1} - E_{1,2}$ est orthogonale à M_1 et de norme $\sqrt{2}$ donc on pose $M_2 = \frac{1}{\sqrt{2}}(E_{2,1} - E_{1,2})$.

La base (M_1, M_2) est orthonormée.

2. La projection orthogonale de J sur le sous-espace F est

$$\langle J | M_1\rangle M_1 + \langle J | M_2\rangle M_2 = \mathrm{I}_2$$

donc $d(J, F) = \|J - \mathrm{I}_2\| = \sqrt{2}$.

Exercice 13

1. Vérifions qu'une base orthogonale est fournie par la famille $(P_1, \dots, P_n)$ avec $P_k = -kX^k + \sum_{j=0}^{k-1} X^j$.

▹ Pour tout $k \in [\![1, n]\!]$, $P_k(1) = -k + \sum_{j=0}^{k-1} 1 = 0$ donc $P_k \in H$.

▹ Pour tout $k > l$, $\langle P_k | P_l\rangle = -l + \sum_{j=0}^{l-1} 1 = 0$. Donc la famille $(P_1, \dots, P_n)$ est orthogonale et ne contient pas le vecteur nul, donc est libre. Comme son cardinal est $n = \dim H$, c'est une base orthogonale de H.

Pour obtenir une base orthonormée, il ne reste plus qu'à diviser chacun de ces polynômes par sa norme. Comme $\|P_k\| = \sqrt{k^2 + \sum_{j=0}^{k-1} 1} = \sqrt{k(k+1)}$, une base orthonormée de H est composée des $\frac{1}{\sqrt{k(k+1)}} P_k$ pour tout $k \in [\![1, n]\!]$.

2. N'utilisons pas la première question et calculons la projection orthogonale sur $H^\perp$. Notons tout d'abord que $P = \sum_{k=0}^{n} X^k$ dirige $H^\perp$ donc en forme une base. La projection orthogonale d'un polynôme Q sur $H^\perp$ est donc

$$p_{H^\perp}(Q) = \frac{\langle Q | P\rangle}{\|P\|^2} P = \frac{Q(1)}{n+1} \sum_{k=0}^{n} X^k.$$

Alors, $d(X, H) = \|p_{H^\perp}(X)\| = \frac{1}{n+1}\|P\| = \frac{1}{\sqrt{n+1}}$.

Exercice 14

1. Raisonnons par l'absurde en supposant que $\mathrm{Vect}(e_1, \dots, e_n) \neq E$ et considérons $x \in \mathrm{Vect}(e_1, \dots, e_n)^\perp$ non nul. L'identité de l'énoncé donne alors $\|x\|^2 = \sum_{i=1}^{n} 0 = 0$ donc $x = 0_E$: contradiction.

Ainsi, $\mathrm{Vect}(e_1, \dots, e_n) = E$ et la famille est génératrice.

2. Pour tout $j \in [\![1, n]\!]$,

$$\|e_j\|^2 = \sum_{i=1}^{n} \langle e_j | e_i\rangle^2.$$

D'où, comme les vecteurs sont unitaires

$$0 = \sum_{\substack{i=1 \\ i \neq j}}^{n} \langle e_j | e_i\rangle^2.$$

Par conséquent, pour tous $i \neq j$, $\langle e_i | e_j \rangle = 0$: la famille est orthogonale et composée de vecteurs unitaires donc est orthonormée et *a fortiori* libre. Comme elle est génératrice d'après la première question, c'est une base orthonormée de E.
3. a. D'après la première question la famille $(e_1, \dots, e_n)$ est génératrice. Or, ici son cardinal est égal à la dimension de l'espace E donc elle est une base de E.
b. Il s'agit d'utiliser une identité de polarisation. Soit $x, y \in E$

$$\begin{aligned}
\langle x|y \rangle &= \frac{1}{2}(\|x\|^2 + \|y\|^2 - \|x-y\|^2) \\
&= \frac{1}{2}\left(\sum_{i=1}^{n} \langle x|e_i \rangle^2 + \sum_{i=1}^{n} \langle y|e_i \rangle^2 - \sum_{i=1}^{n} \langle x-y|e_i \rangle^2\right) \\
&= \frac{1}{2}\sum_{i=1}^{n} \left(\langle x|e_i \rangle^2 + \langle y|e_i \rangle^2 - (\langle x|e_i \rangle - \langle y|e_i \rangle)^2\right) \\
&= \sum_{i=1}^{n} \langle x|e_i \rangle \langle y|e_i \rangle.
\end{aligned}$$

c. En appliquant la formule précédente à $x = e_i$ et $y = e_j$, on a $[M]_{i,j} = \sum_{k=1}^{n} [M]_{i,k}[M]_{k,j}$. Ainsi, $M^2 = M$.
▹ Montrons par l'absurde que la matrice M est inversible : il existe une combinaison linéaire nulle non-triviale de ses vecteurs colonnes $C_1, \dots, C_n$

$$\sum_{k=1}^{n} \lambda_k C_k = 0_{n,1}.$$

Alors, pour tout $j \in [\![1, n]\!]$, $\langle \sum_{k=1}^{n} \lambda_k e_k | e_j \rangle = 0$.
En particulier $\langle \sum_{k=1}^{n} \lambda_k e_k | \sum_{k=1}^{n} \lambda_k e_k \rangle = 0$ et il existe une combinaison linéaire nulle non-triviale de la famille $(e_1, \dots, e_n)$: contradiction avec la propriété de base.
▹ En conclusion, $M = \mathrm{I}_n$ et, en identifiant les coefficients, la famille est orthonormée.

Exercice 15

1. Pour tout $x \in E$, $p(x) = x + (p(x) - x)$ et $\langle p(x) - x | p(x) \rangle = 0$ car la projection p est orthogonale. Ainsi

$$\|p(x)\|^2 = \langle p(x)|p(x) \rangle = \langle p(x)|x \rangle + \langle p(x)|p(x) - x \rangle = \langle p(x)|x \rangle.$$

2. Rappelons que la trace d'un projecteur est son rang (il suffit de se placer dans une base adaptée à la décomposition de ce projecteur). Or,

$$\mathrm{tr}(p) = \sum_{k=1}^{n} \langle p(e_k)|e_k \rangle = \sum_{k=1}^{n} \|p(e_k)\|^2,$$

d'après la première question. En conclusion, $\sum_{k=1}^{n} \|p(e_k)\|^2 = \mathrm{rg}(p)$.

Exercice 16

($\Rightarrow$) Supposons que p est la projection sur F parallèlement à $F^{\perp}$. Alors, pour tout $x \in E$,

$$\|x\|^2 = \|p(x)\|^2 + \|x - p(x)\|^2 \geq \|p(x)\|^2,$$

d'où $\|x\| \geq \|p(x)\|$.
($\Leftarrow$) Soit $x \in \mathrm{Im}\, p$ et $y \in \mathrm{Ker}\, p$. Alors, pour tout $\lambda \in \mathbb{R}$,

$$\|x\|^2 = \|p(x + \lambda y)\|^2 \leq \|x + \lambda y\|^2 = \|x\|^2 + 2\lambda \langle x|y \rangle + \lambda^2 \|y\|^2.$$

Ainsi, pour tout $\lambda \in \mathbb{R}$, $2\lambda\langle x|y\rangle + \lambda^2\|y\|^2 \geq 0$: avec une étude rapide de ce polynôme réel, il apparaît que ceci n'est possible que si $\langle x|y\rangle = 0$. Par conséquent, $\operatorname{Im} p$ et $\operatorname{Ker} p$ sont orthogonaux et p est une projection orthogonale.

Exercice 17

▹ Il faut tout d'abord que chaque colonne soit de norme 1, ce qui entraîne $c^2 = \frac{1}{6}$, $a^2 = 0$ (donc $a = 0$) et $b^2 = \frac{1}{3}$.

▹ Ensuite, il faut exprimer l'orthogonalité des colonnes. En utilisant que $a = 0$, l'orthogonalité entre C_1 et C_2 donne $c = \frac{\sqrt{6}}{6}$, l'orthogonalité entre C_2 et C_3 donne $b = \frac{\sqrt{3}}{3}$. On vérifie qu'avec ces valeurs, les colonnes C_1 et C_3 sont orthogonales donc les colonnes forment une famille orthonormée : la matrice est orthogonale.

Exercice 18

Remarquons tout d'abord que l'on doit avoir $|c| = \sqrt{a^2 + b^2}$ puisqu'une matrice orthogonale conserve la norme. Pour chacune de ces deux valeurs de c, il existe une unique rotation envoyant le vecteur $a\,e_1 + b\,e_2$ sur le vecteur $c\,e_1$ (car ceux des vecteurs ont même norme). La matrice R de cette rotation répond à la question posée.

Exercice 19

1. Cette inégalité n'est autre que l'inégalité de Cauchy-Schwarz avec les matrices $B = (|a_{i,j}|)_{1\leq i,j\leq n}$ et $J = (1)_{1\leq i,j\leq n}$ pour le produit scalaire canonique de $\mathscr{M}_n(\mathbb{R})$ car $\|B\|^2 = n$ et $\|J\|^2 = n^2$.
2. Notons $X \in \mathscr{M}_{n,1}(\mathbb{R})$ le vecteur colonne dont tous les coefficients sont égaux à 1 et remarquons que la quantité du membre de gauche s'exprime avec le produit scalaire canonique $\langle\,|\,\rangle$ sur $\mathscr{M}_{n,1}(\mathbb{R})$:

$$\sum_{1\leq i,j\leq n} a_{i,j} = X^{\mathrm{T}}(AX) = \langle X|AX\rangle.$$

L'inégalité de Cauchy-Schwarz donne alors

$$\left|\sum_{1\leq i,j\leq n} a_{i,j}\right| \leq \|X\|.\|AX\| = \|X\|^2 = n.$$

Exercice 20

Soit $(e_1,\dots,e_p)$ et $(\varepsilon_1,\dots,\varepsilon_p)$ des bases orthonormées de F et G. Complétons ces familles libres en bases de E puis orthonormalisons-les avec le procédé de Gram-Schmidt de sorte à obtenir des bases orthonormées $(e_1,\dots,e_n)$ et $(\varepsilon_1,\dots,\varepsilon_n)$ de E.

Soit $u \in \mathscr{L}(E)$ défini par

$$\forall k \in [\![1,n]\!], \qquad u(e_k) = \varepsilon_k.$$

Alors, par construction,

- $u \in \mathscr{O}(E)$ car u transforme une base orthonormée en base orthonormée ;
- $u(F) = G$ car

$$u(F) = u(\operatorname{Vect}(e_1,\dots,e_p)) = \operatorname{Vect}(u(e_1),\dots,u(e_p)) = \operatorname{Vect}(\varepsilon_1,\dots,\varepsilon_p) = G.$$

Exercice 21

Le sens direct est immédiat puisqu'un endomorphisme orthogonal préserve le produit scalaire. Réciproquement, supposons que, pour tout $(i,j)\in[\![1,p]\!]^2$, $\langle x_i|x_j\rangle=\langle y_i|y_j\rangle$.

▹ Remarquons tout d'abord que toute combinaison linéaire des vecteurs $x_1,\ldots,x_p$ a la même norme que la combinaision linéaire correspondante des $y_1,\ldots,y_p$. En effet, pour tout $\lambda_1,\ldots,\lambda_p\in\mathbb{R}$ tels que $\sum_{k=1}^p\lambda_k x_k=0_E$,

$$\begin{aligned}\Big\|\sum_{k=1}^p\lambda_k y_k\Big\|^2&=\sum_{k=1}^p\sum_{l=1}^p\lambda_k\lambda_l\langle y_k|y_l\rangle\\&=\sum_{k=1}^p\sum_{l=1}^p\lambda_k\lambda_l\langle x_k|x_l\rangle\\&=\Big\|\sum_{k=1}^p\lambda_k x_k\Big\|^2.\end{aligned}$$

Par conséquent, si un vecteur x_i est combinaison linéaire des autres vecteurs x_k pour $k\neq i$, le vecteur correspondant y_i est la même combinaison linéaire des autres vecteurs y_k pour $k\neq i$. On peut donc dorénavant se limiter au cas où les familles $(x_1,\ldots,x_p)$ et $(y_1,\ldots,y_p)$ sont libres.

▹ Complétons $(x_1,\ldots,x_p)$ en une base $(x_1,\ldots,x_p,e_{p+1},\ldots e_n)$ telle que $(e_{p+1},\ldots e_n)$ est une base orthonormée de $\mathrm{Vect}(x_1,\ldots,x_p)^\perp$ puis $(y_1,\ldots,y_p)$ en une base $(y_1,\ldots,y_p,f_{p+1},\ldots f_n)$ telle que $(f_{p+1},\ldots f_n)$ est une base orthonormée de $\mathrm{Vect}(y_1,\ldots,y_p)^\perp$.

Considérons l'unique endomorphisme u qui envoie la base $(x_1,\ldots,x_p,e_{p+1},\ldots e_n)$ sur $(y_1,\ldots,y_p,f_{p+1},\ldots f_n)$ (dans cet ordre). Montrons que u est orthogonal.

Soit $x\in E$ de décomposition

$$x=\sum_{i=1}^p\lambda_i x_i+\sum_{i=p+1}^n\lambda_i e_i.$$

Alors,

$$\begin{aligned}\|u(x)\|^2&=\Big\|\sum_{i=1}^p\lambda_i u(x_i)+\sum_{i=p+1}^n\lambda_i u(e_i)\Big\|^2\\&=\Big\|\sum_{i=1}^p\lambda_i y_i\Big\|^2+\Big\|\sum_{i=p+1}^n\lambda_i f_i\Big\|^2\\&=\Big\|\sum_{i=1}^p\lambda_i y_i\Big\|^2+\sum_{i=p+1}^n\lambda_i^2\\&=\Big\|\sum_{i=1}^p\lambda_i x_i\Big\|^2+\Big\|\sum_{i=p+1}^n\lambda_i e_i\Big\|^2=\|x\|^2.\end{aligned}$$

Ainsi, $u\in\mathcal{O}(E)$.

Corrigés des exercices d'approfondissement du chapitre 20

Exercice A

▹ Supposons que $\|x_n - x\| \to 0$. Alors, pour tout $y \in E$, l'inégalité de Cauchy-Schwarz entraîne

$$|\langle x_n - x | y\rangle| \leq \|x_n - x\|\|y\|.$$

On conclut par encadrement que $\langle x_n - x | y\rangle \to 0$.

▹ Supposons que, pour tout $y \in E$, $\langle x_n - x | y\rangle \to 0$ et considérons une base orthonormée $(e_1, \ldots, e_p)$ de E. Alors, pour tout n, l'expression des coordonnées dans cette base orthonormée donne

$$\|x_n - x\|^2 = \sum_{k=1}^{p} \langle x_n - x | e_k\rangle^2.$$

Or par hypothèse, chacun des termes du membre de droite tend vers 0 donc, par somme, $\|x_n - x\| \to 0$.

Remarque

L'implication retour n'est plus vraie en dimension infinie comme on peut le vérifier avec l'espace $\mathbb{R}[X]$ muni du produit scalaire canonique, x le polynôme nul et pour tout n, x_n le polynôme $\frac{1}{\sqrt{n+1}} \sum_{k=0}^{n} X^k$.

Exercice B

▹ Si $F = G^\perp$, alors, pour tout $x \in E$, $d(x, F)^2 = \|p_G(x)\|^2$ et $d(x, G)^2 = \|x - p_G(x)\|^2$: le résultat découle immédiatement du théorème de Pythagore.

▹ Supposons désormais que $\|x\|^2 = d(x, F)^2 + d(x, G)^2$ pour tout $x \in E$.

- Pour tout $x \in F$, $d(x, F) = 0$ donc $\|x\|^2 = d(x, G)^2 = \|x\|^2 - \|p_G(x)\|^2$ donc $p_G(x) = 0_E$ soit $x \in G^\perp$. Ainsi, $F \subset G^\perp$ (et en particulier F et G sont en somme directe).
- Pour tout $x \in (F \oplus G)^\perp$, $x \in F^\perp$ donc $d(x, F)^2 = \|x\|^2$ et, par symétrie, $d(x, G)^2 = \|x\|^2$. En reportant dans l'égalité, on obtient $\|x\|^2 = 0$ donc $x = 0_E$. Ainsi, $(F \oplus G)^\perp = \{0_E\}$ donc $F \oplus G = E$.

Exercice C

D'après la formule du rang ($\dim \operatorname{Ker} u + \dim \operatorname{Im} u = \dim E$), il suffit de montrer que les sous-espaces $\operatorname{Ker} u$ et $\operatorname{Im} u$ sont en somme directe. Considérons $x \in \operatorname{Ker} u \cap \operatorname{Im} u$ puis $t \in E$ tel que $x = u(t)$. Alors, en utilisant successivement la propriété de l'énoncé avec $x - t$, x et t,

$$\begin{aligned} 0 &= \langle u(x-t) | x - t\rangle \\ &= \underbrace{\langle u(x) | x\rangle}_{=0} + \underbrace{\langle u(x) | t\rangle}_{=0_E} + \underbrace{\langle u(t) | x\rangle}_{=\|x\|^2} + \underbrace{\langle u(t) | t\rangle}_{=0}. \end{aligned}$$

Ainsi, $\|x\|^2 = 0$ donc $x = 0_E$. En conclusion, $\operatorname{Ker} u \cap \operatorname{Im} u = \{0_E\}$ ce qui permet de conclure comme annoncé.

CORRIGÉS

Exercice D

▹ Remarquons tout d'abord que si un tel hyperplan existe alors $x-y$ est normal à H donc $\langle x-y|y\rangle = 0$ ce qui est vrai d'après l'hypothèse. Le seul hyperplan possible est donc $\mathrm{Vect}(x-y)^\perp$.

▹ Montrons que $H = \mathrm{Vect}(x-y)^\perp$ convient. La projection orthogonale de x sur $\mathrm{Vect}(x-y)$ est

$$\frac{\langle x-y|x\rangle}{\|x-y\|^2}(x-y) = x-y.$$

Par conséquent, la projection sur H est $x-(x-y) = y$.

Exercice E

L'orthogonalité de $F^\perp$ et $G^\perp$ entraîne $F^\perp \subset (G^\perp)^\perp = G$. Par conséquent, $E = F+G$.

▹ Montrons que $G = F^\perp \oplus F\cap G$. On sait déjà que $F^\perp$ et $F\cap G$ sont orthogonaux donc la somme est directe. Par ailleurs, tout vecteur $x \in G$ s'écrit $x = x_1 + x_2$ avec $x_1 \in F^\perp$ et $x_2 \in F$. Comme $F^\perp \subset G$, on a $x_2 = x - x_1 \in G$ donc $x_2 \in F\cap G$. D'où $G = F^\perp \oplus F\cap G$.

▹ Par conséquent, $E = F^\perp \oplus (F\cap G) \oplus G^\perp$ et les trois sous-espaces sont orthogonaux deux à deux. Soit $x \in E$. Il existe $x_1 \in F^\perp$, $x_2 \in F\cap G$ et $x_3 \in G^\perp$ tels que $x = x_1 + x_2 + x_3$. Alors,

$$p_F(x) = x_2 + x_3, \qquad p_G(x) = x_1 + x_2, \qquad p_{F\cap G}(x) = x_2$$

d'où $p_F(x) + p_G(x) - p_{F\cap G}(x) = x$. De même,

$$p_F \circ p_G(x) = p_F(x_1 + x_2) = x_2 = p_{F\cap G}(x), \qquad p_G \circ p_F(x) = p_G(x_2 + x_3) = x_2 = p_{F\cap G}(x),$$

d'où $p_F \circ p_G(x) = p_G \circ p_F(x) = p_{F\cap G}(x)$. Comme ces égalités sont valables pour tout $x \in E$, on en déduit les identités recherchées entre endomorphismes.

De même, $F = G^\perp \oplus F\cap G$. Par conséquent, $E = F^\perp \oplus (F\cap G) \oplus G^\perp$.

Exercice F

▹ Pour $F = \{0_E\}$, on obtient $u(E) = \{0_E\}^\perp = E$. Par conséquent, u est un endomorphisme surjectif donc un automorphisme.

▹ Soit x, y deux vecteurs orthogonaux. Alors, $y \in \mathrm{Vect}(x)^\perp$ donc $u(y) \in (u(\mathrm{Vect}(x)))^\perp$. Par conséquent, $u(x)$ et $u(y)$ sont orthogonaux.

▹ Considérons désormais une base orthonormée $(e_1, \dots, e_n)$ de E. La famille $(u(e_1), \dots, u(e_n))$ est orthogonale d'après le point précédent. Par ailleurs, pour tous $i \neq j$, $e_i + e_j$ et $e_i - e_j$ sont orthogonaux donc, d'après le premier point,

$$\langle u(e_i) + u(e_j) | u(e_i) - u(e_j)\rangle = 0$$

soit $\|u(e_i)\| = \|u(e_j)\|$.

Notons $\lambda = \|u(e_i)\|$ (qui est non nul car u injectif). Alors, pour tout $x = \sum_{i=1}^n x_i e_i$, on a

$$\|u(x)\|^2 = \Big\|\sum_{i=1}^n x_i\, u(e_i)\Big\|^2 = \sum_{i=1}^n x_i^2 \lambda^2 = \lambda^2 \|x\|^2.$$

En conclusion, $\frac{1}{\lambda} u \in \mathcal{O}(E)$ donc u est la composition de l'homothétie de rapport λ et d'un endomorphisme orthogonal.

Réciproquement, les compositions d'une homothétie et d'un endomorphisme orthogonal vérifie bien l'hypothèse de l'énoncé.

Corrigés des problèmes du chapitre 20

Problème 1

Partie A – Base ω-orthogonale échelonnée

1. La bilinéarité et la symétrie sont évidentes.
Pour tout $P \in \mathbb{R}[X]$, la quantité $(P|P)_\omega$ est l'intégrale d'une fonction positive donc est positive.
Un polynôme $P \in \mathbb{R}[X]$ vérifie $(P|P)_\omega = 0$ est nul sur un intervalle ouvert non-vide donc P est nul.
2. a. Pour prouver l'existence d'une base ω-orthogonale échelonnée, il suffit d'appliquer le procédé d'orthonormalisation de Gram-Schmidt à la base canonique : la famille obtenue est orthonormale et le degré de chaque polynôme de la nouvelle famille est égal au degré du polynôme correspondant de la base canonique.
b. Remarquons que les polynômes constants P_0 et Q_0 sont colinéaires.
Pour tout $n \geq 0$, on remarque que P_{n+1} et Q_{n+1} sont tous les deux dans le supplémentaire orthogonal de $\mathbb{R}_n[X] = \mathrm{Vect}(P_k)_{k\leq n} = \mathrm{Vect}(Q_k)_{k\leq n}$ dans $\mathbb{R}_{n+1}[X]$ qui est de dimension 1 donc sont colinéaires.
3. Soit Q le polynôme unitaire scindé à racines simples dont les racines sont celles de P_n d'ordre impair dans l'intervalle I. Si $\deg Q < n$, alors $(P_n|Q)_\omega = 0$ or la fonction $t \mapsto P_n(t)Q(t)\omega(t)$ est de signe constant sur I, continue par morceaux et non-nulle sur un intervalle ouvert non-vide : contradiction. Par conséquent, $\deg Q = n$ et P_n admet n racines (simples) dans l'intervalle I.
4. a. Comme (P_0, P_1) est une base de $\mathbb{R}_1[X]$ et $XP_0 \in \mathbb{R}_1[X]$, il existe un unique couple de coordonnées $(\alpha_0, \beta_0) \in \mathbb{R}^2$ tel que

$$XP_0 = \alpha_0 P_1 + \beta_0 P_0.$$

b. Pour tout $Q \in \mathbb{R}_{n-2}[X]$, $(XP_n|Q)_\omega = (P_n|XQ)_\omega = 0$ car P_n est orthogonal à $\mathbb{R}_{n-1}[X] = \mathrm{Vect}(P_k)_{k\leq n-1}$.
c. Le polynôme XP_n s'exprime de manière unique sur la base orthogonale $(P_k)_{k\leq n+1}$ de $\mathbb{R}_{n+1}[X]$. Or, d'après la question précédente, les coordonnées sur les vecteurs $(P_k)_{k\leq n-2}$ sont nulles ; ainsi, il existe un unique triplet de coordonnées $(\alpha_n, \beta_n, \gamma_n) \in \mathbb{R}^3$ tel que

$$XP_n = \alpha_n P_{n+1} + \beta_n P_n + \gamma_n P_{n-1}.$$

d. En étudiant les coefficients de degré $n+1$ et de degré n dans l'équation de la question précédente, on obtient

$$\begin{aligned} a_n &= \alpha_n a_{n+1} \\ b_n &= \alpha_n b_{n+1} + \beta_n a_n \end{aligned}$$

ce qui donne les deux premières relations. Pour la troisième, effectuons le produit scalaire $(XP_n|P_{n-1})_\omega = (P_n|XP_{n-1})_\omega$ qui vaut $\gamma_n \|P_{n-1}\|_\omega^2$ d'après la question précédente au rang n et $\alpha_n \|P_n\|_\omega^2$ d'après la question précédente au rang $n-1$: on en déduit la troisième relation.

Partie B – Formule de quadrature

1. Les formes linéaires $Q \mapsto Q\left(\lambda_k^{(n)}\right)$ forment une base de l'espace dual de $\mathbb{R}_{n-1}[X]$ (c'est notamment la base duale de la base des polynômes d'interpolation de Lagrange associés aux $\left(\lambda_k^{(n)}\right)_k$). Par conséquent, la forme linéaire $Q \mapsto \int_I Q(t)\omega(t)\,\mathrm{d}t$ s'exprime sur cette base et il existe $c_1, \ldots, c_n \in \mathbb{R}$ tels que

$$\forall Q \in \mathbb{R}_{n-1}[X], \qquad \int_I Q(t)\omega(t)\,\mathrm{d}t = \sum_{k=1}^{n} c_k Q\left(\lambda_k^{(n)}\right).$$

2. Pour tout $Q \in \mathbb{R}_{2n-1}[X]$, il existe A, B tels que $Q = AP_n + B$ et $\deg B \leq n-1$. Par conséquent,

$$\begin{aligned}\int_I Q(t)\omega(t)\,\mathrm{d}t &= (A|P_n)_\omega + \int_I B(t)\omega(t)\,\mathrm{d}t \\ &= \sum_{k=1}^{n} c_k B\big(\lambda_k^{(n)}\big) = \sum_{k=1}^{n} c_k Q\big(\lambda_k^{(n)}\big)\end{aligned}$$

en utilisant la question précédente et en remarquant que

$$Q\big(\lambda_k^{(n)}\big) = A\big(\lambda_k^{(n)}\big)P_n\big(\lambda_k^{(n)}\big) + B\big(\lambda_k^{(n)}\big) = B\big(\lambda_k^{(n)}\big).$$

3. Appliquons la relation précédente avec $Q = L_i^2$ le carré d'un polynôme d'interpolation de Lagrange associés aux $\big(\lambda_k^{(n)}\big)_k$. Alors,

$$c_i = \int_I L_i^2(t)\omega(t)\,\mathrm{d}t \geq 0.$$

4. a. L'application linéaire

$$\left\{\begin{array}{ccc} \mathbb{R}_{2n-1}[X] & \to & \mathbb{R}^{2n} \\ P & \mapsto & \big(P\big(\alpha_1^{(n)}\big), P'\big(\alpha_1^{(n)}\big), \ldots, P\big(\alpha_n^{(n)}\big), P'\big(\alpha_n^{(n)}\big)\big) \end{array}\right.$$

est un isomorphisme (son noyau ne contient que le polynôme nul, seul polynôme de degré au plus $2n-1$ à admettre n racines doubles et les deux espaces ont la même dimension). Par conséquent, il existe un unique polynôme $H \in \mathbb{R}_{2n-1}[X]$ tel que, pour tout $k \leq n$,

$$H\big(\alpha_k^{(n)}\big) = f\big(\alpha_k^{(n)}\big), \qquad H'\big(\alpha_k^{(n)}\big) = f'\big(\alpha_k^{(n)}\big).$$

b. Par linéarité, $\mathscr{E}(f-H) = \mathscr{E}(f) - \mathscr{E}(H)$. Or $\mathscr{E}(H) = 0$ d'après la question 2, d'où le résultat.

c. Notons $g = f - H - \lambda P_n^2$. Cette fonction de classe $\mathscr{C}^1$ s'annule en x et $\big(\alpha_k^{(n)}\big)_{k\leq n}$: d'après le théorème de Rolle, g' admet n zéros sur les intervalles ouverts définis par ces points. Or, g' s'annule aussi en $\big(\alpha_k^{(n)}\big)_{k\leq n}$ donc g' admet $2n$ zéros deux à deux distincts. Des applications successives du théorème de Rolle montrent que $g^{(2n)} = f^{(2n)} - \lambda(2n)!d_n^2$ admet un zéro c_x. Donc

$$\lambda = \frac{f^{(2n)}(c_x)}{(2n)!d_n^2}.$$

d. D'après les questions précédentes,

$$\mathscr{E}(f) = \mathscr{E}(f-H) = \int_I \frac{f^{(2n)}(c_x)}{(2n)!d_n^2} P_n(x)^2\omega(x)\,\mathrm{d}x\,.$$

Par conséquent,

$$|\mathscr{E}(f)| \leq \frac{1}{d_n^2(2n)!}\|f^{(2n)}\|_\infty \int_I P_n(x)^2\omega(x)\,\mathrm{d}x\,.$$

Chapitre 21

Calcul intégral

1. Intégrale sur un segment — **2.** Sommes de Riemann — **3.** Intégration et dérivation — **4.** Formules de Taylor

Objectifs et compétences du programme

- Connaître les règles de manipulations des intégrales.
- Comprendre le lien entre dérivation et primitivation.
- Reconnaître les sommes de Riemann et préciser leurs limites.
- Calculer des intégrales de fonctions élémentaires.

Bernhard Riemann, mathématicien allemand (1826-1866)
Difficile de résumer l'immense apport de Riemann par rapport à la brièveté de son existence : fondateur de la théorie de l'intégration (à la suite de Cauchy), de l'étude des fonctions à variables complexes, de la géométrie différentielle (les célèbres variétés riemanniennes)... il a aussi étudié les équations aux dérivées partielles.
Son plus célèbre article date de 1859 « sur le nombre de nombres premiers inférieurs à une taille donnée » : il y reprend l'étude de la fonction ζ (entamée par Euler) avec un argument complexe ce qui lui permet d'énoncer l'« hypothèse de Riemann », l'une des conjectures de la liste des problèmes du millénaire.

COURS

Dans ce chapitre, on construit l'intégrale d'une fonction continue par morceaux sur un segment à partir de l'approximation par des fonctions en escalier. Une fois cette construction achevée, on retrouve toutes les propriétés opératoires de l'intégrale et on explore le lien entre intégration et primitivation. Parmi les sous-produits de cette étude, on obtient notamment la convergence des sommes de Riemann (qui complète l'étude de suites numériques) et les formules de Taylor globales ou locales (qui complètent l'étude des fonctions plusieurs fois dérivables).

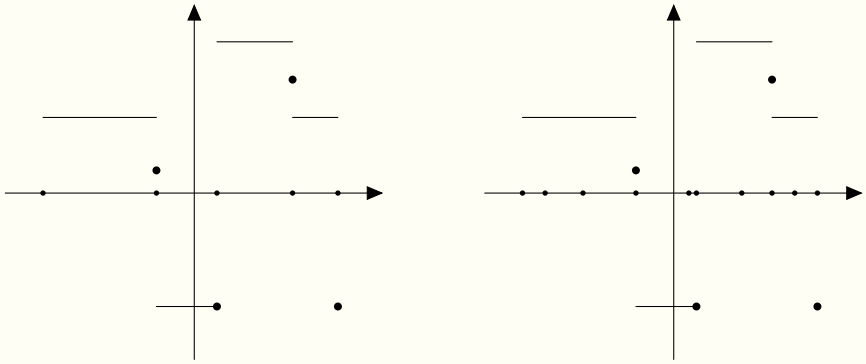
Compléments sur les fonctions en escalier

Définition 21.1.

Une fonction f est en escalier sur l'intervalle $[a,b]$ s'il existe une subdivision $(x_i)_{i\in[0,N]}$ telle que les restrictions de f aux intervalles $]x_{i-1},x_i[$ pour tout $i\in[\![1,N]\!]$ soient constantes.

Remarquons que si une fonction est en escalier pour une subdivision donnée, elle le sera encore pour une subdivision plus « fine » c'est-à-dire à laquelle on a ajouté des points.

Voici, par exemple, la courbe représentative d'une fonction en escalier avec deux subdivisions adaptées (marquées par des points sur l'axe des abscisses) :



En fait, on avait établi le résultat suivant à partir de la continuité uniforme (et donc du théorème de Heine).

Proposition 21.2.

Si f est une fonction continue par morceaux sur le segment I et si $\varepsilon>0$, alors il existe φ et ψ deux fonctions en escalier sur I telles que $\varphi\leq f\leq\psi$ et $\psi-\varphi\leq\varepsilon$.

1. Intégrale sur un segment

1.1. Construction

Définition 21.3.

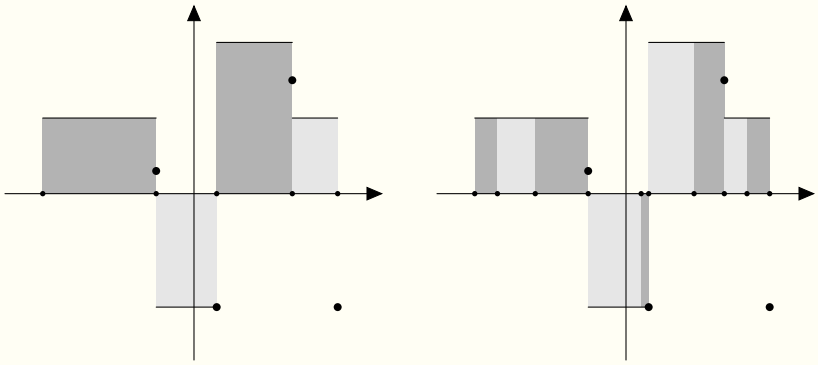
Soit φ une fonction en escalier sur $[a,b]$, $(x_i)_{i\in[0,N]}$ une subdivision adaptée à φ et $(\varphi_i)_{i\in[1,N]}$ les valeurs prises par φ sur chacun des intervalles ouverts associés à cette subdivision.
L'intégrale de φ sur $[a,b]$ est le réel $I(\varphi)$ défini par

$$I(\varphi)=\sum_{i=1}^{N}(x_i-x_{i-1})\varphi_i.$$

Remarque

▹ Le réel $I(\varphi)$ est la somme algébrique des aires des rectangles délimités par la courbe représentative de φ et l'axe des abscisses : on compte en positif l'aire des rectangles au-dessus de l'axe des abscisses, en négatif celles des autres.

▹ Si on choisit une subdivision de $[a,b]$ plus fine que $(x_i)_{i\in[0,N]}$ (c'est-à-dire avec davantage de points), on trouve la même intégrale en rassemblant des termes dans la somme. En effet, ajouter un point revient à partitionner un rectangle en deux rectangles sans que l'aire totale ne varie.



Par transitivité, l'intégrale de φ ne dépend pas de la subdivision choisie tant qu'elle reste adaptée à φ .

Proposition 21.4.

Sur l'espace des fonctions en escalier $\mathscr{E}$, l'application $I:\varphi\mapsto I(\varphi)$ vérifie les propriétés suivantes

- (linéarité) $\forall\varphi,\psi\in\mathscr{E},\ \forall\lambda,\mu\in\mathbb{R},\qquad I(\lambda\varphi+\mu\psi)=\lambda I(\varphi)+\mu I(\psi)$;
- (croissance) $\forall\varphi,\psi\in\mathscr{E},\qquad \varphi\leq\psi\quad\Rightarrow\quad I(\varphi)\leq I(\psi)$;
- (inégalité en valeur absolue) $\forall\varphi\in\mathscr{E},\qquad |I(\varphi)|\leq I(|\varphi|)$.

La démonstration du premier point requiert l'utilisation d'une subdivision adaptée aux deux fonctions en escalier. Le troisième point repose sur l'inégalité triangulaire.

Proposition 21.5.

Soit f une fonction continue par morceaux sur le segment $[a,b]$. Alors, les quantités

$$I_+(f) = \inf\{I(\psi),\ \text{avec } \psi \geq f \text{ et } \psi \in \mathscr{E}\}$$
$$I_-(f) = \sup\{I(\varphi),\ \text{avec } \varphi \leq f \text{ et } \varphi \in \mathscr{E}\}$$

sont finies et égales. Cette valeur commune est alors appelée intégrale de f et notée

$$\int_a^b f(t)\,\mathrm{d}t,\ \int_a^b f \text{ ou } \int_{[a,b]} f.$$

1.2. Propriétés de l'intégrale

Proposition 21.6.

L'application $f \mapsto \int_a^b f(t)\,\mathrm{d}t$ est linéaire de l'espace des fonctions continues par morceaux dans $\mathbb{R}$.

Démonstration

Montrons le résultat pour la somme (la multiplication par un scalaire se traite de même).

▹ Soit f et g continues par morceaux, φ et ϕ des fonctions en escalier inférieures à f et g respectivement. Alors, $\varphi+\phi$ est une fonction en escalier inférieure à $f+g$. Par conséquent,

$$I(\varphi+\phi) \leq \int_a^b (f(t)+g(t))\,\mathrm{d}t\,.$$

Comme $I(\varphi)+I(\phi) = I(\varphi+\phi)$, on obtient

$$I(\varphi)+I(\phi) \leq \int_a^b (f(t)+g(t))\,\mathrm{d}t\,.$$

En passant à la borne supérieure successivement pour φ et ϕ, on obtient

$$\int_a^b f(t)\,\mathrm{d}t + \int_a^b g(t)\,\mathrm{d}t \leq \int_a^b (f(t)+g(t))\,\mathrm{d}t\,.$$

▹ On obtient en utilisant de la même manière des fonctions en escalier supérieures à f et g :

$$\int_a^b f(t)\,\mathrm{d}t + \int_a^b g(t)\,\mathrm{d}t \geq \int_a^b (f(t)+g(t))\,\mathrm{d}t\,.$$

On conclut par double-inégalité. ■

Proposition 21.7.

Soit f et g deux fonctions continues par morceaux sur $[a,b]$ telles que $f \leq g$. Alors,

$$\int_a^b f(t)\,\mathrm{d}t \leq \int_a^b g(t)\,\mathrm{d}t\,.$$

Démonstration
Toute fonction en escalier φ inférieure à f est inférieure à g donc $I(\varphi) \leq \int_a^b g(t)\,\mathrm{d}t$. Ainsi, l'intégrale $\int_a^b g(t)\,\mathrm{d}t$ est un majorant de $\{I(\varphi)$, avec $\varphi \leq f$ et $\varphi \in \mathscr{E}\}$ donc est supérieur ou égal à son plus petit majorant à savoir $\int_a^b f(t)\,\mathrm{d}t$. ■

Remarque
En particulier, si $f \geq 0$ est continue par morceaux sur $[a,b]$, $\int_a^b f(t)\,\mathrm{d}t \geq 0$.

Attention toutefois, on peut avoir $f \geq 0$ et $f \neq 0$ malgré $\int_a^b f(t)\,\mathrm{d}t = 0$: par exemple, pour une fonction nulle partout sauf en un point du segment $[a,b]$.

Avec l'hypothèse supplémentaire de continuité, on obtient le résultat suivant.

Proposition 21.8.

Soit f une fonction continue sur $[a,b]$, positive et non nulle en au moins un point de $[a,b]$. Alors

$$\int_a^b f(t)\,\mathrm{d}t > 0.$$

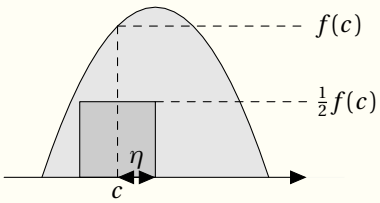
Conseils méthodologiques

On peut utiliser cette proposition (ou plutôt sa contraposée) pour établir la nullité d'une fonction : il suffit de dire que la fonction est continue, positive et d'intégrale nulle.

Démonstration
Supposons qu'il existe $c \in]a,b[$ tel que $f(c) > 0$; par continuité, il existe $\eta > 0$ tel que

$$\forall x \in A = [a,b] \cap]c-\eta, c+\eta[, \qquad f(x) \geq \frac{1}{2} f(c).$$

Par croissance de l'intégrale, $\int_a^b f(t)\,\mathrm{d}t \geq \frac{1}{2} f(c)|A| > 0$ où $|A|$ désigne la longueur de A.



■

Exemple

Cherchons les polynômes P vérifiant $\int_0^1 x(1-x)P(x)^2\,\mathrm{d}x = 0$.
La fonction $f : x \mapsto x(1-x)P(x)^2$ est continue, positive. S'il existe un réel $c \in]0,1[$ tel que $P(c)$ est non nul alors $f(c)$ est aussi non nulle et l'intégrale est strictement positive. Par conséquent, les polynômes recherchés sont nuls sur $]0,1[$ et donc sur $\mathbb{R}$.

Par ailleurs, la propriété de croissance appliquée à l'encadrement $-|f(x)| \leq f(x) \leq |f(x)|$ implique l'inégalité suivante.

Proposition 21.9.

Soit f une fonction continue par morceaux sur $[a,b]$. Alors, $\left|\int_a^b f(t)\,\mathrm{d}t\right| \leq \int_a^b |f(t)|\,\mathrm{d}t$.

Remarque

Montrons que le cas d'égalité pour une fonction continue f n'est possible que si f est de signe constant. Notons $\epsilon \in \{\pm 1\}$ le signe de $\int_a^b f(t)\,\mathrm{d}t$ avec la convention $\epsilon = 1$ si cette intégrale est nulle.

Si $\left|\int_a^b f(t)\,\mathrm{d}t\right| = \int_a^b |f(t)|\,\mathrm{d}t$, alors

$$\int_a^b (|f(t)| - \epsilon f(t))\,\mathrm{d}t = 0.$$

La fonction $t \mapsto |f(t)| - \epsilon f(t)$ est continue, positive et d'intégrale nulle donc est identiquement nulle, c'est-à-dire $|f| = \epsilon.f$.

Terminons avec une autre inégalité classique sur les intégrales.

Proposition 21.10. Inégalité de Cauchy-Schwarz

Soit f et g deux fonctions continues par morceaux sur $[a,b]$. Alors,

$$\left|\int_a^b f(x)g(x)\,\mathrm{d}x\right| \leq \sqrt{\int_a^b f(x)^2\,\mathrm{d}x \int_a^b g(x)^2\,\mathrm{d}x}.$$

Démonstration

La fonction $P : \lambda \mapsto \int_a^b (f(x) + \lambda g(x))^2\,\mathrm{d}x$ est polynomiale et positive.

▹ Si $\int_a^b g(x)^2\,\mathrm{d}x = 0$, alors le polynôme P est de degré au plus 1 et il est positif si, et seulement si,

$$\int_a^b f(x)g(x)\,\mathrm{d}x = 0$$

l'inégalité est vérifiée.

▹ Sinon, le polynôme P est de degré 2, de coefficient dominant positif. P est positif si, et seulement si, son discriminant est négatif ou nul :

$$\left(2\int_a^b f(x)g(x)\,\mathrm{d}x\right)^2 - 4\int_a^b f(x)^2\,\mathrm{d}x \int_a^b g(x)^2\,\mathrm{d}x \leq 0,$$

d'où l'inégalité. ■

Remarque

Supposons que f et g sont continues. En conservant les notations de la démonstration, l'inégalité est alors une égalité si P admet une racine réelle λ_0. La fonction $(f + \lambda_0 g)^2$ est alors continue, positive d'intégrale nulle donc elle est nulle et les fonctions f, g forment une famille liée.

Notation

Soit f une fonction continue par morceaux sur $[a,b]$. On pose

$$\int_b^a f(t)\,\mathrm{d}t = -\int_a^b f(t)\,\mathrm{d}t, \qquad \int_a^a f(t)\,\mathrm{d}t = 0.$$

Remarque

On veillera à remarquer que les inégalités énoncées précédemment supposaient implicitement que $a < b$ et que permuter l'ordre des bornes change les signes donc les inégalités.

Par exemple, pour f et g deux fonctions continues par morceaux telles que $f \leq g$, on obtient

$$\int_a^b f(t)\,\mathrm{d}t \geq \int_a^b g(t)\,\mathrm{d}t$$

lorsque $b < a$.

Avec cette convention, on obtient alors la relation de Chasles.

Proposition 21.11. Relation de Chasles

Soit f continue par morceaux sur I, a, b et $c \in I$. Alors,

$$\int_a^b f(t)\,\mathrm{d}t = \int_a^c f(t)\,\mathrm{d}t + \int_c^b f(t)\,\mathrm{d}t\,.$$

Démonstration

Il suffit d'établir le résultat pour $a \leq c \leq b$ (car en changeant les signes des intégrales avec les bornes mal orientées, on se ramène à cette situation). Or, dans ce cas, il suffit d'ajouter le point c aux subdivisions associées aux fonctions en escalier de la définition. ■

1.3. Généralisation aux fonctions complexes

Pour conserver la linéarité de l'opérateur d'intégration, on adopte la définition suivante.

Définition 21.12.

Soit f une fonction continue par morceau de $[a, b]$ dans $\mathbb{C}$. L'intégrale de f est alors le complexe

$$\int_a^b f(t)\,\mathrm{d}t = \int_a^b \mathfrak{Re}(f)(t)\,\mathrm{d}t + i\int_a^b \mathfrak{Im}(f)(t)\,\mathrm{d}t\,.$$

Ainsi, on sait intégrer les fonctions à valeurs complexes mais avec un argument réel.

On peut généraliser les propriétés précédentes.

Proposition 21.13.

Soit f, g deux fonctions continues par morceaux sur $[a, b]$ et $\lambda, \mu \in \mathbb{C}$.

Linéarité

$$\int_a^b (\lambda f + \mu g)(t)\,\mathrm{d}t = \lambda \int_a^b f(t)\,\mathrm{d}t + \mu \int_a^b g(t)\,\mathrm{d}t\,.$$

Inégalité en module

$$\left|\int_a^b f(t)\,\mathrm{d}t\right| \leq \int_a^b |f(t)|\,\mathrm{d}t\,.$$

Inégalité de Cauchy-Schwarz

$$\left|\int_a^b f(t)\overline{g(t)}\,\mathrm{d}t\right| \leq \sqrt{\int_a^b |f(t)|^2\,\mathrm{d}t \int_a^b |g(t)|^2\,\mathrm{d}t}.$$

Relation de Chasles

$$\forall c \in]a,b[, \quad \int_a^b f(t)\,\mathrm{d}t = \int_a^c f(t)\,\mathrm{d}t + \int_c^b f(t)\,\mathrm{d}t\,.$$

Montrons uniquement l'inégalité de Cauchy-Schwarz.

Démonstration

Considérons un réel θ tel que

$$\int_a^b f(t)\overline{g(t)}\,\mathrm{d}t = e^{i\theta}\left|\int_a^b f(t)\overline{g(t)}\,\mathrm{d}t\right|.$$

Alors, pour tout $\lambda \in \mathbb{R}$,

$$\int_a^b |f(t)+\lambda e^{i\theta} g(t)|^2\,\mathrm{d}t = \lambda^2 \int_a^b |g(t)|^2\,\mathrm{d}t + 2\lambda\left|\int_a^b f(t)\overline{g(t)}\,\mathrm{d}t\right| + \int_a^b |f(t)|^2\,\mathrm{d}t\,.$$

Comme dans le cas réel, on remarque que le polynôme (réel) en λ du second membre est toujours positif donc de discriminant négatif. ■

2. Sommes de Riemann

2.1. Cas d'une fonction continue

Définition 21.14.

Soit f une fonction définie sur $[a,b]$, $(x_k)_{k\in[0,N]}$ une subdivision de $[a,b]$ et $(c_k)_{k\in[0,N-1]}$ une suite de réels telle que $c_k \in [x_k, x_{k+1}]$ pour tout $k \in [\![0, N-1]\!]$. La somme de Riemann associée à ces paramètres est le scalaire

$$R(f) = \sum_{k=0}^{N-1} (x_{k+1}-x_k) f(c_k).$$

Remarque

La quantité $R(f)$ est l'intégrale d'une fonction en escalier définie à partir de f donc est intuitivement une approximation de $\int_a^b f(t)\,\mathrm{d}t$.

Proposition 21.15.

Soit f continue sur $[a,b]$ et $\varepsilon > 0$. Alors, il existe $\eta > 0$ tel que la condition

$$\max\{|x_{k+1}-x_k|,\ k \in [\![0, N-1]\!]\} \leq \eta$$

entraîne

$$\left|\int_a^b f(t)\,\mathrm{d}t - R(f)\right| \leq \varepsilon.$$

En d'autres termes, les sommes de Riemann d'une fonction continue tendent vers l'intégrale quand le pas de la subdivision tend vers 0.

Démonstration

Notons tout d'abord qu'avec la relation de Chasles,

$$\int_a^b f(t)\,\mathrm{d}t = \sum_{k=0}^{N-1} \int_{x_k}^{x_{k+1}} f(t)\,\mathrm{d}t$$

$$R(f) = \sum_{k=0}^{N-1} \int_{x_k}^{x_{k+1}} f(c_k)\,\mathrm{d}t$$

Par conséquent,

$$\left|\int_a^b f(t)\,\mathrm{d}t - R(f)\right| = \left|\sum_{k=0}^{N-1} \int_{x_k}^{x_{k+1}} (f(t)-f(c_k))\,\mathrm{d}t\right|$$

$$\leq \sum_{k=0}^{N-1} \int_{x_k}^{x_{k+1}} |f(t)-f(c_k)|\,\mathrm{d}t\,.$$

Or, f est continue donc uniformément continue sur $[a,b]$ (théorème de Heine) : il existe $\eta > 0$ tel que

$$\forall x, y \in [a,b], \qquad |x-y| \leq \eta \quad \Rightarrow \quad |f(x)-f(y)| \leq \varepsilon.$$

Par conséquent, si $\max\{|x_{k+1}-x_k|,\ k \in [\![0, N-1]\!]\} \leq \eta$, alors

$$\left|\int_a^b f(t)\,\mathrm{d}t - R(f)\right| \leq \varepsilon \sum_{k=0}^{N-1} (x_{k+1}-x_k) = \varepsilon(b-a).$$

■

Précisons quelques méthodes d'approximation d'intégrale qui consistent à calculer des sommes de Riemann.

Définition 21.16.

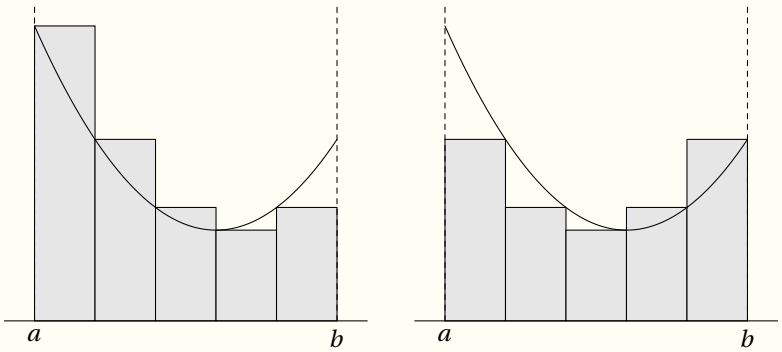
Soit $N \geq 2$. La méthode des rectangles consiste à calculer la somme de Riemann pour une subdivision à pas fixe dont les points sont donc

$$x_k = a + k\frac{b-a}{N} \text{ pour } k \in [\![0, N]\!]$$

et avec les valeurs

- $c_k = a + k\frac{b-a}{N}$ pour $k \in [\![0, N-1]\!]$ pour les rectangles à gauche,
- $c_k = a + (k+1)\frac{b-a}{N}$ pour $k \in [\![0, N-1]\!]$ pour les rectangles à droite.

Voici des représentations des rectangles à gauche puis à droite sur la courbe d'une fonction polynomiale de degré 2.



Réécrivons le résultat d'approximation pour la méthode des rectangles à gauche et à droite.

Corollaire 21.17.

Soit f continue sur $[a,b]$. Alors,

$$\lim_N \frac{b-a}{N}\sum_{k=0}^{N-1} f\left(a+k\frac{b-a}{N}\right)=\int_a^b f(x)\,\mathrm{d}x,$$
$$\lim_N \frac{b-a}{N}\sum_{k=1}^{N} f\left(a+k\frac{b-a}{N}\right)=\int_a^b f(x)\,\mathrm{d}x.$$

Démonstration

Soit $\varepsilon>0$ et $\eta>0$ qui lui est associé par la proposition précédente. Alors, pour tout N tel que $\frac{b-a}{N}\leq\eta$,

$$\left|\int_a^b f(x)\,\mathrm{d}x-\frac{b-a}{N}\sum_{k=0}^{N-1} f\left(a+k\frac{b-a}{N}\right)\right|\leq\varepsilon.$$

■

Exemple

Voici quelques illustrations simples de cette convergence.

▹ En considérant la fonction $x\mapsto x^2$,

$$\lim_{N\to\infty}\frac{1}{N}\sum_{k=0}^{N-1}\left(\frac{k}{N}\right)^2=\int_0^1 x^2\,\mathrm{d}x=\frac{1}{3}.$$

▹ En considérant la fonction $x\mapsto x(1-x)$,

$$\begin{aligned}\lim_{N\to\infty}\frac{1}{N^3}\sum_{k=0}^{N-1}k(N-k)&=\lim_{N\to\infty}\frac{1}{N}\sum_{k=0}^{N-1}\frac{k}{N}\left(1-\frac{k}{N}\right)\\&=\int_0^1 x(1-x)\,\mathrm{d}x=\frac{1}{6}.\end{aligned}$$

Exemple

Soit $b>a$. Calculons $\int_a^b e^x\,\mathrm{d}x$ avec les sommes de Riemann.

$$\begin{aligned}\int_a^b e^x\,\mathrm{d}x&=\lim_{N\to\infty}\frac{b-a}{N}\sum_{k=0}^{N-1}e^{a+k\frac{b-a}{N}}=e^a\lim_{N\to\infty}\frac{b-a}{N}\sum_{k=0}^{N-1}(e^{\frac{b-a}{N}})^k\\&=e^a\lim_{N\to\infty}\frac{b-a}{N}\frac{1-e^{b-a}}{1-e^{\frac{b-a}{N}}}=e^b-e^a.\end{aligned}$$

Exemple

Calculons pour $|a|>1$, $\int_0^\pi \ln|a^2-2a\cos x+1|\,\mathrm{d}x$.

$$\begin{aligned}\int_0^\pi \ln|a^2-2a\cos x+1|\,\mathrm{d}x &= \lim_{N\to\infty}\frac{\pi}{N}\sum_{k=1}^{N}\ln|a^2-2a\cos\frac{k\pi}{N}+1|\\ &= \lim_{N\to\infty}\frac{\pi}{N}\ln|\prod_{k=1}^{N}(a^2-2a\cos\frac{k\pi}{N}+1)|\\ &= \lim_{N\to\infty}\frac{\pi}{N}\ln(|a+1|^2\prod_{k=1}^{N-1}|a^2-2a\cos\frac{k\pi}{N}+1|)\\ &= \lim_{N\to\infty}\frac{\pi}{N}\ln(|a+1|^2\frac{|a^{2N}-1|}{|a^2-1|})\\ &= \lim_{N\to\infty}\frac{\pi}{N}\left(\ln\frac{|a+1|^2}{|a^2-1|}+\ln|1-a^{-2N}|+2N\ln|a|\right)\\ &= 2\pi\ln|a|.\end{aligned}$$

Remarquons que le même calcul donne, pour $|a|<1$ (donc $a^{2N}\to 0$),

$$\int_0^\pi \ln|a^2-2a\cos x+1|\,\mathrm{d}x = \lim_{N\to\infty}\frac{\pi}{N}\left(\ln\frac{|a+1|^2}{|a^2-1|}+\ln|a^{2N}-1|\right)=0.$$

2.2. Cas des fonctions lipschitziennes

Les résultats précédents sont intéressants mais inutiles en pratique car on n'a pas quantifié l'erreur. En revanche, on peut obtenir facilement une majoration de l'erreur dans le cas d'une fonction f K-lipschitzienne ; en effet,

$$\begin{aligned}\left|\int_a^b f(t)\,\mathrm{d}t-R(f)\right| &\le K\sum_{k=0}^{N-1}\int_{x_k}^{x_{k+1}}|x-c_k|\,\mathrm{d}x\\ &\le \frac{K}{2}\sum_{k=0}^{N-1}\left(\left[-(c_k-x)^2\right]_{x_k}^{c_k}+\left[(x-c_k)^2\right]_{c_k}^{x_{k+1}}\right)\\ &\le \frac{K}{2}\sum_{k=0}^{N-1}\left((x_{k+1}-c_k)^2+(c_k-x_k)^2\right).\end{aligned}$$

Proposition 21.18.

Si f est K-lipschitzienne et $R(f)$ est la somme de Riemann obtenue par la méthode des rectangles (à gauche ou à droite), alors

$$\left|\int_a^b f(t)\,\mathrm{d}t-R(f)\right| \le \frac{K(b-a)^2}{2N}.$$

Remarque

Le résultat précédent est énoncé pour les fonctions lipschitziennes sur un segment. Il est important de noter que c'est évidemment le cas pour toutes les fonctions de classe $\mathscr{C}^1$ sur un segment d'après l'inégalité des accroissements finis.

3. Intégration et dérivation

3.1. Primitives

Définition 21.19.

Une primitive d'une fonction f est une fonction F dérivable telle que $F' = f$.

Proposition 21.20.

Deux primitives d'une fonction définies sur un même intervalle ne diffèrent que d'une constante.

Remarque

Le résultat est faux si l'on ne précise pas que l'on travaille sur un intervalle puisque $y' = 0$ admet des solutions non constantes sur une réunion d'intervalles disjoints.

Exemples

▹ Les primitives de $x \mapsto \frac{1}{x}$ sont, sur $\mathbb{R}_+^*$ ou $\mathbb{R}_-^*$, $x \mapsto \ln|x| + C$ avec $C \in \mathbb{R}$.
▹ Les primitives de $x \mapsto \frac{1}{\sqrt{1-x^2}}$ sont $x \mapsto \arcsin x + C$ avec $C \in \mathbb{R}$.
▹ Plus généralement, avec les fonctions réciproques, on obtient, à constante près, les primitives suivantes (pour $a > 0$).

$f(x)$	$F(x)$
$\dfrac{1}{a^2+x^2}$	$\dfrac{1}{a}\arctan\left(\dfrac{x}{a}\right)$
$\dfrac{1}{\sqrt{a^2-x^2}}$	$\arcsin\left(\dfrac{x}{a}\right)$

Conseils méthodologiques

Dans les calculs de primitives simples, il est important de reconnaître les fonctions de la forme $u'u$, $\frac{u'}{u}$ et plus généralement de la forme $u'u^n$ avec $n \in \mathbb{Z}$.

Exemple

La méthode précédente avec $u = \ln(x)$ permet d'exhiber, toujours à constante près, les primitives suivantes (pour $n \geq 2$).

$f(x)$	$F(x)$
$\frac{1}{x}\ln(x)$	$\frac{1}{2}(\ln(x))^2$
$\frac{1}{x\ln(x)}$	$\ln\lvert\ln(x)\rvert$
$\frac{1}{x\ln(x)^n}$	$\frac{1}{1-n}\cdot\frac{1}{(\ln(x))^{n-1}}$

3.2. Intégrale fonction des bornes

Proposition 21.21.

Soit f continue par morceaux sur l'intervalle I et $a \in I$. Posons

$$F : x \mapsto \int_a^x f(t)\,\mathrm{d}t\,.$$

▹ La fonction F est continue.
▹ La fonction F est dérivable à droite et à gauche sur I et ses dérivées sont

$$F_g'(x) = \lim_{t\to x^-} f(t), \qquad F_d'(x) = \lim_{t\to x^+} f(t).$$

▹ La fonction F est dérivable en tout point de continuité de f.

Démonstration

▹ Remarquons tout d'abord que f est bornée. Ainsi, l'inégalité de la moyenne donne pour tout $x \in I$ et tout $h \in \mathbb{R}$ tel que $x+h \in I$:

$$|F(x+h) - F(x)| \leq \|f\|_\infty |h|.$$

Ainsi, $\lim\limits_{h\to 0} F(x+h) = F(x)$ et F est continue en x.

▹ Utilisons désormais le fait que f admet une limite à gauche et à droite en tout point de I. Par exemple, considérons $x \in I$ et notons ℓ la limite à gauche de f en x ; pour tout $\varepsilon > 0$, il existe $\eta > 0$ tel que

$$\forall t \in I \cap]x-\eta, x], \qquad |f(t) - \ell| \leq \varepsilon.$$

Ainsi, pour tout $y \in I \cap]x-\eta, x]$,

$$\begin{aligned} |F(x) - F(y) - \ell(x-y)| &= \left| \int_y^x f(t)\,\mathrm{d}t - \int_y^x \ell\,\mathrm{d}t \right| \\ &\leq \int_y^x |f(t) - \ell|\,\mathrm{d}t \leq (x-y)\varepsilon. \end{aligned}$$

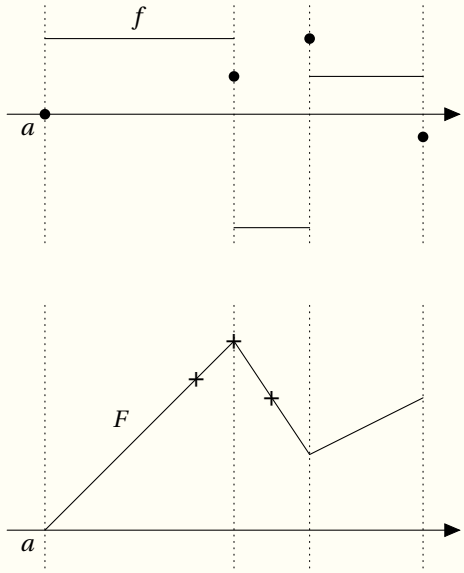
Par conséquent, le taux d'accroissement de F admet pour limite à gauche le réel ℓ : F est dérivable à gauche.

▹ Pour la dérivabilité en tout point de continuité en f, on remplace les définitions de limite à gauche et à droite par la continuité de f. ■

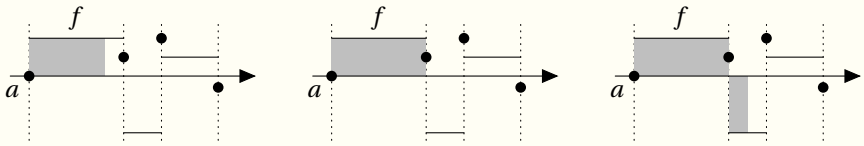
Remarque

Illustrons ce résultat.

Traçons les courbes représentatives d'une fonction en escalier f puis de la fonction $F : x \mapsto \int_a^x f(t)\,\mathrm{d}t$.



Les trois points marqués d'une croix sur la courbe représentative de F indiquent la valeur des aires sous la courbe représentative de f (entre a et l'abscisse du point considéré) telles qu'elles sont indiquées sur le schéma suivant



On vérifie bien sur ces dessins comment un point de discontinuité de f entraîne un point de non-dérivabilité pour F.

Corollaire 21.22.

Soit f une fonction continue sur un intervalle I et $a \in I$. Alors, $x \mapsto \int_a^x f(t)\,\mathrm{d}t$ est l'unique primitive de f nulle en a.

Proposition 21.23.

Soit f une fonction continue sur $[a,b]$ et F une primitive de f. Alors,

$$\int_a^b f(t)\,\mathrm{d}t = F(b) - F(a).$$

Démonstration

Les fonctions F et $G : x \mapsto \int_a^x f(t)\,\mathrm{d}t$ ne diffèrent que d'une constante d'où

$$G(b) - F(b) = G(a) - F(a).$$

Ainsi, comme $G(a) = 0$,

$$\int_a^b f(t)\,\mathrm{d}t = F(b) - F(a).$$

■

Corollaire 21.24.

Soit f une fonction continue sur I, u et v deux fonctions de classe $\mathscr{C}^1$ de J dans I.
Alors $G : x \mapsto \int_{u(x)}^{v(x)} f(t)\,\mathrm{d}t$ est de classe $\mathscr{C}^1$ sur J et

$$\forall x \in J, \qquad G'(x) = v'(x) f(v(x)) - u'(x) f(u(x)).$$

Exemple

Étudions la fonction $f : x \mapsto \int_x^{2x} \frac{4}{\sqrt{1+t^4}}\,\mathrm{d}t$.
Cette fonction est définie sur tout $\mathbb{R}$ en tant qu'intégrale d'une fonction continue ; elle est impaire car

$$\forall x \in \mathbb{R}, \qquad f(-x) = \int_{-x}^{-2x} \frac{4}{\sqrt{1+t^4}}\,\mathrm{d}t = -\int_x^{2x} \frac{4}{\sqrt{1+u^4}}\,\mathrm{d}u = -f(x),$$

en posant $u = -t$: il suffit donc d'étudier cette fonction sur $\mathbb{R}_+$.
D'après le corollaire précédent, f est de classe $\mathscr{C}^1$ de dérivée égale, pour tout $x \in \mathbb{R}$, à

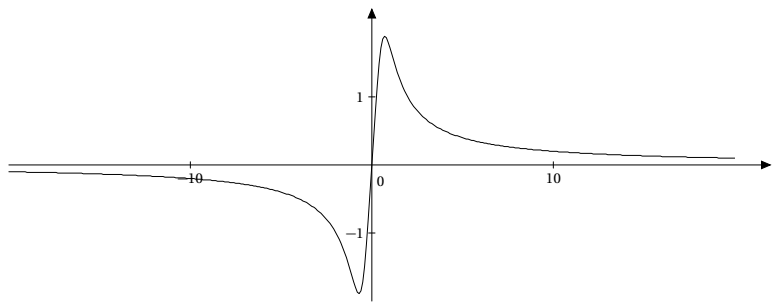
$$\begin{aligned} f'(x) &= \frac{8}{\sqrt{1+2^4x^4}} - \frac{4}{\sqrt{1+x^4}} \\ &= \frac{12(1-4x^4)}{\sqrt{1+2^4x^4}\sqrt{1+x^4}(\sqrt{1+2^4x^4}+2\sqrt{1+x^4})}. \end{aligned}$$

On en déduit le tableau de variations suivant

x	0		$2^{-\frac{1}{2}}$		$+\infty$
$f'(x)$		$+$	0	$-$	
f	0	$\nearrow$		$\searrow$	0

Pour la limite en $+\infty$, on utilise l'encadrement suivant (dont chaque membre tend vers 0 lorsque x tend vers $+\infty$)

$$\frac{4x}{\sqrt{1+2^4x^4}} \leq f(x) \leq \frac{4x}{\sqrt{1+x^4}}.$$



3.3. Intégration par parties

Proposition 21.25. Formule d'intégration par parties

Soit u, v deux fonctions de classe $\mathscr{C}^1$ sur $[a,b]$. Alors,

$$\int_a^b u'(t)v(t)\,\mathrm{d}t = \big[u(t)v(t)\big]_a^b - \int_a^b u(t)v'(t)\,\mathrm{d}t\,.$$

Démonstration
Il suffit d'intégrer la relation $(u.v)'(t) = u'(t)v(t) + u(t)v'(t)$ entre a et b. ■

Exemple

▹ Calculons une primitive de ln par intégration par parties : pour tout $x \in \mathbb{R}_+^*$,

$$\begin{aligned}\int_1^x \ln t\,\mathrm{d}t &= \big[t\ln t\big]_1^x - \int_1^x t\frac{1}{t}\,\mathrm{d}t\\ &= x\ln x - x + 1.\end{aligned}$$

Par suite, les primitives de ln sont les fonctions $x \mapsto x\ln x - x + C$ avec $C \in \mathbb{R}$.
▹ De même, on obtient par intégration par parties que les primitives de arctan sont les fonctions $x \mapsto x\arctan x - \frac{1}{2}\ln(x^2+1) + C$ avec $C \in \mathbb{R}$.

Exemple Intégrales de Wallis

Pour tout $n \in \mathbb{N}$, on pose $I_n = \int_0^{\pi/2} \cos^n t\,\mathrm{d}t$. Calculons I_n à l'aide d'une intégration par parties.
Pour tout entier n,

$$\begin{aligned}I_n - I_{n+2} &= \int_0^{\pi/2} \cos^n t \sin^2 t\,\mathrm{d}t\\ &= \left[\frac{-\cos^{n+1} t}{n+1}\sin t\right]_0^{\frac{\pi}{2}} - \int_0^{\pi/2} \frac{-\cos^{n+1} t}{n+1}\cos t\,\mathrm{d}t = \frac{1}{n+1} I_{n+2}.\end{aligned}$$

D'où $I_{n+2} = \frac{n+1}{n+2} I_n$ pour tout $n \in \mathbb{N}$. Comme on a $I_0 = \frac{\pi}{2}$ et $I_1 = 1$, pour tout $p \in \mathbb{N}$,

$$I_{2p} = \frac{2p-1}{2p}\cdots\frac{1}{2}I_0 = \frac{(2p)!}{(2^p p!)^2}\cdot\frac{\pi}{2}, \qquad I_{2p+1} = \frac{2p}{2p+1}\cdots\frac{2}{3}I_1 = \frac{(2^p p!)^2}{(2p+1)!}.$$

Exemple

Soit P un polynôme et $a \neq 0$. Montrons par récurrence sur le degré $n = \deg P$ que la fonction $x \mapsto P(x)e^{ax}$ admet une primitive de la forme $x \mapsto Q(x)e^{ax}$ avec Q polynôme de degré n.
▹ Pour $n = 0$, P est une constante C, et une primitive de $x \mapsto P(x)e^{ax}$ est $x \mapsto \frac{C}{a}e^{ax}$ d'où le résultat.
▹ Soit $n \in \mathbb{N}$ tel que le résultat est vrai au rang n et P polynôme de degré $n+1$. Calculons par intégration par parties : pour tout $x \in \mathbb{R}$,

$$\int_0^x P(t)e^{at}\,\mathrm{d}t = \left[\tfrac{1}{a}P(t)e^{at}\right]_0^x - \int_0^x \tfrac{1}{a}P'(t)e^{at}\,\mathrm{d}t\,.$$

D'après l'hypothèse de récurrence pour le polynôme P' de degré n, il existe un polynôme Q de degré n tel que

$$\int_0^x P(t)e^{at}\,\mathrm{d}t = \left[\tfrac{1}{a}P(t)e^{at} - \tfrac{1}{a}Q(t)e^{at}\right]_0^x = \frac{1}{a}(P(x)-Q(x))e^{ax} - \frac{1}{a}(P(0)-Q(0)).$$

Comme $\frac{1}{a}(P-Q)$ est de degré $n+1$, on a que $x \mapsto \frac{1}{a}(P(x)-Q(x))e^{ax}$ est une primitive de la forme souhaitée de $x \mapsto P(x)e^{ax}$.

Conseils méthodologiques

D'après les exemples précédents, on en déduit que la formule d'intégration par parties est pratique pour effectuer des calculs notamment dans les cas suivants :

- faire disparaître une fonction « compliquée » à dérivée simple (comme ln, arctan...)
- obtenir des relations de récurrence (lorsqu'il y a un paramètre entier n)
- faire baisser le degré d'un polynôme apparaissant dans l'intégrale.

3.4. Changement de variables

Proposition 21.26.

Soit f une fonction continue sur un intervalle I et φ une fonction de classe $\mathscr{C}^1$ de $[a,b]$ dans I. Alors,

$$\int_a^b f(\varphi(x))\varphi'(x)\,\mathrm{d}x = \int_{\varphi(a)}^{\varphi(b)} f(y)\,\mathrm{d}y\,.$$

Démonstration

Les deux fonctions $b \mapsto \int_a^b f(\varphi(x))\varphi'(x)\,\mathrm{d}x$ et $b \mapsto \int_{\varphi(a)}^{\varphi(b)} f(y)\,\mathrm{d}y$ sont des primitives de $x \mapsto f(\varphi(x))\varphi'(x)$ nulles en a donc sont égales. ■

Exemple

En posant $y = x^2$ avec x entre 0 et 1, on obtient

$$\int_0^1 x\tan(x^2)\,\mathrm{d}x = \frac{1}{2}\int_0^1 \tan y\,\mathrm{d}y = \frac{1}{2}\left[-\ln|\cos(y)|\right]_0^1 = -\frac{1}{2}\ln(\cos 1).$$

Exemple

En posant $y = \sin x$ avec x entre $\frac{\pi}{2}$ et 0, on obtient

$$\int_0^1 \sqrt{1-y^2}\,\mathrm{d}y = \int_0^{\frac{\pi}{2}} \cos^2(x)\,\mathrm{d}x = \int_0^{\frac{\pi}{2}} \frac{\cos(2x)+1}{2}\,\mathrm{d}x = \frac{\pi}{4}.$$

Proposition 21.27.

L'intégrale d'une fonction impaire sur un segment symétrique par rapport à 0 est nulle.

Démonstration

Soit f impaire. En posant $y = -x$,

$$\int_{-a}^{a} f(x)\,\mathrm{d}x = \int_{a}^{-a} f(-y)(-\mathrm{d}y) = -\int_{-a}^{a} f(y)\,\mathrm{d}y\,.$$

Donc $\int_{-a}^{a} f(x)\,\mathrm{d}x = 0$. ■

3.5. Intégration des fractions rationnelles

Conseils méthodologiques

Pour intégrer une fraction rationnelle, on commence par la décomposer en éléments simples puis on intègre chacun des termes indépendamment.

▹ Voici une primitive pour les éléments simples associés à un polynôme irréductible réel de degré 1 avec $a \in \mathbb{R}$ et $n \geq 2$.

$f(x)$	$F(x)$
$\dfrac{1}{x-a}$	$\ln\lvert x-a\rvert$
$\dfrac{1}{(x-a)^n}$	$\dfrac{1}{1-n}\cdot\dfrac{1}{(x-a)^{n-1}}$

Remarque

Attention, pour un polynôme complexe de degré 1, il faut se ramener au cas réel en utilisant la quantité conjuguée.

$f(x)$	$F(x)$
$\dfrac{1}{x-a-ib} = \dfrac{x-a+ib}{(x-a)^2+b^2}$	$\dfrac{1}{2}\ln((x-a)^2+b^2) + i\arctan\dfrac{x-a}{b}$

▹ Pour les éléments simples associés à un polynôme irréductible réel de degré 2, on se ramène à des fonctions de la forme $\frac{u'}{u}$ et des primitives en arctan. Avec $n \in \mathbb{N}\setminus\{0\}$, $a > 0$ et $b^2 - 4ac < 0$, on a

$f(x)$	$F(x)$
$\dfrac{2ax+b}{ax^2+bx+c}$	$\dfrac{1}{2}\ln(ax^2+bx+c)$
$\dfrac{2ax+b}{(ax^2+bx+c)^n}$	$\dfrac{1}{1-n}\cdot\dfrac{1}{(ax^2+bx+c)^{n-1}}$
$\dfrac{1}{ax^2+bx+c}=\alpha\dfrac{1}{(x-\beta)^2+\gamma^2}$	$\dfrac{\alpha}{\gamma}\arctan\dfrac{x-\beta}{\gamma}$

Conseils méthodologiques

Pour les termes restants (de la forme $\frac{1}{(ax^2+bx+c)^n}$), on pose

$$I_n=\int_0^x \frac{1}{(at^2+bt+c)^n}\,\mathrm{d}t$$

et on trouve une relation de récurrence entre I_n et I_{n+1} par une intégration par parties.

Exemple

En décomposant la fraction $\frac{X+2}{X^2-5X+6}$ en éléments simples, on obtient

$$\int_0^1 \frac{x+2}{x^2-5x+6}\,\mathrm{d}x=\int_0^1 \frac{-4}{x-2}\,\mathrm{d}x+\int_0^1 \frac{5}{x-3}\,\mathrm{d}x\,.$$

Exemple

En écrivant $x+1=\frac{1}{2}(2x+1)+\frac{1}{2}$, on obtient

$$\begin{aligned}\int_0^1 \frac{x+1}{x^2+x+1}\,\mathrm{d}x&=\frac{1}{2}\int_0^1 \frac{2x+1}{x^2+x+1}\,\mathrm{d}x+\frac{1}{2}\int_0^1 \frac{1}{x^2+x+1}\,\mathrm{d}x\\&=\frac{1}{2}\big[\ln(x^2+x+1)\big]_0^1+\frac{1}{2}\int_0^1 \frac{1}{(x+\frac{1}{2})^2+\frac{3}{4}}\,\mathrm{d}x\,.\end{aligned}$$

Conseils méthodologiques

Dans le cas où la fonction à intégrer est de la forme $F(\sin x,\cos x)$ avec F une fraction rationnelle en ces arguments, on peut poser $y=\tan\frac{x}{2}$ et on utilise les expressions rationnelles de $\sin x$ et $\cos x$ en la variable y :

$$\sin x=\frac{2y}{1+y^2},\qquad \cos x=\frac{1-y^2}{1+y^2},\qquad \tan x=\frac{2y}{1-y^2}.$$

On est alors ramené à une intégrale d'une fraction rationnelle car $\mathrm{d}x=\frac{2\,\mathrm{d}y}{1+y^2}$.

4. Formules de Taylor

Définition 21.28.

Soit f une fonction n fois dérivable et a un réel appartenant au domaine de définition de f. Le polynôme de Taylor d'ordre n de la fonction f en a est défini par

$$T_n(X)=\sum_{k=0}^{n}\frac{f^{(k)}(a)}{k!}(X-a)^k.$$

On étudie désormais l'erreur commise lorsque l'on remplace une fonction f par la fonction polynomiale associée T_n. On distingue deux types de résultats :

- les résultats globaux où l'on décrit l'erreur sur l'ensemble de l'intervalle de définition ;
- les résultats locaux où l'on étudie le comportement de l'erreur au voisinage du point où l'on considère le polynôme de Taylor.

4.1. Formules globales

Proposition 21.29. Formule de Taylor avec reste intégral

Soit f une fonction de classe $\mathscr{C}^{n+1}$ sur I et $a\in I$. Alors,

$$\forall x\in I,\qquad f(x)=\sum_{k=0}^{n}\frac{f^{(k)}(a)}{k!}(x-a)^k+\int_a^x\frac{(x-t)^n}{n!}f^{(n+1)}(t)\,\mathrm{d}t\,.$$

Démonstration

On procède par récurrence sur n.

▹ À l'ordre 0, cette formule se réécrit $f(x)=f(a)+\int_a^x f'(t)\,\mathrm{d}t$ et est évidemment vraie.

▹ Soit $n\in\mathbb{N}$. Supposons que la propriété est vérifiée pour les fonctions de classe $\mathscr{C}^{n+1}$ et considérons f de classe $\mathscr{C}^{n+2}$. En appliquant l'hypothèse de récurrence, on obtient

$$f(x)=\sum_{k=0}^{n}\frac{f^{(k)}(a)}{k!}(x-a)^k+\int_a^x\frac{(x-t)^n}{n!}f^{(n+1)}(t)\,\mathrm{d}t\,.$$

En intégrant par parties le reste intégral, on obtient

$$\begin{aligned}\int_a^x\frac{(x-t)^n}{n!}f^{(n+1)}(t)\,\mathrm{d}t&=\left[-\frac{(x-t)^{n+1}}{(n+1)!}f^{(n+1)}(t)\right]_a^x+\int_a^x\frac{(x-t)^{n+1}}{(n+1)!}f^{(n+2)}(t)\,\mathrm{d}t\\&=\frac{(x-a)^{n+1}}{(n+1)!}f^{(n+1)}(a)+\int_a^x\frac{(x-t)^{n+1}}{(n+1)!}f^{(n+2)}(t)\,\mathrm{d}t\,.\end{aligned}$$

D'où le résultat. ■

Exemple

En appliquant cette formule en 0, à l'ordre 3, à la fonction cos, on vérifie

$$\forall x\in[0,\pi],\qquad \cos(x)\geq 1-\frac{x^2}{2}.$$

En effet, le reste intégral de Taylor correspondant à la différence des deux membres est $\int_0^x\frac{(x-t)^2}{2}\sin(t)\,\mathrm{d}t$ et il est positif (comme intégrale de fonction positive) sur l'intervalle considéré.

L'expression du reste est compliquée et on se contente souvent de la majoration suivante :

Corollaire 21.30. Inégalité de Taylor-Lagrange

Soit f une fonction de classe $\mathscr{C}^{n+1}$ sur I et $a \in I$. Alors,

$$\forall x \in I, \qquad \left| f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k \right| \leq \sup_{t\in[a,x]} |f^{(n+1)}(t)| \frac{|x-a|^{n+1}}{(n+1)!}.$$

Démonstration

Il suffit de majorer le reste en utilisant la croissance de l'intégrale.

- Si $x > a$, alors

$$\begin{aligned}\left| \int_a^x \frac{(x-t)^n}{n!} f^{(n+1)}(t)\,\mathrm{d}t \right| &\leq \int_a^x \frac{(x-t)^n}{n!} |f^{(n+1)}(t)|\,\mathrm{d}t \\ &\leq \sup_{t\in[a,x]} |f^{(n+1)}(t)| \int_a^x \frac{(x-t)^n}{n!}\,\mathrm{d}t \\ &= \sup_{t\in[a,x]} |f^{(n+1)}(t)| \frac{(x-a)^{n+1}}{(n+1)!}.\end{aligned}$$

- Si $x < a$, alors

$$\begin{aligned}\left| \int_a^x \frac{(x-t)^n}{n!} f^{(n+1)}(t)\,\mathrm{d}t \right| &\leq \int_x^a \frac{(t-x)^n}{n!} |f^{(n+1)}(t)|\,\mathrm{d}t \\ &\leq \sup_{t\in[x,a]} |f^{(n+1)}(t)| \int_x^a \frac{(t-x)^n}{n!}\,\mathrm{d}t \\ &= \sup_{t\in[x,a]} |f^{(n+1)}(t)| \frac{(a-x)^{n+1}}{(n+1)!}.\end{aligned}$$

■

Voici quelques exemples d'applications de ces formules.

Exemple Inégalité de Kolmogorov

Soit f une fonction de classe $\mathscr{C}^2$ telle que f est bornée par M_0 et f'' est bornée par M_2. Alors, f' est bornée, plus précisément, on a

$$\forall x \in I, \qquad |f'(x)| \leq 2\sqrt{M_0 M_2}.$$

En effet, l'inégalité de Taylor-Lagrange entre a et x nous donne

$$|f(x) - f(a) - (x-a)f'(a)| \leq M_2 \frac{|x-a|^2}{2}.$$

Or, $|f(x) - f(a) - (x-a)f'(a)| \geq |(x-a)f'(a)| - |f(x) - f(a)|$ par inégalité triangulaire. Donc

$$|f'(a)| \leq M_2 \frac{|x-a|}{2} + M_0 \frac{2}{|x-a|}.$$

En étudiant les variations de $x \mapsto M_2 \frac{|x-a|}{2} + M_0 \frac{2}{|x-a|}$, on obtient que le minimum de cette fonction est $2\sqrt{M_0 M_2}$. Par conséquent, en choisissant la valeur de x réalisant ce minimum,

$$|f'(a)| \leq 2\sqrt{M_0 M_2}.$$

Exemple Développements en série entière

En appliquant l'inégalité de Taylor-Lagrange à la fonction cos entre 0 et x, on obtient

$$\left|\cos(x)-\sum_{k=0}^{n}\frac{(-1)^k}{(2k)!}x^{2k}\right|\leq\frac{|x|^{2n+2}}{(2n+2)!}.$$

Or, le majorant tend vers 0 lorsque n tend vers $+\infty$ d'où, pour tout $x\in\mathbb{R}$,

$$\lim_{n\to\infty}\sum_{k=0}^{n}\frac{(-1)^k}{(2k)!}x^{2k}=\cos(x).$$

On montre, de même, que

$$\lim_{n\to\infty}\sum_{k=0}^{n}\frac{(-1)^k}{(2k+1)!}x^{2k+1}=\sin(x),$$

$$\lim_{n\to\infty}\sum_{k=0}^{n}\frac{1}{k!}x^k=e^x.$$

4.2. Formule locale

Dans cette sous-section, on rappelle puis on établit la formule de Taylor-Young pour les développements limités.

Proposition 21.31. Formule de Taylor-Young

Soit f une fonction de classe $\mathscr{C}^n$ sur I et $a\in I$. Alors,

$$f(x)\underset{a}{=}\sum_{k=0}^{n}\frac{f^{(k)}(a)}{k!}(x-a)^k+o((x-a)^n).$$

Démonstration

Commençons par écrire la formule de Taylor avec reste intégral à l'ordre $n-1$

$$f(x)-\sum_{k=0}^{n}\frac{f^{(k)}(a)}{k!}(x-a)^k=\int_a^x\frac{(x-t)^{n-1}}{(n-1)!}f^{(n)}(t)\,\mathrm{d}t-\frac{f^{(n)}(a)}{n!}(x-a)^n.$$

Le second membre se réécrit $\int_a^x\frac{(x-t)^{n-1}}{(n-1)!}(f^{(n)}(t)-f^{(n)}(a))\,\mathrm{d}t$. Utilisons dorénavant la continuité de $f^{(n)}$ en a. Soit $\varepsilon>0$. Il existe $\eta>0$ tel que

$$\forall x\in I,\qquad |x-a|\leq\eta\quad\Rightarrow\quad|f^{(n)}(x)-f^{(n)}(a)|\leq\varepsilon.$$

Par conséquent, pour tout $x\in[a,a+\eta]$,

$$\left|\int_a^x\frac{(x-t)^{n-1}}{(n-1)!}(f^{(n)}(t)-f^{(n)}(a))\,\mathrm{d}t\right|\leq\varepsilon\frac{(x-a)^n}{n!}.$$

On conclut par encadrement. ■

FICHE

SYNTHÈSE

L'intégrale des fonctions continues par morceaux sur un segment est définie à partir de l'intégrale, plus simple, des fonctions en escalier. On utilise pour cela l'approximation uniforme déduite du théorème de Heine pour les fonctions continues sur un segment.

Propriétés de l'intégrale

▹ Soit f, g deux fonctions continues par morceaux sur $[a, b]$ à valeurs dans $\mathbb{C}$ et $\lambda, \mu \in \mathbb{C}$.

Linéarité

$$\int_a^b (\lambda f + \mu g)(t)\,\mathrm{d}t = \lambda \int_a^b f(t)\,\mathrm{d}t + \mu \int_a^b g(t)\,\mathrm{d}t\,.$$

Inégalité en module

$$\left|\int_a^b f(t)\,\mathrm{d}t\right| \leq \int_a^b |f(t)|\,\mathrm{d}t\,.$$

Inégalité de Cauchy-Schwarz

$$\left|\int_a^b f(t)\overline{g(t)}\,\mathrm{d}t\right| \leq \sqrt{\int_a^b |f(t)|^2\,\mathrm{d}t \int_a^b |g(t)|^2\,\mathrm{d}t}.$$

Relation de Chasles

$$\forall c \in]a, b[, \quad \int_a^b f(t)\,\mathrm{d}t = \int_a^c f(t)\,\mathrm{d}t + \int_c^b f(t)\,\mathrm{d}t\,.$$

▹ Soit f, g deux fonctions continues par morceaux sur $[a, b]$ à valeurs dans $\mathbb{R}$ telles que $f \leq g$.

Croissance de l'intégrale

$$\int_a^b f(t)\,\mathrm{d}t \leq \int_a^b g(t)\,\mathrm{d}t\,.$$

Voici un cas particulier fréquemment utilisé de convergence des sommes de Riemann.

Sommes de Riemann

Soit f continue sur $[a, b]$. Alors,

$$\lim_n \frac{b-a}{n} \sum_{k=0}^{n-1} f\left(a + k\frac{b-a}{n}\right) = \int_a^b f(x)\,\mathrm{d}x$$

$$\lim_n \frac{b-a}{n} \sum_{k=1}^{n} f\left(a + k\frac{b-a}{n}\right) = \int_a^b f(x)\,\mathrm{d}x\,.$$

Une fois l'intégrale construite, on obtient le résultat liant intégration et primitivation.

Intégrale fonction des bornes

Soit f continue par morceaux sur l'intervalle I et $a \in I$. La fonction

$$F : x \longmapsto \int_a^x f(t)\,\mathrm{d}t$$

est continue partout et dérivable en tout point de continuité de f.

Les formules d'intégration par partie et de changement de variables fournissent des méthodes de calcul d'intégrale.

Méthodes de calcul

▹ Soit u, v deux fonctions de classe $\mathscr{C}^1$ sur $[a,b]$. Alors,

$$\int_a^b u'(t)v(t)\,\mathrm{d}t = \big[u(t)v(t)\big]_a^b - \int_a^b u(t)v'(t)\,\mathrm{d}t\,.$$

▹ Soit f une fonction continue sur un intervalle I et φ une fonction de classe $\mathscr{C}^1$ de $[a,b]$ dans I. Alors,

$$\int_a^b f(\varphi(x))\varphi'(x)\,\mathrm{d}x = \int_{\varphi(a)}^{\varphi(b)} f(y)\,\mathrm{d}y\,.$$

Il existe deux types de formules de Taylor : celles globales (avec reste intégral, inégalité de Taylor-Lagrange) et celle locale (Taylor-Young).

Formules de Taylor

▹ Soit f une fonction de classe $\mathscr{C}^{n+1}$ sur I et $a \in I$. Alors

$$\forall x \in I, \qquad f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + \int_a^x \frac{(x-t)^n}{n!} f^{(n+1)}(t)\,\mathrm{d}t\,.$$

▹ Soit f une fonction de classe $\mathscr{C}^{n+1}$ sur I et $a \in I$. Alors

$$\forall x \in I, \qquad \left| f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k \right| \leq \sup_{t\in[a,x]} |f^{(n+1)}(t)| \frac{|x-a|^{n+1}}{(n+1)!}.$$

▹ Soit f une fonction de classe $\mathscr{C}^n$ sur I et $a \in I$. Alors,

$$f(x) \underset{a}{=} \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + o((x-a)^n).$$

EXERCICES

Vrai ou faux ?

		Vrai	Faux
a)	$\int_1^2 x\,\mathrm{d}x = \frac{5}{2}$.	☐	☐
b)	Une primitive de $x \mapsto \ln x$ est $x \mapsto x\ln x - x + 1$.	☐	☐
c)	Une primitive de $x \mapsto \dfrac{1}{a^2+x^2}$ est $x \mapsto \arctan \dfrac{x}{a}$.	☐	☐
d)	Une primitive de tan est $x \mapsto -\ln\lvert\cos x\rvert$.	☐	☐
e)	Une primitive de $x \mapsto \dfrac{x}{\sqrt{1-x^2}}$ est $x \mapsto \arcsin x$.	☐	☐
f)	Une primitive de $x \mapsto \dfrac{1}{x\ln x}$ est $x \mapsto \ln\lvert\ln x\rvert$.	☐	☐
g)	Soit f continue sur $[0,1]$ telle que $\left\lvert\int_0^1 f(t)\,\mathrm{d}t\right\rvert = \int_0^1 \lvert f(t)\rvert\,\mathrm{d}t$. Alors, f est de signe constant.	☐	☐
h)	Soit f continue sur $[0,1]$ et F une primitive de f. Alors, $\int_0^1 \lvert f(t)\rvert\,\mathrm{d}t = \left[\lvert F(t)\rvert\right]_0^1$.	☐	☐
i)	Soit f continue sur $[a,b]$. Alors, $$\lim \frac{b-a}{n}\sum_{k=0}^{n-1} f\left(k\frac{b-a}{n}\right) = \int_a^b f(t)\,\mathrm{d}t\,.$$	☐	☐
j)	Soit f continue sur $[0,1]$. Alors, $$\lim \frac{1}{n}\sum_{k=1}^{n} f\left(\frac{k}{n}\right) = \int_0^1 f(t)\,\mathrm{d}t\,.$$	☐	☐
k)	Pour f de classe $\mathscr{C}^{n+1}$, $$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!}x^k + \int_0^x \frac{u^n}{n!} f^{(n+1)}(x-u)\,\mathrm{d}u\,.$$	☐	☐
l)	Pour f de classe $\mathscr{C}^{n+1}$, $$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!}x^k + x^{n+1}\int_0^1 \frac{(1-t)^n}{n!} f^{(n+1)}(tx)\,\mathrm{d}t\,.$$	☐	☐

Exercices d'application du chapitre 21

●●○ **Exercice 1**

Soit $f : x \mapsto x\lfloor x\rfloor - \lfloor x\rfloor^2$. Calculer $\int_0^n f(t)\,\mathrm{d}t$ où $n \in \mathbb{N}$.

●●○ **Exercice 2**

Considérons les intégrales

$$I = \int_0^{\frac{\pi}{2}} \frac{\sin t}{\sqrt{1+2\sin t\cos t}}\,\mathrm{d}t, \qquad J = \int_0^{\frac{\pi}{2}} \frac{\cos t}{\sqrt{1+2\sin t\cos t}}\,\mathrm{d}t\,.$$

1. Calculer $I+J$.
2. Montrer que $I = J$ par un judicieux changement de variables.
3. En déduire la valeur de I.

●●○ **Exercice 3**

Soit $P = \sum_{k=0}^{n} a_k X^k \in \mathbb{C}[X]$. Montrer que

$$\int_0^{2\pi} |P(e^{i\theta})|^2\,\mathrm{d}\theta = 2\pi \sum_{k=0}^{n} |a_k|^2.$$

●○○ **Exercice 4**

Montrer qu'une fonction $f : [0,1] \to \mathbb{R}$ continue admet exactement une primitive d'intégrale nulle sur le segment $[0,1]$.

●○○ **Exercice 5**

Soit $f : \mathbb{R} \to \mathbb{R}$ continue telle qu'il existe $T > 0$ et $C \in \mathbb{R}$ vérifiant

$$\forall x \in \mathbb{R}, \qquad \int_x^{x+T} f(t)\,\mathrm{d}t = C.$$

Montrer que f est T-périodique.

●●○ **Exercice 6**

Soit $f : [0,1] \to \mathbb{R}$ continue telle que $\int_0^1 f(t)\,\mathrm{d}t = 0$, $a = \min\{f(x),\ x \in [a,b]\}$, et $b = \max\{f(x),\ x \in [a,b]\}$. Montrer que $\int_0^1 f(t)^2\,\mathrm{d}t \leq -ab$.

●○○ **Exercice 7**

Soit $f : [0,1] \to \mathbb{R}$ continue telle que $\int_0^1 f(t)\,\mathrm{d}t = \frac{1}{2}$. Montrer que f admet un point fixe.

●●○ **Exercice 8**

Soit $f : [a,b] \to \mathbb{R}$ continue.

1. Montrer que

$$(b-a)\min_{x\in[a,b]} f(x) \leq \int_a^b f(t)\,\mathrm{d}t \leq (b-a)\max_{x\in[a,b]} f(x).$$

2. En déduire qu'il existe $c \in [a,b]$ tel que
$$f(c) = \frac{1}{b-a}\int_a^b f(t)\,\mathrm{d}t\,.$$
3. Soit g continue par morceaux positive d'intégrale non nulle. Montrer qu'il existe $c \in [a,b]$ tel que
$$f(c)\int_a^b g(t)\,\mathrm{d}t = \int_a^b f(t)g(t)\,\mathrm{d}t\,.$$

●○○ **Exercice 9**

Soit $f : [0,1] \to \mathbb{R}$ continue et, pour tout $n \in \mathbb{N}$, $I_n = \int_0^1 f(x^n)\,\mathrm{d}x$. Déterminer la limite de la suite $(I_n)_n$.

●●○ **Exercice 10**

Soit $f : [a,b] \to \mathbb{R}$ continue non identiquement nulle, telle que
$$\forall\ k \in [\![0, n-1]\!], \qquad \int_a^b t^k f(t)\,\mathrm{d}t = 0.$$
Montrer que f change au moins n fois de signe sur $]a,b[$.

●●○ **Exercice 11**

Soit $a,\ b \in \mathbb{R}_+^*$ et $f : [0,a] \to [0,b]$ une bijection de classe $\mathscr{C}^1$, strictement croissante. Montrer que
$$ab = \int_0^a f(x)\,\mathrm{d}x + \int_0^b f^{-1}(y)\,\mathrm{d}y\,.$$
Interpréter géométriquement ce résultat.

●●● **Exercice 12**

Soit $f : [a,b] \to \mathbb{R}$ continue. Montrer que $\int_a^b f(t)\cos(nt)\,\mathrm{d}t \to 0$.

●○○ **Exercice 13**

Soit $f : [0,1] \to \mathbb{R}$ continue. Déterminer la limite des expressions suivantes

1. $\frac{1}{n}\sum_{k=1}^{n} f\left(\frac{k}{n^2}\right)$
2. $\frac{1}{n}\sum_{k=1}^{42n} f\left(\frac{k}{n^2}\right)$

●●● **Exercice 14**

Pour tout $n \geq 1$, posons $u_n = \sum_{k=1}^{n} \frac{n}{k^2+n^2}$.

1. Montrer que la suite $(u_n)_n$ converge et déterminer sa limite ℓ.
2. Préciser un équivalent simple de $u_n - \ell$.

●●● **Exercice 15**

Soit $f : [a,b] \to \mathbb{R}_+^*$ continue.

1. Montrer qu'il existe une subdivision $(x_k^{(n)})_{k\in[0,n]}$ telle que
$$\forall k \in [\![1,n]\!], \qquad \int_{x_{k-1}^{(n)}}^{x_k^{(n)}} f(t)\,\mathrm{d}t = \frac{1}{n}\int_a^b f(t)\,\mathrm{d}t\,.$$
2. Déterminer la limite de $\frac{1}{n}\sum_{k=1}^{n} x_k^{(n)}$.

EXERCICES

●●○ **Exercice 16**

Soit $f:[0,1]\to\mathbb{R}$ de classe $\mathscr{C}^1$. Montrer qu'il existe $c\in[0,1]$ tel que

$$\int_0^1 f(t)\,\mathrm{d}t = f(0)+\frac{1}{2}f'(c).$$

●●○ **Exercice 17**

Soit $f:I\to R$ une fonction de classe $\mathscr{C}^{n+1}$ et $a<b\in I$.

1. Définissons l'application

$$\varphi : x\mapsto f(b)-\sum_{k=0}^{n}\frac{f^{(k)}(x)}{k!}(b-x)^k+\frac{A}{(n+1)!}(b-x)^{n+1}$$

   où A est une constante choisie de sorte à ce que $\varphi(a)=0$.
   Montrer qu'il existe $c\in[a,b]$ tel que $\varphi'(c)=0$.
2. En déduire qu'il existe $c\in]a,b[$ tel que

$$f(b)=\sum_{k=0}^{n}\frac{f^{(k)}(a)}{k!}(b-a)^k+\frac{f^{(n+1)}(c)}{(n+1)!}(b-a)^{n+1}.$$

●●● **Exercice 18**

Soit une fonction $f:I\to\mathbb{R}$ qui admet une dérivée d'ordre n sur I et $a\in I$. Montrer que

$$\lim_{x\to a}\frac{1}{(x-a)^n}\left(f(x)-\sum_{k=0}^{n}\frac{f^{(k)}(a)}{k!}(x-a)^k\right)=0.$$

Il s'agit d'établir la formule de Taylor-Young pour une fonction n fois dérivable et non plus de classe $\mathscr{C}^n$.

Exercices d'approfondissement du chapitre 21

●●● **Exercice A**

Soit P un polynôme complexe non constant. Montrer que l'ensemble

$$\mathscr{Z}=\left\{x\in\mathbb{R},\ \int_0^x P(t)e^{it}\,\mathrm{d}t=0\right\}$$

est fini.

●●○ **Exercice B**

Soit $f:[0,1]\to\mathbb{R}$ continue. Montrer que $n^2\int_0^1(x^n-x^{n+1})f(x)\,\mathrm{d}x\to f(1)$.

●●○ **Exercice C**

Soit $f:[0,1]\to\mathbb{R}$ continue. Déterminer la limite des expressions suivantes :

1. $\frac{1}{n^2}\sum\limits_{1\le k<l\le n} f\left(\frac{k}{n}\right)f\left(\frac{l}{n}\right)$
2. $\frac{1}{n^2}\sum\limits_{1\le k\le l\le n} f\left(\frac{k}{n}\right)f\left(\frac{l}{n}\right)$

●●○ **Exercice D**

Soit $A_1A_2\ldots A_n$ un polygone régulier inscrit dans le cercle unité tel que $A_1(1,0)$. Calculer la limite de la quantité $\frac{1}{n}\sum_{k=2}^{n}A_1A_k$.

●●● **Exercice E**

Soit x_n l'argument du premier maximum local de $f_n : x \mapsto \sum_{k=1}^{n} \frac{\sin(kx)}{k}$ sur $\mathbb{R}_+$. Montrer que

$$f_n(x_n) \to \int_0^{\pi} \frac{\sin t}{t}\,\mathrm{d}t\,.$$

●●● **Exercice F**

Déterminer la borne inférieure de la partie de $\mathbb{R}$ définie par

$$E = \left\{ c \in \mathbb{R},\ \forall f \in \mathscr{C}^0([0,1],\mathbb{R}_+),\ \int_0^1 f(\sqrt[n]{t})\,\mathrm{d}t \leq c \int_0^1 f(t)\,\mathrm{d}t \right\}.$$

Problèmes du chapitre 21

●○○ **Problème 1 Inégalité de Wirtinger**

Soit $f : [0,\pi] \to \mathbb{R}$ une fonction de classe $\mathscr{C}^1$ vérifiant $f(0) = f(\pi) = 0$.

1. a. Montrer que la fonction $x \mapsto f(x)\operatorname{cotan}(x)$ est prolongeable par continuité sur $[0,\pi]$.
   b. Montrer l'existence et déterminer la valeur de la limite en 0 de
   $$x \mapsto (f^2(x) + f^2(\pi - x))\operatorname{cotan}(x).$$
2. Montrer que, pour tout $x \in]0,\pi[$,
   $$\int_x^{\pi-x} (f'(t) - f(t)\operatorname{cotan}(t))^2\,\mathrm{d}t = (f^2(x) + f^2(\pi - x))\operatorname{cotan}(x) + \int_x^{\pi-x} (f'^2(t) - f^2(t))\,\mathrm{d}t\,.$$
3. En déduire l'inégalité suivante dite de Wirtinger
   $$\int_0^{\pi} f^2(t)\,\mathrm{d}t \leq \int_0^{\pi} f'^2(t)\,\mathrm{d}t\,.$$
4. Étudions le cas d'égalité : supposons $\int_0^{\pi} f^2(t)\,\mathrm{d}t = \int_0^{\pi} f'^2(t)\,\mathrm{d}t$.
   a. Montrer que, pour tout $x \in]0,\pi[$, $f'(x) = f(x)\operatorname{cotan}(x)$.
   b. En déduire que f est colinéaire à la fonction sin.

●●● **Problème 2**

Partie A – Noyau de Jackson

Définissons, pour tout $n \in \mathbb{N}^*$, $U_n : \theta \mapsto \sum_{k=0}^{n-1} \sum_{l=-k}^{k} e^{il\theta}$.

1. Vérifier que la fonction U_n est réelle, continue sur $\mathbb{R}$, paire et 2π-périodique.
2. Montrer que, pour tout $\theta \neq 0[\pi]$,
   $$U_n(\theta) = \left(\frac{\sin\frac{n\theta}{2}}{\sin\frac{\theta}{2}}\right)^2.$$
3. En déduire qu'il existe un unique $c_n > 0$ tel que $\int_{-\pi}^{\pi} c_n U_n^2(t)\,\mathrm{d}t = 1$.
   Posons désormais $J_n = c_n U_n^2$.
4. Le but de cette question est d'obtenir une estimation asymptotique de $I_n = \int_{-\pi}^{\pi} |t| J_n(t)\,\mathrm{d}t$.

PROBLÈMES

a. Justifier l'inégalité $\sin\frac{t}{2} \leq \frac{t}{2}$ pour tout $t \in [0,\pi]$, puis montrer que la suite $(n^3 c_n)_n$ est bornée.
Indication : pour le second point, on pourra effectuer le changement de variable $x = \frac{nt}{2}$.
b. Justifier l'inégalité $\frac{t}{\pi} \leq \sin\frac{t}{2}$ pour tout $t \in [0,\pi]$, puis montrer que la suite $(\frac{I_n}{n^2 c_n})_n$ est bornée.
c. Conclure que la suite $(nI_n)_n$ est bornée.

On note dorénavant, pour toute fonction $f : [-1,1] \to \mathbb{R}$ continue,

$$\|f\|_\infty = \sup\{|f(x)|,\ x \in [-1,1]\}.$$

Définition 21.32.

Une fonction $f : [0,1] \to \mathbb{R}$ est limite uniforme de la suite de fonctions $(f_n)_n$ si la suite réelle $(\|f - f_n\|_\infty)_n$ converge vers 0 c'est-à-dire si

$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N},\ \forall n \geq N, \qquad \|f - f_n\|_\infty \leq \varepsilon.$$

Partie B – Théorème de Jackson-Weierstrass

1. Soit $f : [-1,1] \to \mathbb{R}$ de classe $\mathscr{C}^1$.
   a. Montrer qu'il existe une constante κ telle que, pour tout $\theta \in \mathbb{R}$,
   $$\left| f(\cos\theta) - \int_{-\pi}^{\pi} f(\cos(\theta - t)) J_n(t)\,\mathrm{d}t \right| \leq I_n \|f'\|_\infty \leq \frac{\kappa \|f'\|_\infty}{n}.$$
   b. Montrer que
   $$\int_{-\pi}^{\pi} f(\cos(\theta - t)) J_n(t)\,\mathrm{d}t = \frac{1}{2}\int_{-\pi}^{\pi} f(\cos(t))(J_n(\theta + t) + J_n(\theta - t))\,\mathrm{d}t\,.$$
   c. Montrer que, pour tout $n \in \mathbb{N}$, il existe un unique polynôme $T_n \in \mathbb{R}[X]$ tel que
   $$\forall \theta \in \mathbb{R}, \qquad T_n(\cos\theta) = \cos(n\theta).$$
   Préciser le coefficient dominant et le degré de T_n.
   d. Déduire des questions précédentes l'existence d'un polynôme $P_n \in \mathbb{R}_{2n}[X]$ tel que
   $$\int_{-\pi}^{\pi} f(\cos(\theta - t)) J_n(t)\,\mathrm{d}t = P_n(\cos\theta).$$
   Indication : on pourra commencer par exprimer $J_n(\theta)$ sous la forme $\sum_{k=0}^{2n} a_k \cos(k\theta)$ avec des coefficients a_k réels.
   e. En déduire que f est limite uniforme d'une suite de polynômes sur $[-1,1]$.
2. Soit $f : [-1,1] \to \mathbb{R}$ continue.
   Posons, pour tout $h \geq 0$, $\omega(h) = \sup\{|f(x) - f(y)|,\ x, y \in [-1,1]$ tels que $|x - y| \leq h\}$.
   a. Montrer que la fonction ω est croissante sur $\mathbb{R}_+$.
   b. Montrer que, pour tout $\lambda \in \mathbb{R}$ et tout $h \in \mathbb{R}_+$, $\omega(|\lambda| h) \leq (|\lambda| + 1)\omega(h)$.
   Indication : on pourra subdiviser un intervalle de longueur $|\lambda| h$ en intervalles de longueur au plus h.
   c. En déduire qu'il existe une constante κ' telle que
   $$\forall \theta \in \mathbb{R}, \qquad \left| f(\cos\theta) - \int_{-\pi}^{\pi} f(\cos(\theta - t)) J_n(t)\,\mathrm{d}t \right| \leq \kappa' \omega\left(\frac{1}{n}\right).$$
   d. En déduire que f est limite uniforme d'une suite de polynômes sur $[-1,1]$.

●●○ **Problème 3 Transformée de Cesàro**

Définition 21.33.

La transformée de Cesàro d'une fonction continue $f : \mathbb{R}_+ \to \mathbb{R}$ est la fonction $f^* : \mathbb{R}_+ \to \mathbb{R}$ définie par

$$f^* : x \mapsto \begin{cases} \frac{1}{x}\int_0^x f(t)\,\mathrm{d}t & \text{si } x > 0, \\ f(0) & \text{si } x = 0. \end{cases}$$

Partie A – Exemples

1. Considérons dans cette question $f : x \mapsto \arctan x$. Calculer f^* puis donner les limites de f et f^* en $+\infty$.
2. Considérons dans cette question $f : x \mapsto |\sin x|$. Calculer f^* en fonction de $\cos x$ et de $\lfloor \frac{x}{\pi} \rfloor$. Préciser la limite de f^* en $+\infty$.
3. Considérons dans cette question $f : x \mapsto \sqrt{x}\sin x$ (prolongée par continuité en 0).
   a. Montrer que f n'est pas bornée sur $\mathbb{R}_+$.
   b. Montrer que f^* tend vers 0 quand x tend vers $+\infty$.

Partie B – Propriétés de f^*

Dorénavant, f est une application continue de $\mathbb{R}_+$ dans $\mathbb{R}$.

1. Montrer que f^* est de classe $\mathscr{C}^1$ sur $\mathbb{R}_+^*$, puis que f^* est continue en 0.
2. Montrer que si f admet une limite finie ℓ lorsque x tend vers $+\infty$, alors f^* admet la même limite quand x tend vers $+\infty$.
3. Soit $\alpha \in \mathbb{R}^*$ et $\beta \in \mathbb{R}$. Montrer que si la droite d'équation $y = \alpha x + \beta$ est asymptote à la courbe représentative de f au voisinage de $+\infty$, alors la courbe représentative de f^* admet une asymptote au voisinage de $+\infty$ dont on déterminera une équation en fonction de (α, β).
4. Montrer que si f admet un développement limité à l'ordre n en 0, alors f^* admet un développement limité à l'ordre n en 0 que l'on précisera.
5. Montrer que si f est monotone sur $\mathbb{R}_+$, f^* est aussi monotone sur $\mathbb{R}_+$.
6. Montrer que si f est périodique, de période $T > 0$, alors f^* admet une limite quand $x \to +\infty$ puis donner une expression de cette limite en fonction de T et de $\int_0^T f(t)\,\mathrm{d}t$.

Partie C – Inégalité de Hardy

Soit $f : \mathbb{R}_+ \to \mathbb{R}$ continue. Montrer que, pour tout $x > 0$,

$$\int_0^x (f^*)^2(t)\,\mathrm{d}t \leq 4\int_0^x f^2(t)\,\mathrm{d}t\,.$$

PROBLÈMES

CORRIGÉS

Corrigés des Vrai/Faux

a) Faux, le résultat est $\frac{3}{2}$.

b) Vrai.

c) Faux.

d) Vrai.

e) Faux.

f) Vrai.

g) Vrai.

h) Faux, il suffit de considérer $f : x \mapsto 1$ et $F : x \mapsto x - \frac{1}{2}$: on obtient alors $\int_0^1 |f(t)|\,\mathrm{d}t = 1$ mais $[|F(t)|]_0^1 = 0$.

i) Faux.

j) Vrai.

k) Vrai.

l) Vrai.

Corrigés des exercices d'application du chapitre 21

Exercice 1

D'après la relation de Chasles,

$$\begin{aligned}\int_0^n f(t)\,\mathrm{d}t = \int_0^n f(x)\,\mathrm{d}x &= \sum_{k=0}^{n-1}\int_k^{k+1} f(x)\,\mathrm{d}x \\ &= \sum_{k=0}^{n-1}\int_k^{k+1}(kx-k^2)\,\mathrm{d}x = \sum_{k=0}^{n-1}\left(\frac{1}{2}k(2k+1)-k^2\right) \\ &= \frac{n(n-1)}{4}.\end{aligned}$$

Exercice 2

1. Commençons par réécrire le dénominateur : pour tout t,

$$\begin{aligned}1+2\sin t\cos t &= \cos^2 t + \sin^2 t + 2\sin t\cos t \\ &= (\cos t + \sin t)^2.\end{aligned}$$

Remarquons ensuite que, sur l'intervalle étudié, $\cos t + \sin t > 0$ donc

$$\sqrt{1+2\sin t\cos t} = \cos t + \sin t.$$

Par conséquent,

$$I + J = \int_0^{\frac{\pi}{2}} \frac{\sin t + \cos t}{\sqrt{1+2\sin t\cos t}}\,\mathrm{d}t = \int_0^{\frac{\pi}{2}} \mathrm{d}t = \frac{\pi}{2}.$$

2. Utilisons dans I le changement de variable $t = \frac{\pi}{2} - x$:

$$\begin{aligned} I &= \int_{\frac{\pi}{2}}^{0} \frac{\sin(\frac{\pi}{2}-x)}{\sqrt{1+2\sin(\frac{\pi}{2}-x)\cos(\frac{\pi}{2}-x)}}(-\mathrm{d}x) \\ &= \int_0^{\frac{\pi}{2}} \frac{\cos x}{\sqrt{1+2\sin x\cos x}}\,\mathrm{d}x = J. \end{aligned}$$

3. En conclusion, $\frac{\pi}{2} = I + J = 2I$ donc $I = \frac{\pi}{4}$.

Exercice 3

Écrivons dans l'intégrale que $|P(e^{i\theta})|^2 = P(e^{i\theta})\overline{P(e^{i\theta})}$:

$$\int_0^{2\pi} |P(e^{i\theta})|^2\,\mathrm{d}\theta = \sum_{k=0}^{n}\sum_{l=0}^{n} a_k\overline{a_l}\int_0^{2\pi} e^{i(k-l)\theta}\,\mathrm{d}\theta\,.$$

Or, pour tout $m \in \mathbb{Z}$,

$$\int_0^{2\pi} e^{im\theta}\,\mathrm{d}\theta = \begin{cases} 0 & \text{si } n \neq 0, \\ 2\pi & \text{sinon.} \end{cases}$$

Par conséquent,

$$\int_0^{2\pi} |P(e^{i\theta})|^2\,\mathrm{d}\theta = \sum_{k=0}^{n} a_k\overline{a_k}2\pi = 2\pi\sum_{k=0}^{n}|a_k|^2.$$

Exercice 4

Considérons une primitive F de f. Alors, les primitives de f sont les fonctions $F_C : x \mapsto F(x) + C$ pour une constante $C \in \mathbb{R}$.

Pour tout $C \in \mathbb{R}$,

$$\int_0^1 F_C(t)\,\mathrm{d}t = \int_0^1 F(t)\,\mathrm{d}t + C.$$

Ainsi, une primitive F_C est d'intégrale sur $[0,1]$ nulle si, et seulement si, $C = -\int_0^1 F(t)\,\mathrm{d}t$. Comme il existe une unique valeur C vérifiant cette condition, il existe une unique primitive de f d'intégrale sur $[0,1]$ nulle.

Exercice 5

La fonction $x \mapsto \int_x^{x+T} f(t)\,\mathrm{d}t$ est dérivable (comme intégrale dépendant des bornes d'une fonction continue) et constante donc à dérivée nulle. Or, sa dérivée est $x \mapsto f(x+T) - f(x)$, donc f est T-périodique.

Exercice 6

Par définition la fonction $(f-a)(b-f)$ est positive donc d'intégrale positive. Or,

$$\int_0^1 (f(t)-a)(b-f(t))\,\mathrm{d}t = -\int_0^1 f(t)^2\,\mathrm{d}t + (a+b)\int_0^1 f(t)\,\mathrm{d}t - ab = -\int_0^1 f(t)^2\,\mathrm{d}t - ab.$$

D'où le résultat annoncé.

Exercice 7

Supposons que la fonction f n'admette pas de point fixe, c'est-à-dire que la fonction $g : x \mapsto f(x)-x$ ne s'annule pas. Comme la fonction g est continue et ne s'annule pas, elle est donc de signe constant d'après la théorème des valeurs intermédiaires. Supposons, sans perte de généralité, que g est strictement positive. Par croissance de l'intégrale

$$0 < \int_0^1 g(t)\,\mathrm{d}t = \int_0^1 f(t)\,\mathrm{d}t - \int_0^1 t\,\mathrm{d}t = \frac{1}{2} - \frac{1}{2} = 0,$$

d'où la contradiction.

Exercice 8

1. Il suffit d'appliquer la propriété de croissance de l'intégrale aux inégalités

$$\min_{x\in[a,b]} f(x) \leq f \leq \max_{x\in[a,b]} f(x).$$

2. D'après le théorème des bornes pour la fonction continue f, les valeurs $\min\limits_{x\in[a,b]} f(x)$ et $\max\limits_{x\in[a,b]} f(x)$ sont atteintes par f. Par la propriété des valeurs intermédiaires, $\frac{1}{b-a}\int_a^b f(t)\,\mathrm{d}t$ est aussi une valeur atteinte par f.
3. Le réel $\frac{1}{\int_a^b g(t)\,\mathrm{d}t}\int_a^b f(t)g(t)\,\mathrm{d}t$ appartient à l'intervalle $\left[\min\limits_{x\in[a,b]} f(x), \max\limits_{x\in[a,b]} f(x)\right]$ donc est atteint par f d'après le théorème des valeurs intermédiaires.

Exercice 9

Montrons que la suite converge vers $f(0)$; pour cela, on utilise la continuité de f en 0 et une relation de Chasles pour les intégrales.
Soit $\varepsilon > 0$. Il existe $\eta > 0$ tel que, pour tout $y \in [0,\eta]$, on ait $|f(y)-f(0)| \leq \frac{\varepsilon}{2}$. Considérons $\alpha \in]0,1[$ tel que $2(1-\alpha)\|f\|_\infty \leq \frac{\varepsilon}{2}$ puis $N \in \mathbb{N}$ tel que $\alpha^N \leq \eta$, c'est-à-dire $N \geq \frac{\ln\eta}{\ln\alpha}$.
Alors, pour tout $n \geq N$,

$$\begin{aligned}
|I_n - f(0)| &= \left|\int_0^1 f(x^n)\,\mathrm{d}x - f(0)\right| = \left|\int_0^1 f(x^n) - f(0)\,\mathrm{d}x\right| \\
&\leq \int_0^1 |f(x^n)-f(0)|\,\mathrm{d}x \\
&\leq \int_0^\alpha |f(x^n)-f(0)|\,\mathrm{d}x + \int_\alpha^1 |f(x^n)-f(0)|\,\mathrm{d}x \\
&\leq \alpha\frac{\varepsilon}{2} + 2(1-\alpha)\|f\|_\infty \\
&\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.
\end{aligned}$$

En conclusion, $\lim I_n = f(0)$.

Exercice 10

Supposons par l'absurde de f change de signe au plus $p < n$ fois sur $]a,b[$ en $x_1 < \ldots < x_p$. Alors, la fonction $g : x \mapsto f(x)(x-x_1)\ldots(x-x_p)$ est continue et de signe constant. Par ailleurs, l'intégrale $\int_a^b g(t)\,\mathrm{d}t$ est nulle comme combinaison linéaire des intégrales $\int_a^b t^k f(t)\,\mathrm{d}t$ avec $k \in [\![0,p]\!]$. Par conséquent, g est la fonction nulle (continue, de signe constant et d'intégrale nulle). Ainsi, f est nulle sur $[a,b]\setminus\{x_1,\ldots,x_p\}$ donc sur $[a,b]$ par continuité : ceci contredit l'hypothèse sur f et donc f change au moins n fois de signe sur $]a,b[$.

Exercice 11

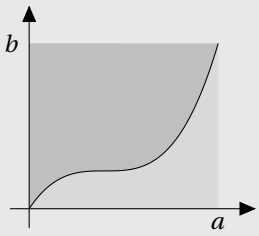
Posons $y = f(x)$ dans la seconde intégrale puis effectuons une intégration par parties

$$\int_0^b f^{-1}(y)\,\mathrm{d}y = \int_0^a x f'(x)\,\mathrm{d}x = \big[x f(x)\big]_0^a - \int_0^a f(t)\,\mathrm{d}t\,.$$

Ainsi,

$$\int_0^a f(x)\,\mathrm{d}x + \int_0^b f^{-1}(y)\,\mathrm{d}y = \big[x f(x)\big]_0^a = ab.$$

La première intégrale correspond à l'aire sous la courbe représentative de f ; la seconde à l'aire sous la courbe représentative de f^{-1}, ce qui est aussi, par symétrie avec la première bissectrice, l'aire entre la courbe représentative de f et l'axe des ordonnées. Ainsi, la somme correspond à l'aire du rectangle de sommets opposés $(0,0)$ et (a,b).



Exercice 12

▹ Si f est la fonction constante égale à c, alors

$$\int_a^b f(t)\cos(nt)\,\mathrm{d}t = c\,\frac{1}{n}\big[\sin(nt)\big]_a^b \to 0.$$

▹ Si f est une fonction en escalier, le résultat découle du premier point en utilisant la relation de Chasles.

▹ Supposons f continue. Soit $\varepsilon > 0$ et φ en escalier sur $[a,b]$ telle que $\max\limits_{t\in[a,b]} |f(t)-\varphi(t)| \leq \varepsilon$. D'après le première point, il existe $N \in \mathbb{N}$ tel que

$$\forall n \geq N, \qquad \left|\int_a^b \varphi(t)\cos(nt)\,\mathrm{d}t\right| \leq \varepsilon.$$

CORRIGÉS

Alors,

$$\begin{aligned}\left|\int_a^b f(t)\cos(nt)\,\mathrm{d}t\right| &\leq \left|\int_a^b (f(t)-\varphi(t))\cos(nt)\,\mathrm{d}t\right| + \left|\int_a^b \varphi(t)\cos(nt)\,\mathrm{d}t\right| \\ &\leq \int_a^b |f(t)-\varphi(t)|.|\cos(nt)|\,\mathrm{d}t + \varepsilon \\ &\leq \varepsilon(b-a)+\varepsilon.\end{aligned}$$

D'après la définition de limite, $\int_a^b f(t)\cos(nt)\,\mathrm{d}t \to 0$.

Exercice 13

Montrons, dans les deux cas, que ces sommes tendent vers une valeur qui dépend de $f(0)$. Fixons $\varepsilon > 0$ et considérons $\eta > 0$ tel que, pour tout $x \in [0,\eta]$, $|f(x)-f(0)| \leq \varepsilon$.

1. Soit $N \in \mathbb{N}$ tel que $\frac{1}{N} \leq \eta$. Alors, pour tout $n \geq N$ et tout $k \in [\![1,n]\!]$, $\frac{k}{n^2} \leq \eta$ donc

$$\left|\frac{1}{n}\sum_{k=1}^{n} f\left(\frac{k}{n^2}\right) - f(0)\right| \leq \frac{1}{n}\sum_{k=1}^{n}\left|f\left(\frac{k}{n^2}\right)-f(0)\right| \leq \frac{1}{n}\sum_{k=1}^{n}\varepsilon = \varepsilon.$$

Ainsi, $\lim \frac{1}{n}\sum_{k=1}^{n} f\left(\frac{k}{n^2}\right) = f(0)$.

2. Soit $N' \in \mathbb{N}$ tel que $\frac{42}{N'} \leq \eta$. Alors, pour tout $n \geq N'$ et tout $k \in [\![1,42n]\!]$, $\frac{k}{n^2} \leq \eta$ donc

$$\left|\frac{1}{n}\sum_{k=1}^{42n} f\left(\frac{k}{n^2}\right) - 42f(0)\right| \leq \frac{1}{n}\sum_{k=1}^{42n}\left|f\left(\frac{k}{n^2}\right)-f(0)\right| \leq \frac{1}{n}\sum_{k=1}^{n}\varepsilon = \varepsilon.$$

Ainsi, $\lim \frac{1}{n}\sum_{k=1}^{42n} f\left(\frac{k}{n^2}\right) = 42f(0)$.

Exercice 14

1. En écrivant

$$u_n = \frac{1}{n}\sum_{k=1}^{n}\frac{1}{\left(\frac{k}{n}\right)^2+1},$$

on reconnaît une somme de Riemann de la fonction $f : x \mapsto \frac{1}{1+x^2}$ continue sur $[0,1]$ (somme des rectangles à droite). Par conséquent,

$$\lim u_n = \int_0^1 \frac{1}{1+x^2}\,\mathrm{d}x = [\arctan x]_0^1 = \frac{\pi}{4}.$$

2. Pour avoir le terme suivant du développement asymptotique de u_n, écrivons la formule de Taylor-Lagrange pour f entre $\frac{k}{n}$ et x :

$$\left|f(x)-f\left(\frac{k}{n}\right)-\left(x-\frac{k}{n}\right)f'\left(\frac{k}{n}\right)\right| \leq \frac{M_2}{2}\left|x-\frac{k}{n}\right|^2,$$

avec $M_2 = \sup_{[0,1]}|f''|$. Intégrons entre $\frac{k-1}{n}$ et $\frac{k}{n}$ en utilisant que la valeur absolue de l'intégrale est inférieure à l'intégrale de la valeur absolue,

$$\left|\int_{\frac{k-1}{n}}^{\frac{k}{n}} f(x)\,\mathrm{d}x - \frac{1}{n}f\left(\frac{k}{n}\right) - \frac{1}{2n^2}f'\left(\frac{k}{n}\right)\right| \leq \frac{M_2}{2n^3}.$$

Sommons maintenant sur k avec l'inégalité triangulaire et la relation de Chasles :

$$\left|\int_0^1 f(x)\,\mathrm{d}x - \frac{1}{n}\sum_{k=1}^{n} f\Big(\frac{k}{n}\Big) - \frac{1}{2n^2}\sum_{k=1}^{n} f'\Big(\frac{k}{n}\Big)\right| \leq \frac{M_2}{2n^2}.$$

Ainsi,

$$u_n = \int_0^1 f(x)\,\mathrm{d}x + \frac{1}{2n^2}\sum_{k=1}^{n} f'\Big(\frac{k}{n}\Big) + O\Big(\frac{1}{n^2}\Big).$$

Or, f' est continue sur $[0,1]$ donc

$$\lim \frac{1}{n}\sum_{k=1}^{n} f'\Big(\frac{k}{n}\Big) = \int_0^1 f'(t)\,\mathrm{d}t = f(1) - f(0) = -\frac{1}{2}.$$

En conclusion,

$$u_n = \int_0^1 f(x)\,\mathrm{d}x - \frac{1}{4n} + o\Big(\frac{1}{n}\Big).$$

Exercice 15

1. L'application $F : x \mapsto \int_a^x f(t)\,\mathrm{d}t$ est continue, strictement croissante (car sa dérivée, f, est strictement positive) donc réalise une bijection de $[a,b]$ dans $[F(a), F(b)] = [0, F(b)]$. Par conséquent, pour tout $k \in [\![0,n]\!]$, il existe un unique antécédent $x_k^{(n)}$ par F à la valeur $\frac{k}{n}F(b)$. Par construction, pour tout $k \in [\![1,n]\!]$,

$$\int_{x_{k-1}^{(n)}}^{x_k^{(n)}} f(t)\,\mathrm{d}t = F(x_k^{(n)}) - F(x_{k-1}^{(n)}) = \frac{1}{n}F(b) = \frac{1}{n}\int_a^b f(t)\,\mathrm{d}t\,.$$

2. D'après la première question, $x_k^{(n)} = F^{-1}\big(\frac{k}{n}F(b)\big)$ pour tout $k \in [\![0,n]\!]$. Ainsi, la somme recherchée est une somme de Riemann

$$\frac{1}{n}\sum_{k=1}^{n} x_k^{(n)} = \frac{1}{n}\sum_{k=1}^{n} F^{-1}\left(\frac{k}{n}F(b)\right) \to \int_0^1 F^{-1}(tF(b))\,\mathrm{d}t\,.$$

Simplifions cette dernière expression, en posant $x = F^{-1}(tF(b))$,

$$\int_0^1 F^{-1}(tF(b))\,\mathrm{d}t = \frac{1}{\int_a^b f(t)\,\mathrm{d}t}\int_a^b x f(x)\,\mathrm{d}x\,.$$

Exercice 16

Notons $F : x \mapsto \int_0^x f(t)\,\mathrm{d}t$. Comme f est de classe $\mathscr{C}^1$, F est de classe $\mathscr{C}^2$; la formule de Taylor avec reste intégral donne alors

$$F(x) = F(0) + F'(0) + \int_0^x (x-t)f'(t)\,\mathrm{d}t,$$

soit en appliquant en 1,

$$\int_0^1 f(t)\,\mathrm{d}t = f(0) + \int_0^1 (1-t)f'(t)\,\mathrm{d}t\,.$$

Or,

$$\frac{1}{2}\min_{[0,1]} f' = \int_0^1 (1-t)\,\mathrm{d}t \min_{[0,1]} f' \leq \int_0^1 (1-t)f'(t)\,\mathrm{d}t \leq \int_0^1 (1-t)\,\mathrm{d}t \max_{[0,1]} f' = \frac{1}{2}\max_{[0,1]} f'.$$

D'après le théorème des valeurs intermédiaires appliqué à la fonction continue $\frac{1}{2}f'$, il existe $c \in [0,1]$ tel que

$$\int_0^1 (1-t)f'(t)\,\mathrm{d}t = \frac{1}{2}f'(c),$$

ce qui donne bien le résultat désiré.

Exercice 17

1. La fonction φ est dérivable sur $[a,b]$ (car f est de classe $\mathscr{C}^{n+1}$) et $\varphi(a)=\varphi(b)$, on peut donc appliquer le théorème de Rolle et il existe $c \in]a,b[$ tel que $\varphi'(c)=0$.
2. Le calcul explicite de la dérivée de φ donne pour $x \in I$

$$\begin{aligned}\varphi'(x) &= \sum_{k=1}^{n} \frac{f^{(k)}(x)}{(k-1)!}(b-x)^{k-1} - \sum_{k=0}^{n} \frac{f^{(k+1)}(x)}{k!}(b-x)^k - \frac{A}{n!}(b-x)^n \\ &= -\frac{f^{(n+1)}(x)}{n!}(b-x)^n - \frac{A}{n!}(b-x)^n.\end{aligned}$$

La condition $\varphi'(c)=0$ entraîne alors $A=-f^{(n+1)}(c)$.
La formule de l'énoncé n'est alors que l'égalité $\varphi(a)=0$.

Exercice 18

Montrons cette propriété par récurrence sur $n \in \mathbb{N} \setminus \{0\}$.

- La propriété est vraie pour $n=1$ par définition de la dérivabilité en a.
- Soit $n \in \mathbb{N} \setminus \{0\}$ telle que la propriété soit vraie au rang n et f une fonction $n+1$ fois dérivable sur I.

Appliquons l'hypothèse de récurrence à f' et posons ε la fonction continue de limite nulle en a définie par

$$\varepsilon(x) = \frac{1}{(x-a)^n}\left(f'(x) - \sum_{k=0}^{n} \frac{f^{(k+1)}(a)}{k!}(x-a)^k\right).$$

Intégrons cette relation entre a et x. Pour tout $\varepsilon > 0$, il existe $\eta > 0$ tel que

$$\forall x \in I, \qquad |x-a| \leq \eta \quad \Rightarrow \quad |\varepsilon(x)| \leq \varepsilon.$$

Ainsi, pour tout $x \in I$ tel que $|x-a| \leq \eta$,

$$\left|\frac{1}{(x-a)^{n+1}} \int_a^x \varepsilon(t)(t-a)^n\,\mathrm{d}t\right| \leq \frac{1}{n+1}\varepsilon,$$

d'où le résultat.

Corrigés des exercices d'approfondissement du chapitre 21

Exercice A

▹ Montrons dans un premier temps que, pour tout $P \in \mathbb{C}[X]$, il existe $Q \in \mathbb{C}[X]$ de même degré que P tel que $(Q(x)e^{ix})' = P(x)e^{ix}$. Procédons par récurrence sur le degré de P.

- Le résultat est immédiat si P est un polynôme constant.

• Soit n tel que le résultat soit vrai pour les polynômes de degré n. Considérons P de degré $n+1$ et Q le polynôme associé à P' par hypothèse de récurrence. Par intégration par parties, pour tout $x \in \mathbb{R}$,

$$\begin{aligned}\int_0^x P(t)e^{it}\,\mathrm{d}t &= \left[-iP(t)e^{it}\right]_0^x + i\int_0^x P'(t)e^{it}\,\mathrm{d}t\\ &= -iP(X)e^{ix} + iP(0) + i\left[Q(t)e^{it}\right]_0^x\\ &= (-iP(x)+iQ(x))e^{ix} + i(P(0)-Q(0)).\end{aligned}$$

Ainsi, $x \mapsto P(x)e^{ix}$ admet une primitive de la forme $x \mapsto (-iP(x)+iQ(x))e^{ix}$ et $\deg(-iP+iQ) = \deg P$ ce qui achève la récurrence.

▹ Revenons à l'exercice : avec le polynôme Q correspondant à P d'après le premier point, on obtient

$$\int_0^x P(t)e^{it}\,\mathrm{d}t = \left[Q(t)e^{it}\right]_0^x = Q(x)e^{ix} - Q(0).$$

Un réel x appartient à l'ensemble $\mathscr{Z}$ si, et seulement si, $Q(x) = Q(0)e^{-ix}$. Cette dernière condition entraîne $|Q(x)|^2 = |Q(0)|^2$ qui se réécrit à son tour, en notant $Q = \sum_{k=0}^{n} a_k X^k$,

$$\sum_{k=0}^{n}\sum_{l=0}^{n} a_k\overline{a_l}x^{k+l} = |Q(0)|^2.$$

Cette condition est polynomiale donc admet un nombre fini de solutions ce qui entraîne, par conséquent, le caractère fini de $\mathscr{Z}$.

Exercice B

Nous allons montrer le résultat équivalent suivant

$$\lim_n (n+2)(n+1)\int_0^1 (x^n - x^{n+1})f(x)\,\mathrm{d}x = f(1).$$

En effet, $\int_0^1 (x^n - x^{n+1})\,\mathrm{d}x = \frac{1}{(n+1)(n+2)}$ donc

$$f(1) = (n+1)(n+2)\int_0^1 (x^n - x^{n+1})f(1)\,\mathrm{d}x\,.$$

Par suite,

$$\left|(n+2)(n+1)\int_0^1 (x^n - x^{n+1})f(x)\,\mathrm{d}x - f(1)\right| \leq (n+2)(n+1)\int_0^1 (x^n - x^{n+1})|f(x)-f(1)|\,\mathrm{d}x\,.$$

Fixons $\varepsilon > 0$ et considérons, par continuité de f en 1, $\eta > 0$ tel que, pour tout $x \in [\eta, 1]$, $|f(x)-f(1)| \leq \varepsilon$. Ainsi,

$$\left|(n+2)(n+1)\int_0^1 (x^n - x^{n+1})f(x)\,\mathrm{d}x - f(1)\right| \leq (n+2)(n+1)\int_0^\eta (x^n - x^{n+1})|f(x)-f(1)|\,\mathrm{d}x + \varepsilon.$$

Majorons f par M sur $[0,1]$,

$$(n+2)(n+1)\int_0^\eta (x^n - x^{n+1})|f(x)-f(1)|\,\mathrm{d}x \leq 2M(n+2)(n+1)\int_0^\eta (1-x)x^n\,\mathrm{d}x \leq 2M(n+2)\eta^{n+1}.$$

Ce dernier terme tend vers 0 donc à partir d'un certain rang N, il est inférieur à ε. Donc pour $n \geq N$,

$$\left|(n+2)(n+1)\int_0^1 (x^n - x^{n+1})f(x)\,\mathrm{d}x - f(1)\right| \leq 2\varepsilon,$$

ce qui est la définition de la limite annoncée par l'énoncé.

Exercice C

Commençons par remarquer le développement suivant pour des complexes $a_1, \dots, a_n$

$$\left(\sum_{k=1}^{n} a_k\right)^2 = \sum_{k=1}^{n} a_k \sum_{l=1}^{n} a_l = \sum_{1\le k<l\le n} a_k a_l + \sum_{k=1}^{n} a_k^2 + \sum_{1\le l<k\le n} a_k a_l$$
$$= 2\sum_{1\le k<l\le n} a_k a_l + \sum_{k=1}^{n} a_k^2.$$

1. D'après la remarque liminaire,

$$\frac{1}{n^2}\sum_{1\le k<l\le n} f\left(\frac{k}{n}\right) f\left(\frac{l}{n}\right) = \frac{1}{2}\left(\frac{1}{n}\sum_{k=1}^{n} f\left(\frac{k}{n}\right)\right)^2 - \frac{1}{2n^2}\sum_{k=1}^{n} f\left(\frac{k}{n}\right)^2.$$

Comme f et f^2 sont des fonctions continues, on a

$$\frac{1}{n}\sum_{k=1}^{n} f\left(\frac{k}{n}\right) \to \int_0^1 f(t)\,\mathrm{d}t, \qquad \frac{1}{n}\sum_{k=1}^{n} f\left(\frac{k}{n}\right)^2 \to \int_0^1 f(t)^2\,\mathrm{d}t\,.$$

En combinant ces résultats,

$$\frac{1}{n^2}\sum_{1\le k<l\le n} f\left(\frac{k}{n}\right) f\left(\frac{l}{n}\right) \to \frac{1}{2}\left(\int_0^1 f(t)^2\,\mathrm{d}t\right)^2.$$

2. Il suffit de remarquer que

$$\frac{1}{n^2}\sum_{1\le k\le l\le n} f\left(\frac{k}{n}\right) f\left(\frac{l}{n}\right) = \frac{1}{n^2}\sum_{1\le k<l\le n} f\left(\frac{k}{n}\right) f\left(\frac{l}{n}\right) + \frac{1}{n^2}\sum_{k=1}^{n} f\left(\frac{k}{n}\right)^2.$$

En passant à la limite, on trouve que la limite est la même.

Exercice D

Quitte a effectuer une rotation centrée sur l'origine, on peut supposer que les points A_k ont pour affixe les complexes $e^{\frac{2ik\pi}{n}}$. Alors,

$$\frac{1}{n}\sum_{k=2}^{n} A_1A_k = \frac{1}{n}\sum_{k=1}^{n}\left|e^{\frac{2ik\pi}{n}} - 1\right| = \frac{1}{n}\sum_{k=1}^{n}\left|e^{\frac{ik\pi}{n}} - e^{-\frac{ik\pi}{n}}\right| = \frac{1}{n}\sum_{k=1}^{n} 2\sin\frac{k\pi}{n}.$$

On reconnaît alors une somme de Riemann associée à la fonction continue $x \mapsto 2\sin(x\pi)$. La limite est donc

$$\int_0^1 2\sin(\pi x)\,\mathrm{d}x = \frac{2}{\pi}\int_0^{\pi} \sin u\,\mathrm{d}u = \frac{4}{\pi}.$$

Exercice E

▹ Commençons par déterminer x_n. La fonction f_n est dérivable (comme somme de fonctions dérivables) et, pour tout $x \in \mathbb{R}_+$,

$$f_n'(x) = \sum_{k=1}^{n} \cos(kx).$$

Pour tout $x \notin 2\pi\mathbb{Z}$, cette dernière somme se calcule comme partie réelle d'une somme géométrique complexe de raison différente de 1. Après calculs, on obtient

$$f_n'(x) = \frac{\sin\frac{nx}{2}}{\sin\frac{x}{2}}\cos\frac{nx}{2}.$$

Cette expression permet d'obtenir le tableau de signe suivant pour $f_n'(x)$ au voisinage de 0 :

x	0		$\frac{\pi}{n}$		$\frac{2\pi}{n}$
$f_n'(x)$	n	$+$	0	$-$	0

Par conséquent, le premier maximum local de f_n sur $\mathbb{R}_+$ est atteint en $x_n = \frac{\pi}{n}$.
▹ Remarquons que $f_n(x_n)$ est une somme de Riemann (rectangle à droite) pour la fonction continue

$$f : x \mapsto \begin{cases} 0 & \text{si } x = 0, \\ \frac{\sin x}{x} & \text{sinon.} \end{cases}$$

Plus précisément,

$$f_n(x_n) = \sum_{k=1}^{n} \frac{\sin\frac{k\pi}{n}}{k} = \frac{\pi}{n}\sum_{k=1}^{n}\frac{\sin\frac{k\pi}{n}}{\frac{k\pi}{n}}.$$

En conclusion, $f_n(x_n) \to \int_0^\pi \frac{\sin t}{t}\,\mathrm{d}t$.

Exercice F

▹ Remarquons tout d'abord que pour $f \in \mathscr{C}^0([0,1],\mathbb{R}_+)$,

$$\int_0^1 f(\sqrt[n]{t})\,\mathrm{d}t = \int_0^1 f(y)ny^{n-1}\,\mathrm{d}y \leq n\int_0^1 f(y)\,\mathrm{d}y.$$

Ainsi, $n \in E$ et $\inf E \leq n$.
▹ Pour la fonction $f : x \mapsto x^m$ avec $m > 0$ et pour $c \in E$, on a

$$\frac{1}{\frac{m}{n}+1} \leq c\,\frac{1}{m+1}.$$

Ainsi, $\inf E \geq \frac{m+1}{\frac{m}{n}+1}$ pour tout $m > 0$. En faisant tendre m vers $+\infty$, on obtient $\inf E \geq n$.
En conclusion, $\inf E = n$.

Corrigés des problèmes du chapitre 21

Problème 1

1. a. ▹ Écrivons $f(x)\operatorname{cotan}(x) = \frac{f(x)}{x}\,x\operatorname{cotan}(x)$ puis remarquons que

$$\frac{f(x)}{x} \underset{x\to 0}{\longrightarrow} f'(0), \qquad x\operatorname{cotan}(x) \underset{x\to 0}{\longrightarrow} 1$$

car $\operatorname{cotan}(x) \sim \frac{1}{x}$. Par conséquent, la fonction $x \mapsto f(x)\operatorname{cotan}(x)$ est prolongeable par continuité en 0.
▹ On procède de même en π en partant de l'écriture

$$f(x)\operatorname{cotan}(x) = \frac{f(x)}{\pi - x}(\pi - x)\operatorname{cotan}(x).$$

CORRIGÉS

b. D'après les règles générales de manipulation des limites, la quantité

$$(f^2(x)+f^2(\pi-x))\operatorname{cotan}(x)=f(x)(f(x)\operatorname{cotan}(x))+f(\pi-x)(f(\pi-x)\operatorname{cotan}(\pi-x))$$

converge vers $f(0)f'(0)+f(\pi)f'(\pi)=0$ en 0.

2. Il suffit de remarquer que la dérivée de $x\mapsto f(x)\operatorname{cotan}(x)$ est donnée par

$$\begin{aligned}2f(x)f'(x)\operatorname{cotan}(x)+f(x)^2\operatorname{cotan}'(x)&=2f(x)f'(x)\operatorname{cotan}(x)-f(x)^2(1+\operatorname{cotan}(x)^2\\&=f'^2(x)-f^2(x)-(f(x)\operatorname{cotan}(x)-f'(x))^2.\end{aligned}$$

3. En faisant tendre x vers 0 dans l'égalité précédente, on obtient que

$$\int_0^\pi (f'^2(t)-f^2(t))\,\mathrm{d}t=\int_0^\pi (f'(t)-f(t)\operatorname{cotan}(t))^2\,\mathrm{d}t\,.$$

Ce dernier terme est positif comme intégrale d'une fonction positive (avec les bornes dans le bon ordre) donc

$$\int_0^\pi f'^2(t)\,\mathrm{d}t\geq\int_0^\pi f^2(t)\,\mathrm{d}t\,.$$

4. a. D'après l'identité de la question précédente, l'égalité se réécrit $\int_0^\pi (f'(t)-f(t)\operatorname{cotan}(t))^2\,\mathrm{d}t=0$. Or, la fonction $x\mapsto(f'(x)-f(x)\operatorname{cotan}(x))^2$ est continue positive donc est identiquement nulle.

b. Résolvons l'équation différentielle ainsi obtenue en remarquant qu'une primitive de cotan sur $]0,\pi[$ est $\ln\circ\sin$. Ainsi, il existe $C\in\mathbb{R}$ tel que $f:x\mapsto C\,e^{\ln(\sin(x))}=C\sin(x)$.

Problème 2

Partie A – Noyau de Jackson

1. ▹ Commençons par vérifier que U_n est une fonction réelle car on a, pour tout $t\in[-1,1]$

$$\overline{\sum_{k=0}^{n-1}\sum_{l=-k}^{k}e^{il\theta}}=\sum_{k=0}^{n-1}\sum_{l=-k}^{k}e^{-il\theta}=\sum_{k=0}^{n-1}\sum_{m=-k}^{k}e^{im\theta}.$$

Par conséquent, U_n est une fonction réelle.

▹ De même, la fonction U_n est paire (il suffit de changer l en $-l$ dans la seconde somme comme dans l'étape intermédiaire du calcul précédent).

▹ Enfin, U_n est continue et 2π-périodique comme somme et produit de fonctions continues et 2π-périodiques.

2. Utilisons quelques calculs sur les sommes géométriques. Tout d'abord, pour tout $\theta\neq 0[\pi]$

$$\sum_{l=-k}^{k}e^{il\theta}=\frac{e^{-ik\theta}-e^{i(k+1)\theta}}{1-e^{i\theta}}.$$

Par conséquent

$$\begin{aligned}U_n(\theta)&=\frac{1}{1-e^{i\theta}}\Big(\sum_{k=0}^{n-1}e^{-ik\theta}-\sum_{k=0}^{n-1}e^{i(k+1)\theta}\Big)\\&=\frac{1}{1-e^{i\theta}}\Big(\frac{1-e^{-in\theta}}{1-e^{-i\theta}}-\frac{e^{i\theta}-e^{i(n+1)\theta}}{1-e^{i\theta}}\Big)\\&=\frac{(2-2\cos(n\theta))}{|1-e^{i\theta}|^2}=\Big(\frac{\sin\frac{n\theta}{2}}{\sin\frac{\theta}{2}}\Big)^2.\end{aligned}$$

3. D'après la première question, la fonction U_n^2 est continue, positive, non nulle sur $[-\pi,\pi]$ donc son intégrale sur cet intervalle est strictement positive. La constante c_n recherchée est alors l'inverse de cette intégrale.
4. a. ▹ D'après la concavité de la fonction $t \longmapsto \sin\frac{t}{2}$ sur $[0,\pi]$, la courbe est au-dessous de sa tangente à l'origine, $\sin\frac{t}{2} \leq \frac{t}{2}$ (on aurait également pu étudier les variations de la fonction $t \longmapsto \sin\frac{t}{2} - \frac{t}{2}$).
▹ Par définition de c_n,

$$\begin{aligned} 1 &= 2c_n \int_0^{\pi} \Big(\frac{\sin\frac{nt}{2}}{\sin\frac{t}{2}}\Big)^4 \mathrm{d}t \geq 2^5 c_n \int_0^{\pi} \Big(\frac{\sin\frac{nt}{2}}{t}\Big)^4 \mathrm{d}t \\ &\geq 4n^3 c_n \int_0^{\frac{n\pi}{2}} \Big(\frac{\sin u}{u}\Big)^4 \mathrm{d}u\,. \end{aligned}$$

Ainsi, $n^3 c_n \leq \frac{1}{4\int_0^{\frac{\pi}{2}} \left(\frac{\sin u}{u}\right)^4 \mathrm{d}u}$ donc $n^3 c_n = O(1)$.
b. ▹ D'après la concavité de la fonction $t \longmapsto \sin\frac{t}{2}$ sur $[0,\pi]$, la courbe représentative est au-dessus de sa corde joignant les points d'abscisse 0 et π, $\frac{t}{\pi} \leq \sin\frac{t}{2}$ (on aurait également pu étudier les variations de la fonction $t \longmapsto \sin\frac{t}{2} - \frac{t}{\pi}$).
▹ Par conséquent,

$$\begin{aligned} I_n &= 2c_n \int_0^{\pi} t\Big(\frac{\sin\frac{nt}{2}}{\sin\frac{t}{2}}\Big)^4 \mathrm{d}t \leq 2\pi^4 c_n \int_0^{\pi} t\Big(\frac{\sin\frac{nt}{2}}{t}\Big)^4 \mathrm{d}t \\ &\leq \frac{\pi^4}{2} n^2 c_n \int_0^{\frac{n\pi}{2}} \frac{\sin^4 u}{u^3}\,\mathrm{d}u\,. \end{aligned}$$

Montrons que la dernière intégrale est bornée indépendamment de n :

$$\begin{aligned} \int_0^{\frac{n\pi}{2}} \frac{\sin^4 u}{u^3}\,\mathrm{d}u &= \int_0^{\frac{\pi}{2}} \frac{\sin^4 u}{u^3}\,\mathrm{d}u + \int_{\frac{\pi}{2}}^{\frac{n\pi}{2}} \frac{\sin^4 u}{u^3}\,\mathrm{d}u \\ &\leq \int_0^{\frac{\pi}{2}} \frac{\sin^4 u}{u^3}\,\mathrm{d}u + \int_{\frac{\pi}{2}}^{\frac{n\pi}{2}} \frac{1}{u^3}\,\mathrm{d}u \\ &\leq \int_0^{\frac{\pi}{2}} \frac{\sin^4 u}{u^3}\,\mathrm{d}u + \frac{4}{3\pi^2}. \end{aligned}$$

Ainsi, $I_n = O(n^2 c_n)$.
c. D'après les questions précédentes, $I_n = O(n^2 c_n)$ et $n^2 c_n = O(\frac{1}{n})$; par transitivité de la relation de domination, on obtient $I_n = O(\frac{1}{n})$.

Partie B – Théorème de Jackson-Weierstrass

1. a. Utilisons que $\int_{-\pi}^{\pi} J_n(t)\,\mathrm{d}t = 1$ pour rassembler les deux termes

$$\begin{aligned} \Big|f(\cos\theta) - \int_{-\pi}^{\pi} f(\cos(\theta-t))J_n(t)\,\mathrm{d}t\Big| &= \Big|f(\cos\theta)\int_{-\pi}^{\pi} J_n(t)\,\mathrm{d}t - \int_{-\pi}^{\pi} f(\cos(\theta-t))J_n(t)\,\mathrm{d}t\Big| \\ &= \Big|\int_{-\pi}^{\pi} (f(\cos\theta) - f(\cos(\theta-t)))J_n(t)\,\mathrm{d}t\Big| \\ &\leq \int_{-\pi}^{\pi} |f(\cos\theta) - f(\cos(\theta-t))|J_n(t)\,\mathrm{d}t \\ &\leq \|f'\|_\infty I_n, \end{aligned}$$

en utilisant l'inégalité des accroissements finis pour la fonction $x \mapsto f(\cos x)$ qui est de classe $\mathscr{C}^1$ sur $[-\pi,\pi]$ comme composée de fonctions de classe $\mathscr{C}^1$. On conclut désormais en rappelant que $I_n = O(\frac{1}{n})$.
b. Commençons par effectuer successivement les changements de variables $t \mapsto \theta - t$ et $t \mapsto -t$

$$\begin{aligned}\int_{-\pi}^{\pi} f(\cos(\theta-t))J_n(t)\,\mathrm{d}t &= \int_{\theta-\pi}^{\theta+\pi} f(\cos(u))J_n(\theta-u)\,\mathrm{d}u\\ &= \int_{-\pi}^{\pi} f(\cos(t))J_n(\theta-t)\,\mathrm{d}t\\ &= \int_{-\pi}^{\pi} f(\cos(t))J_n(\theta+t)\,\mathrm{d}t,\end{aligned}$$

car les fonctions $x \mapsto f(\cos(x))$ et J_n sont 2π-périodiques. Pour avoir le résultat annoncé, il suffit de considérer la demi-somme des expressions obtenues.
c. ▹ Montrons l'existence en exhibant un polynôme avec la formule de De Moivre ; pour tout $\theta \in \mathbb{R}$,

$$\begin{aligned}\cos(n\theta) = \mathfrak{Re}((e^{i\theta})^n) &= \mathfrak{Re}((\cos\theta + i\sin\theta)^n)\\ &= \sum_{l=0}^{\lfloor \frac{n}{2}\rfloor} \binom{n}{2l}(-1)^l \cos^{n-2l}\theta \sin^{2l}\theta\\ &= \sum_{l=0}^{\lfloor \frac{n}{2}\rfloor} \binom{n}{2l}(-1)^l \cos^{n-2l}\theta (1-\cos^2\theta)^l.\end{aligned}$$

Donc le polynôme $T_n(X) = \sum_{l=0}^{\lfloor \frac{n}{2}\rfloor} \binom{n}{2l}(-1)^l X^{n-2l}(1-X^2)^l$ convient. Son degré est n et son coefficient dominant est $\sum_{l=0}^{\lfloor \frac{n}{2}\rfloor} \binom{n}{2l} = \sum_{l=0}^{\lfloor \frac{n-1}{2}\rfloor} \binom{n}{2l+1} = 2^{n-1}$.
▹ Montrons l'unicité par l'absurde en supposons qu'il existe une autre polynôme $S_n \in \mathbb{R}_n[X]$ vérifiant la condition de l'énoncé. Alors, $T_n - S_n$ est nul sur l'intervalle $[-1,1]$ donc uniformément nul.
d. D'après les questions précédentes,

$$\int_{-\pi}^{\pi} f(\cos(\theta-t))J_n(t)\,\mathrm{d}t = \int_{-\pi}^{\pi} f(\cos(t))\frac{1}{2}\big(J_n(\theta+t)+J_n(\theta-t)\big)\,\mathrm{d}t\,.$$

Écrivons désormais $J_n(u)$ sous la forme

$$\sum_{k=0}^{2n} a_k(e^{iku}+e^{-iku}) = 2\sum_{k=0}^{2n} a_k\cos(ku)$$

avec des coefficients a_k réels (écriture qui est obtenue en développant l'expression de J_n et en regroupant les termes par parité). Par conséquent,

$$\begin{aligned}\int_{-\pi}^{\pi} f(\cos(\theta-t))J_n(t)\,\mathrm{d}t &= \sum_{k=0}^{2n} a_k \int_{-\pi}^{\pi} f(\cos(t))(\cos(k(\theta+t))+\cos(k(\theta+t)))\,\mathrm{d}t\\ &= \sum_{k=0}^{2n} 2a_k\cos(k\theta)\int_{-\pi}^{\pi} f(\cos(t))\cos(kt)\,\mathrm{d}t\,.\end{aligned}$$

Concluons en rappelant que $\cos(k\theta) = T_k(\cos\theta)$ et en posant

$$P_n(X) = \sum_{k=0}^{2n} 2a_k \int_{-\pi}^{\pi} f(\cos(t))\cos(kt)\,\mathrm{d}t\; T_k(X).$$

e. Pour tout $n \in \mathbb{N}^*$,

$$\|f - P_n\|_\infty = \sup_{\theta \in \mathbb{R}} |f(\cos(\theta)) - P_n(\cos(\theta))| \leq \frac{\kappa \|f'\|_\infty}{n}.$$

Donc, par encadrement, $\|f - P_n\|_\infty \to 0$. En conclusion, f est limite uniforme de la suite de polynômes $(P_n)_n$ sur $[-1,1]$.

2. a. Pour tous $0 < h < h'$, l'ensemble

$$\{|f(x) - f(y)|,\ x, y \in [-1,1] \text{ avec } |x - y| \leq h\}$$

est inclus dans l'ensemble

$$\{|f(x) - f(y)|,\ x, y \in [-1,1] \text{ avec } |x - y| \leq h'\},$$

donc $\omega(h) \leq \omega(h')$.

b. Soit $x < y \in [-1,1]$ tels que $|x - y| \leq |\lambda| h$. Introduisons les réels $x_k = x + kh$ pour $k \in [\![0, \lfloor \frac{y-x}{h} \rfloor]\!]$. Alors, par inégalité triangulaire,

$$\begin{aligned} |f(y) - f(x)| &\leq \sum_{k=0}^{\lfloor \frac{y-x}{h} \rfloor - 1} |f(x_{k+1}) - f(x_k)| + |f(y) - f(x_{\lfloor \frac{y-x}{h} \rfloor})| \\ &\leq \Big(\Big\lfloor \frac{y-x}{h} \Big\rfloor + 1\Big)\omega(h) \leq (|\lambda| + 1)\omega(h) \end{aligned}$$

d'où, par passage à la borne supérieure, $\omega(|\lambda| h) \leq (|\lambda| + 1)\omega(h)$.

c. Reprenons la question précédente en remplaçant l'inégalité des accroissements finis par la majoration avec ω

$$\begin{aligned} \left| f(\cos\theta) - \int_{-\pi}^{\pi} f(\cos(\theta - t)) J_n(t)\,\mathrm{d}t \right| &= \left| \int_{-\pi}^{\pi} (f(\cos\theta) - f(\cos(\theta - t))) J_n(t)\,\mathrm{d}t \right| \\ &\leq \int_{-\pi}^{\pi} |f(\cos\theta) - f(\cos(\theta - t))| J_n(t)\,\mathrm{d}t \\ &\leq \int_{-\pi}^{\pi} \omega(|\cos\theta - \cos(\theta - t)|) J_n(t)\,\mathrm{d}t \\ &\leq \int_{-\pi}^{\pi} \omega(t) J_n(t)\,\mathrm{d}t, \end{aligned}$$

en utilisant conjointement, l'inégalité des accroissements finis pour cos et la croissance de la fonction ω. Écrivons ensuite $\omega(t) = \omega\big(n t \frac{1}{n}\big) \leq (n t + 1)\omega(\frac{1}{n})$.

$$\begin{aligned} \left| f(\cos\theta) - \int_{-\pi}^{\pi} f(\cos(\theta - t)) J_n(t)\,\mathrm{d}t \right| &\leq \omega\Big(\frac{1}{n}\Big) \int_{-\pi}^{\pi} (n t + 1) J_n(t)\,\mathrm{d}t \\ &\leq \omega\Big(\frac{1}{n}\Big)(n I_n + 1). \end{aligned}$$

D'où le résultat, en rappelant que la suite $(n I_n)_n$ est bornée.

d. Soit $\varepsilon > 0$. La fonction f est continue sur le segment $[-1,1]$ donc uniformément continue (théorème de Heine). Ainsi, il existe $\eta > 0$ tel que $\omega(\eta) \leq \varepsilon$.

Par conséquent, il existe un rang N au-delà duquel $\omega\big(\frac{1}{n}\big) \leq \varepsilon$ donc

$$\forall n \geq N, \qquad \left| f(\cos\theta) - \int_{-\pi}^{\pi} f(\cos(\theta - t)) J_n(t)\,\mathrm{d}t \right| \leq \kappa' \varepsilon.$$

On conclut comme à la question précédente.

Problème 3

Partie A – Exemples

1. Par définition, $f^*(0)=0$. En procédant par intégration par parties, on obtient

$$\forall x>0, \qquad f^*(x)=\arctan x-\frac{1}{2x}\ln(1+x^2).$$

Comme $\ln(1+x^2)=2\ln x+\ln(1+\frac{1}{x^2})\underset{+\infty}{=}o(x)$, on en déduit que $f(x)\underset{x\to+\infty}{\longrightarrow}\frac{\pi}{2}$ et $f^*(x)\underset{x\to+\infty}{\longrightarrow}\frac{\pi}{2}$.

2. Par définition, $f^*(0)=0$. Utilisons la relation de Chasles pour ramener le calcul de f^* à des intégrales sur des intervalles sur lesquelles la fonction sin est de signe constant

$$\begin{aligned}
f^*(x)&=\frac{1}{x}\Big(\sum_{k=0}^{\lfloor\frac{x}{\pi}\rfloor-1}\int_{k\pi}^{(k+1)\pi}|\sin t|\,\mathrm{d}t+\int_{\lfloor\frac{x}{\pi}\rfloor\pi}^{x}|\sin t|\,\mathrm{d}t\Big)\\
&=\frac{1}{x}\Big(\sum_{k=0}^{\lfloor\frac{x}{\pi}\rfloor-1}\int_{0}^{\pi}|\sin t|\,\mathrm{d}t+\int_{\lfloor\frac{x}{\pi}\rfloor\pi}^{x}|\sin t|\,\mathrm{d}t\Big)\\
&=\frac{1}{x}\Big(\sum_{k=0}^{\lfloor\frac{x}{\pi}\rfloor-1}\int_{0}^{\pi}\sin t\,\mathrm{d}t+(-1)^{\lfloor\frac{x}{\pi}\rfloor}\int_{\lfloor\frac{x}{\pi}\rfloor\pi}^{x}\sin t\,\mathrm{d}t\Big)\\
&=\frac{1}{x}\Big(2\Big\lfloor\frac{x}{\pi}\Big\rfloor+1-(-1)^{\lfloor\frac{x}{\pi}\rfloor}\cos x\Big).
\end{aligned}$$

D'après les inégalités définissant la partie entière, $\lfloor\frac{x}{\pi}\rfloor\underset{+\infty}{\sim}\frac{x}{\pi}$ et on conclut que $f^*(x)\underset{x\to+\infty}{\longrightarrow}\frac{2}{\pi}$.

3. a. $f\big(\frac{\pi}{2}+2n\pi\big)=\sqrt{\frac{\pi}{2}+2n\pi}\to+\infty$ donc la fonction f n'est pas bornée sur $\mathbb{R}_+$.

b. Commençons par trouver une expression plus maniable de f^* en utilisant la relation de Chasles en 1 et une intégration par parties (en dérivant la racine).

$$f^*(x)=\frac{1}{x}\Big(\int_0^1 f(t)\,\mathrm{d}t+\big[\sqrt{t}(-\cos t)\big]_1^x-\int_1^x\frac{1}{2\sqrt{t}}(-\cos t)\,\mathrm{d}t\Big).$$

D'une part, $\lim\limits_{x\to\infty}\frac{1}{x}\Big(\int_0^1 f(t)\,\mathrm{d}t+\big[\sqrt{t}(-\cos t)\big]_1^x\Big)=0$ (produit d'un terme de limite nulle et d'un terme borné). D'autre part,

$$0\le\Big|\frac{1}{x}\int_1^x\frac{1}{2\sqrt{t}}(-\cos t)\,\mathrm{d}t\Big|\le\frac{1}{x}\int_1^x\frac{1}{2\sqrt{t}}\,\mathrm{d}t=\frac{\sqrt{x}-\sqrt{1}}{x}$$

et ce dernier terme tend vers 0.
Par encadrement, f^* tend vers 0 quand x tend vers $+\infty$.

Partie B – Propriétés de f^*

1. ▹ La fonction f^* est de classe $\mathscr{C}^1$ sur $\mathbb{R}_+^*$ comme produit de deux fonctions de classe $\mathscr{C}^1$ sur $\mathbb{R}_+^*$.
▹ Fixons $\varepsilon>0$. Comme f est continue en 0, il existe $\eta>0$ tel que

$$\forall x\in[0,\eta[, \qquad |f(x)-f(0)|\le\varepsilon.$$

Ainsi, pour tout $\varepsilon>0$ fixé, il existe η tel que pour tout $x\in]0,\eta[$:

$$|f^*(x)-f^*(0)|=\frac{1}{x}\Big|\int_0^x f(t)-f(0)\,\mathrm{d}t\Big|\le\frac{1}{x}\int_0^x|f(t)-f(0)|\,\mathrm{d}t\le\varepsilon$$

Par conséquent, f^* est continue en 0.

2. Soit $\varepsilon > 0$. Comme $\lim\limits_{x\to+\infty} f(x) = \ell$, il existe $A > 0$ tel que

$$\forall x \geq A, \qquad |f(x)-\ell| \leq \varepsilon.$$

Ainsi, pour tout $x \geq A$,

$$\begin{aligned}
|f^*(x)-\ell| &= \frac{1}{x}\left|\int_0^x (f(t)-\ell)\,\mathrm{d}t\right| \\
&\leq \frac{1}{x}\int_0^A |f(t)-\ell|\,\mathrm{d}t + \frac{1}{x}\int_A^x |f(t)-\ell|\,\mathrm{d}t \\
&\leq \frac{1}{x}\int_0^A |f(t)-\ell|\,\mathrm{d}t + \varepsilon.
\end{aligned}$$

Or, $\lim\limits_{x\to+\infty} \frac{1}{x}\int_0^A |f(t)-\ell|\,\mathrm{d}t = 0$ donc il existe $B > 0$ tel que

$$\forall x \geq B, \qquad \frac{1}{x}\int_0^A |f(t)-\ell|\,\mathrm{d}t \leq \varepsilon.$$

Par conséquent, pour tout $x \geq \max(A, B)$, $|f^*(x)-\ell| \leq 2\varepsilon$ et donc $\lim\limits_{x\to+\infty} f^*(x) = \ell$.

3. L'hypothèse de l'énoncé se traduit par $\lim\limits_{x\to+\infty} f(x)-\alpha x-\beta = 0$. On utilise alors la question précédente avec la fonction $g : x \mapsto f(x)-\alpha x-\beta$: on obtient que $\lim\limits_{x\to+\infty} g^*(x) = 0$. Or, pour tout x, $g^*(x) = f^*(x)-\frac{\alpha}{2}x-\beta$. En conclusion, la courbe représentative de la fonction f^* admet l'asymptote d'équation $y = \frac{\alpha}{2}x+\beta$ au voisinage de $+\infty$.

4. Notons P_n la partie régulière du développement limité de f à l'ordre n en 0 ; alors $\int_0^x f(t)\,\mathrm{d}t$ admet un développement limité à l'ordre $n+1$ en 0 de partie régulière $\int_0^x P_n(t)\,\mathrm{d}t$. Ainsi, f^* admet un développement limité à l'ordre n en 0 de partie régulière $\frac{1}{x}\int_0^x P_n(t)\,\mathrm{d}t = P_n^*(x)$.

5. Commençons par effectuer le changement de variables $t = xu$: alors, pour tout $x > 0$,

$$f^*(x) = \int_0^1 f(xu)\,\mathrm{d}u.$$

Ainsi, pour tous $x \geq x'$,

$$f^*(x)-f^*(x') = \int_0^1 (f(xu)-f(x'u))\,\mathrm{d}u.$$

Supposons sans perte de généralité que f est croissante, alors le second membre de l'identité précédente est positif (comme intégrale d'une fonction positive) : par conséquent, f^* est croissante.
En conclusion, f^* est monotone de même sens de variation que f.

6. Commençons par trouver une expression plus maniable de f^* en utilisant la relation de Chasles aux multiples entiers de T

$$\begin{aligned}
f^*(x) &= \frac{1}{x}\Big(\sum_{k=0}^{\lfloor \frac{x}{T}\rfloor-1}\int_{kT}^{(k+1)T} f(t)\,\mathrm{d}t + \int_{\lfloor \frac{x}{T}\rfloor T}^{x} f(t)\,\mathrm{d}t\Big) \\
&= \frac{1}{x}\Big(\sum_{k=0}^{\lfloor \frac{x}{T}\rfloor-1}\int_0^T f(t+kT)\,\mathrm{d}t + \int_{\lfloor \frac{x}{T}\rfloor T}^{x} f(t)\,\mathrm{d}t\Big) \\
&= \frac{1}{x}\left\lfloor \frac{x}{T}\right\rfloor\int_0^T f(t)\,\mathrm{d}t + \frac{1}{x}\int_{\lfloor \frac{x}{T}\rfloor T}^{x} f(t)\,\mathrm{d}t.
\end{aligned}$$

Or, $\left|\frac{1}{x}\int_{\lfloor\frac{x}{T}\rfloor T}^{x} f(t)\,\mathrm{d}t\right| \leq \frac{1}{x}\int_0^T |f(t)|\,\mathrm{d}t$ donc, par encadrement,

$$\frac{1}{x}\int_{\lfloor\frac{x}{T}\rfloor T}^{x} f(t)\,\mathrm{d}t \underset{x\to+\infty}{\longrightarrow} 0.$$

En conclusion,

$$f^*(x) \underset{x\to+\infty}{\longrightarrow} \frac{1}{T}\int_0^T f(t)\,\mathrm{d}t\,.$$

Partie C – Inégalité de Hardy

Remarquons tout d'abord que

$$\int_0^x (4f^2(t)-(f^*)^2(t))\,\mathrm{d}t = \int_0^x (2f(t)-f^*(t))^2\,\mathrm{d}t + 2\int_0^x 2f(t)f^*(t)-(f^*)^2(t)\,\mathrm{d}t$$

Remarquons ensuite que la fonction $x \mapsto x(f^*)^2(x)$ est dérivable sur $\mathbb{R}_+^*$ d'après les règles usuelles et en 0 car prolongeable par continuité en 0 avec un taux d'accroissement en 0 qui converge vers $f^2(0)$. Sa dérivée est $x \mapsto 2f(x)f^*(x)-(f^*)^2(x)$. Ainsi, on peut calculer la seconde intégrale et obtenir

$$\begin{aligned}\int_0^x (4f^2(t)-(f^*)^2(t))\,\mathrm{d}t &= \int_0^x (2f(t)-f^*(t))^2\,\mathrm{d}t + 2[t(f^*)^2(t)]_0^x \\ &= \int_0^x (2f(t)-f^*(t))^2\,\mathrm{d}t + 2x(f^*)^2(x) \geq 0.\end{aligned}$$

Chapitre 22

Séries numériques

1. Séries générales — **2.** Séries à termes positifs — **3.** Applications

Objectifs et compétences du programme

- Adapter les critères de convergence des suites réelles aux séries à terme général positif.
- Déterminer la convergence absolue d'une série.
- Utiliser la comparaison série-intégrale pour obtenir des équivalents.
- Reconnaître un produit de Cauchy.

Niels Henrik Abel, mathématicien norvégien (1802-1829)
Ses contributions majeures concernent la non-résolubilité de l'équation de degré 5 (il n'existe pas de formule générale donnant les racines d'un polynôme de degré 5) d'où émergera le concept de groupe abélien, l'étude d'une famille d'intégrales également appelées en son honneur abéliennes et enfin, l'étude de la convergence des séries. Plus précisément, il a établi des critères de convergence assez fins pour les séries qui ne sont pas absolument convergentes en passant par un analogue discret de la formule d'intégration par parties.

COURS

Dans un chapitre précédent, nous avons étudié la convergence des suites à valeurs réelles ou complexes. Dans ce chapitre, on va compléter cette première approche en développant des outils spécifiques pour l'étude des suites particulières, appelées séries, obtenues comme somme des termes d'une suite auxiliaire, appelée terme général.

À première vue, on semble réviser notre arsenal en utilisant tour à tour le théorème de convergence monotone, les lemmes d'encadrements, les relations de comparaison asymptotique. En y regardant de plus près, on développe d'autres points de vue, notamment la riche analogie avec le calcul intégral.

1. Séries générales

1.1. Définitions

Définition 22.1.

Soit $(u_n)_n \in \mathbb{C}^{\mathbb{N}}$.
▹ La série de terme général u_n est la suite $\left(S_n = \sum_{k=0}^{n} u_k\right)_n$.
▹ La $n^{\text{ième}}$ somme partielle de la série de terme général u_n est le terme S_n.
▹ La somme de la série de terme général u_n est la limite de la suite $\left(S_n = \sum_{k=0}^{n} u_k\right)_n$, quand elle existe. On la note alternativement

$$\sum_{n=0}^{\infty} u_n \text{ ou } \sum u_n.$$

Si la somme est bien définie, le $n^{\text{ième}}$ reste partiel est la différence entre la somme et la $n^{\text{ième}}$ somme partielle.

Remarque

On fera attention que la somme d'une série est définie comme une limite donc que les manipulations usuelles avec les sommes finies ne se généralisent pas sans argument supplémentaire.

On a besoin de préciser quelques termes pour la convergence d'une série.

Définition 22.2.

Soit $(u_n)_n \in \mathbb{C}^{\mathbb{N}}$.
▹ La série de terme général u_n est convergente si la suite $\left(S_n = \sum_{k=0}^{n} u_k\right)_n$ converge.
▹ La série de terme général u_n est absolument convergente si la série de terme général $|u_n|$ est convergente.
▹ La série de terme général u_n est semi-convergente si elle est convergente sans être absolument convergente.

Un premier résultat liant suite et série est le suivant.

Proposition 22.3.

Soit $(u_n)_n \in \mathbb{C}^{\mathbb{N}}$. La suite $(u_n)_n$ et la série de terme général $u_n - u_{n-1}$ sont de même nature.

Démonstration
Par télescopage,

$$\sum_{k=1}^{n}(u_k - u_{k-1}) = u_n - u_0,$$

donc le membre de droite définit une suite convergente si, et seulement si, le membre de gauche aussi. ■

Proposition 22.4.

Soit $(u_n)_n$, $(v_n)_n \in \mathbb{C}^{\mathbb{N}}$, λ, $\mu \in \mathbb{C}$ tels que les séries de termes généraux u_n et v_n convergent. Alors, la série de terme général $\lambda u_n + \mu v_n$ converge.

Démonstration
Il suffit de remarquer que, pour tout $n \in \mathbb{N}$,

$$\sum_{k=0}^{n}(\lambda u_k + \mu v_k) = \lambda \sum_{k=0}^{n} u_k + \mu \sum_{k=0}^{n} v_k.$$

Comme les termes de la combinaison linéaire à droite convergent, la somme dans le terme de gauche converge également. ■

1.2. Premiers exemples

Exemple Séries télescopiques

▹ La série de terme général $\frac{1}{n(n+1)}$ est convergente. En effet, la décomposition $\frac{1}{n(n+1)} = \frac{1}{n} - \frac{1}{n+1}$ entraîne par télescopage que la $n^{\text{ième}}$ somme partielle est

$$\sum_{k=1}^{n} \frac{1}{k(k+1)} = 1 - \frac{1}{n+1},$$

et que cette suite converge vers 1.
▹ Le télescopage peut être parfois plus compliqué à détecter. Considérons par exemple, la série de terme général $\frac{1}{(n+1)\sqrt{n}+n\sqrt{n+1}}$. Elle est télescopique puisque, en utilisant la quantité conjuguée,

$$\begin{aligned}
\frac{1}{(n+1)\sqrt{n}+n\sqrt{n+1}} &= \frac{1}{\sqrt{n(n+1)}(\sqrt{n+1}+\sqrt{n})} \\
&= \frac{\sqrt{n+1}-\sqrt{n}}{\sqrt{n(n+1)}(\sqrt{n+1}^2-\sqrt{n}^2)} \\
&= \frac{\sqrt{n+1}-\sqrt{n}}{\sqrt{n(n+1)}} \\
&= \frac{1}{\sqrt{n}} - \frac{1}{\sqrt{n+1}}.
\end{aligned}$$

Par conséquent, la $n^{\text{ième}}$ somme partielle est

$$\sum_{k=1}^{n} \frac{1}{(k+1)\sqrt{k}+k\sqrt{k+1}} = 1 - \frac{1}{\sqrt{n+1}}.$$

La série converge à nouveau vers 1.

Conseils méthodologiques

Pour calculer une somme partielle par télescopage, il faut pouvoir mettre le terme général sous la forme $u_n = a_n - a_{n+1}$ avec $(a_n)_n$ une suite numérique.

Exemple Série géométrique

Considérons la série dont le terme général est q^n avec $q \neq 1$ (que l'on appelle série géométrique). La $n^{\text{ième}}$ somme partielle est alors

$$\sum_{k=0}^{n} q^k = \frac{1-q^{n+1}}{1-q}.$$

Donc la série converge si, et seulement si, $|q| < 1$ et sa somme est alors $\frac{1}{1-q}$; on vérifie que si $q = 1$, la série est divergente.

Exemple Série géométrique dérivée

Repartons, comme dans l'exemple précédent, de la formule pour $x \neq 1$, $\sum_{k=0}^{n} x^k = \frac{1-x^{n+1}}{1-x}$. Dérivons cette relation par rapport à la variable x ; on obtient l'identité suivante sur $\mathbb{R} \setminus \{1\}$ entre fonctions polynomiales

$$(1-x)^2 \sum_{k=1}^{n} k x^{k-1} = -(n+1)x^n(1-x) + (1-x^{n+1}).$$

Par conséquent, les deux polynômes sont égaux sur tout $\mathbb{C}$. En appliquant cette égalité en q un complexe de module strictement inférieur à 1 fixé, on obtient

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

Par conséquent, la série de terme général kq^{k-1} avec $|q| < 1$ converge (en utilisant les croissances comparées) et sa somme est $\frac{1}{(1-q)^2}$.

Exemple Série exponentielle

Soit $a \in \mathbb{R}$. La série de terme général $\frac{a^n}{n!}$ est convergente de somme e^a. En effet, l'inégalité de Taylor-Lagrange entraîne que :

$$\left| e^a - \sum_{k=0}^{n} \frac{a^k}{k!} \right| \leq \max(1, e^a) \frac{a^{n+1}}{(n+1)!}.$$

Comme le second membre tend vers 0 lorsque n tend vers $+\infty$ (croissances comparées), alors les sommes partielles considérées convergent vers e^a. Un raisonnement analogue avec la formule de Taylor-Lagrange pour les fonctions cos et sin permet de montrer que

▹ la série de terme général $\frac{(-1)^k a^{2k}}{(2k)!}$ converge et que sa somme vaut $\cos(a)$;

▹ la série de terme général $\frac{(-1)^k a^{2k+1}}{(2k+1)!}$ converge et que sa somme vaut $\sin(a)$.

Exemple Série harmonique

Montrons que la série de terme général $\frac{1}{n}$ diverge. En effet, pour tout $n \geq 1$, comme $t \mapsto \frac{1}{t}$ est

décroissante sur $\mathbb{R}_+^*$,

$$\frac{1}{n} \geq \int_n^{n+1} \frac{\mathrm{d}t}{t},$$

et, par la relation de Chasles,

$$\sum_{k=1}^{n} \frac{1}{k} \geq \sum_{k=1}^{n} \int_k^{k+1} \frac{\mathrm{d}t}{t} = \int_1^{n+1} \frac{\mathrm{d}t}{t} = \ln(n+1).$$

Or, la suite $(\ln(n))_n$ diverge vers $+\infty$ donc, par comparaison, la suite des sommes partielles de la série harmonique diverge aussi.

Remarque

Pour montrer la divergence de la série harmonique, nous avons choisi de comparer la somme partielle à une intégrale. Cette méthode se généralise comme nous le verrons dans la partie suivante. On doit toutefois indiquer une méthode plus rapide dans ce cas précis basée sur la minoration suivante pour tout $n \in \mathbb{N}$,

$$\sum_{k=n+1}^{2n} \frac{1}{k} \geq \sum_{k=n+1}^{2n} \frac{1}{2n} = n\frac{1}{2n} = \frac{1}{2}.$$

Par conséquent, en séparant la somme par paquets,

$$\sum_{k=1}^{2n} \frac{1}{k} = 1 + \sum_{l=1}^{n} \sum_{k=l+1}^{2l} \frac{1}{k} \geq 1 + \sum_{l=1}^{n} \frac{1}{2} = 1 + \frac{n}{2}.$$

Par comparaison, $\sum_{k=1}^{2n} \frac{1}{k} \to +\infty$.

1.3. Propriétés élémentaires

Proposition 22.5.

Soit $(u_n)_n \in \mathbb{C}^{\mathbb{N}}$. Si la série de terme général u_n converge, alors la suite $(u_n)_n$ tend vers 0.

La réciproque est fausse comme vu avec la série harmonique qui diverge alors que son terme général tend vers 0.

Conseils méthodologiques

Cette proposition permet de montrer rapidement la divergence des séries dont le terme général ne tend pas vers 0.
Par exemple, la série de terme général $\cos(n)$ diverge grossièrement.

Démonstration

Notons S_n la $n^{\text{ième}}$ somme partielle de la série de terme général u_n. Écrivons, pour tout $n \in \mathbb{N} \setminus \{0\}$, $u_n = S_n - S_{n-1}$. Par hypothèse, la suite $(S_n)_n$ (et donc la suite extraite $(S_{n-1})_n$) converge vers $\ell \in \mathbb{C}$. Par conséquent, la suite $(u_n)_n$ converge vers $\ell - \ell = 0$. ■

Proposition 22.6.

Soit $(u_n)_n \in \mathbb{C}^{\mathbb{N}}$. Si la série de terme général u_n converge absolument, alors elle converge.

Démonstration

▹ Démontrons tout d'abord ce résultat pour une suite réelle $(u_n)_n$.

Supposons que la série de terme général $|u_n|$ converge et considérons, pour tout n,

$$S_n^+ = \sum_{k=0}^{n} (u_k)^+, \qquad S_n^- = \sum_{k=0}^{n} (u_k)^-.$$

La suite $(S_n^+)_n$ est croissante et majorée par $\sum_{k=0}^{+\infty} |u_k|$ donc converge. Comme, pour tout $n \in \mathbb{N}$,

$$S_n^- = -S_n^+ + \sum_{k=0}^{n} |u_k|,$$

la suite $(S_n^-)_n$ est convergente (comme différence de suites convergentes). Ainsi, la série de terme général u_n est convergente d'après la relation suivante donnant la somme partielle

$$\sum_{k=0}^{n} u_k = S_n^+ - S_n^-.$$

▹ Revenons au cas général d'une suite $(u_n)_n \in \mathbb{C}^{\mathbb{N}}$ telle que la série de terme général $|u_n|$ converge. Les suites des sommes partielles des séries de termes généraux $|\mathfrak{Re}(u_n)|$ et $|\mathfrak{Im}(u_n)|$ sont majorées par $\sum_{k=0}^{+\infty} |u_k|$ (car les valeurs absolues des parties réelle et imaginaire sont majorées par le module) donc convergent. D'après le point précédent, les séries (réelles) de termes généraux $\mathfrak{Re}(u_n)$ et $\mathfrak{Im}(u_n)$ convergent, ainsi, la série de terme général $u_n = \mathfrak{Re}(u_n) + i\mathfrak{Im}(u_n)$ converge également. ■

2. Séries à termes positifs

Dans cette section, on ne s'intéresse qu'aux séries dont le terme général est positif (ou, en adaptant légèrement, négatif ou même de signe constant à partir d'un certain rang). Ce cas est extrêmement important puisqu'il correspond à l'étude de la convergence absolue d'une série.

Commençons par un résultat élémentaire mais essentiel.

Proposition 22.7.

Soit $(u_n)_n \in (\mathbb{R}_+)^{\mathbb{N}}$. La série de terme général u_n converge si, et seulement si, la suite des sommes partielles est majorée.

Démonstration

Il suffit d'appliquer le critère de convergence des suites monotones à la suite croissante (car $u_n \geq 0$) des sommes partielles. ■

2.1. Inégalités

Proposition 22.8.

Soit $(u_n)_n \in (\mathbb{R}_+)^{\mathbb{N}}$ et $(v_n)_n \in (\mathbb{R}_+)^{\mathbb{N}}$ telles que $u_n \leq v_n$ (au moins à partir d'un certain rang).
▹ Si la série de terme général v_n converge, alors la série de terme général u_n converge.
▹ Si la série de terme général u_n diverge, alors la série de terme général v_n diverge.

Démonstration
Écrivons les arguments dans le cas où l'inégalité est vérifiée pour tout indice n.
▹ Si la série de terme général v_n converge, alors les sommes partielles de la série de terme général u_n sont majorées par la somme de la série de terme général v_n donc la série de terme général u_n converge.
▹ Si la série de terme général u_n diverge, alors ses sommes partielles divergent vers $+\infty$ et par comparaison celles de la série de terme général v_n aussi : la série de terme général v_n diverge. ■

Exemple

La série de terme général $\frac{1}{2^n(2+\cos n)}$ est convergente car $0 \leq \frac{1}{2^n(2+\cos n)} \leq \frac{1}{2^n}$ et $\frac{1}{2^n}$ est le terme général d'une série convergente.

Exemple

Soit $x \in \mathbb{R}_+$. On a établi l'existence d'une unique suite d'entiers naturels $(x_n)_n$ telle que

$$\forall n \in \mathbb{N}, \qquad \sum_{k=0}^{n} x_k 10^{-k} \leq x < \sum_{k=0}^{n} x_k 10^{-k} + 10^{-n}.$$

De plus, la suite $(x_n)_n$ est à valeurs dans $[\![0,9]\!]$ pour $n \geq 1$ et n'est pas constante égale à 9 à partir d'un certain rang.
Alors, pour tout $n \geq 1$, $0 \leq x_n 10^{-n} \leq 9.10^{-n}$. Par comparaison, la série de terme général $x_n 10^{-n}$ est convergente. En passant à la limite dans l'encadrement précédent, on obtient que

$$x = \sum_{k=0}^{+\infty} \frac{x_k}{10^k}.$$

Voici un résultat de comparaison aux séries géométriques.

Exemple Critère de D'Alembert

Soit $(u_n)_n \in (\mathbb{R}_+)^{\mathbb{N}}$ une suite telle que $\frac{u_{n+1}}{u_n}$ admette une limite $\ell \in \overline{\mathbb{R}}$.
- Si $\ell < 1$, alors la série de terme général u_n est convergente.
- Si $\ell > 1$, alors la série de terme général u_n est divergente.

Démontrons ce résultat.
▹ Si $\ell < 1$, alors il existe N tel que : $\forall n \geq N, \qquad u_{n+1} \leq u_n \frac{1+\ell}{2}$.
Par une récurrence immédiate, pour tout $n \geq N$, on a $u_n \leq u_N \left(\frac{1+\ell}{2}\right)^{n-N}$.
Or, la série géométrique de raison $\frac{1+\ell}{2} \in]0,1[$ est convergente. D'où le résultat par comparaison.
▹ Si $\ell > 1$, alors on procède de même avec N tel que $\forall n \geq N, \qquad u_{n+1} \geq u_n \frac{1+\ell}{2}$.

Remarque
On ne peut pas conclure avec le critère de D'Alembert si $\ell = 1$ comme on peut le voir en considérant les séries de terme général $\frac{1}{n}$ et $\frac{1}{n^2}$ qui sont de nature différente.

Proposition 22.9.

Soit $(u_n)_n$, $(v_n)_n \in \mathbb{C}^{\mathbb{N}}$, λ, $\mu \in \mathbb{C}$ tels que les séries de termes généraux u_n et v_n convergent absolument.
Alors, la série de terme général $\lambda u_n + \mu v_n$ converge absolument.

Démonstration
Il suffit de remarquer que, pour tout $n \in \mathbb{N}$, $|\lambda u_n + \mu v_n| \leq |\lambda||u_n| + |\mu||v_n|$. La série de terme général $|\lambda||u_n|+|\mu||v_n|$ converge donc par inégalité, la série de terme général $|\lambda u_n+\mu v_n|$ converge également. ■

2.2. Comparaison série-intégrale

Commençons par quelques exemples importants.

Proposition 22.10. Séries de Riemann

Soit $\alpha \in \mathbb{R}$. La série de terme général $\frac{1}{n^\alpha}$ converge si, et seulement si, $\alpha > 1$.

Démonstration
Adaptons l'exemple de la série harmonique.
▹ Si $\alpha < 1$, alors on utilise que la fonction $t \longmapsto \frac{1}{t^\alpha}$ est décroissante sur $\mathbb{R}_+^*$ pour obtenir que la minoration

$$\frac{1}{k^\alpha} = \int_k^{k+1} \frac{\mathrm{d}t}{k^\alpha} \geq \int_k^{k+1} \frac{\mathrm{d}t}{t^\alpha}$$

qui donne en additionnant avec la relation de Chasles,

$$\sum_{k=1}^{n} \frac{1}{k^\alpha} \geq \int_1^{n+1} \frac{\mathrm{d}t}{t^\alpha} = \frac{1}{1-\alpha}((n+1)^{1-\alpha}-1),$$

d'où la divergence de la série.
▹ Si $\alpha > 1$, alors on utilise cette fois-ci une majoration analogue, pour tout $n \geq 1$,

$$\int_n^{n+1} \frac{\mathrm{d}t}{t^\alpha} \geq \frac{1}{(n+1)^\alpha},$$

et, toujours avec la relation de Chasles, on obtient pour $n \geq 1$,

$$1 + \frac{1}{1-\alpha}((n+1)^{1-\alpha}-1) \geq \sum_{k=1}^{n} \frac{1}{k^\alpha}.$$

Le majorant admet cette fois-ci une limite finie ; ainsi, la suite des sommes partielles est croissante (car le terme général est positif), majorée donc converge. ■

Généralisons la méthode de cette démonstration.

Proposition 22.11. Comparaison série-intégrale

Soit $f : \mathbb{R}_+ \to \mathbb{R}_+$ décroissante et $a \in \mathbb{R}_+$. Alors, pour tout $n \geq 1$,

$$\int_0^n f(t)\,\mathrm{d}t \geq \sum_{k=1}^{n} f(k) \geq \int_1^{n+1} f(t)\,\mathrm{d}t\,.$$

En particulier, la série de terme général $f(n)$ et la suite $\left(\int_a^n f(t)\,\mathrm{d}t\right)_n$ sont de même nature.

Il y a évidemment un résultat équivalent pour f croissante.

Démonstration

Pour tout $k \in [\![1, n]\!]$ et tout $t \in [k, k+1]$, $f(k) \geq f(t) \geq f(k+1)$. Ainsi, en intégrant, pour tout $k \in [\![1, n]\!]$,

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

En renversant ces inégalités, pour tout $k \in [\![1, n]\!]$,

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

On obtient l'inégalité recherchée avec la relation de Chasles

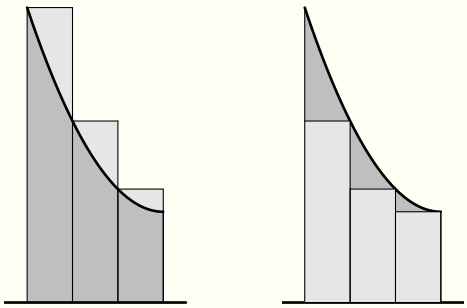
$$\int_0^n f(t)\,\mathrm{d}t = \sum_{k=1}^{n} \int_{k-1}^{k} f(t)\,\mathrm{d}t \geq \sum_{k=1}^{n} f(k) \geq \sum_{k=1}^{n} \int_k^{k+1} f(t)\,\mathrm{d}t = \int_1^{n+1} f(t)\,\mathrm{d}t\,.$$

▹ Si la série de terme général $f(n)$ est convergente, alors la suite $\left(\int_a^n f(t)\,\mathrm{d}t\right)_n$ est croissante majorée (avec l'inégalité de droite) donc convergente.

▹ Si la suite $\left(\int_a^n f(t)\,\mathrm{d}t\right)_n$ converge, alors la suite des sommes partielles de la série de terme général $f(n)$ est majorée donc la série converge. ■

Remarque

L'inégalité établie correspond à l'encadrement de l'aire sous la courbe à l'aide de rectangles construits à droite ou à gauche de la valeur aux arguments entiers.



Conseils méthodologiques

La comparaison série-intégrale est un outil puissant que l'on peut mobiliser pour les séries dont le terme général est de la forme $u_n = f(n)$ avec f monotone sur un voisinage de $+\infty$.

Exemple Séries de Bertrand

Soit $\beta \geq 0$. La série de terme général $\frac{1}{n(\ln n)^\beta}$ est convergente si, et seulement si, $\beta > 1$.
En effet, la fonction $f : x \mapsto \frac{1}{x(\ln x)^\beta}$ admet la primitive $x \mapsto \frac{1}{1-\beta}(\ln x)^{1-\beta}$ pour $\beta \neq 1$ et $x \mapsto \ln(\ln(x))$ sinon ; il suffit alors de vérifier que cette primitive admet une limite finie en $+\infty$ si, et seulement si, $\beta > 1$ et conclure avec la proposition précédente.

Remarque
La proposition précédente permet de déterminer la nature d'une série. On peut toutefois dire un peu plus en exploitant simultanément les deux bornes de l'encadrement.
Par exemple, la monotonie de la fonction $x \mapsto \frac{1}{x}$ entraîne l'encadrement suivant

$$\forall n \geq 2, \qquad 1 + \ln n = 1 + \int_1^n \frac{\mathrm{d}t}{t} \geq \sum_{k=1}^n \frac{1}{k} \geq \int_1^{n+1} \frac{\mathrm{d}t}{t} = \ln(n+1).$$

Comme les deux termes extrémaux sont équivalents à $\ln n$, on obtient l'équivalent suivant des sommes partielles de la série harmonique

$$\sum_{k=1}^n \frac{1}{k} \sim \ln n.$$

2.3. Relations de comparaison

Proposition 22.12.

Soit $(u_n)_n \in (\mathbb{R}_+)^\mathbb{N}$ et $(v_n)_n \in (\mathbb{R}_+)^\mathbb{N}$ telles que $u_n \sim v_n$. La série de terme général u_n converge si, et seulement si, la série de terme général v_n converge.

Démonstration
D'après la définition de l'équivalence entre suites, il existe un rang N tel que

$$\forall n \geq N, \qquad \frac{1}{2} v_n \leq u_n \leq \frac{3}{2} v_n,$$

et on conclut avec la section précédente. ■

Exemple

La série de terme général $\frac{1}{n^2}$ est convergente car son terme général est positif et équivalent à $\frac{1}{n(n+1)}$ qui est le terme général d'une série convergente.

On veillera à vérifier l'importante hypothèse de signe constant dans cette proposition. Toutefois, il arrive que l'on puisse utiliser ce résultat pour étudier des suites qui ne sont pas de signe constant. Voici un exemple classique.

Exemple Série harmonique alternée

La série de terme général $\frac{(-1)^n}{n}$ est convergente (mais pas absolument convergente car la série harmonique diverge).

En effet, si on note S_n sa $n^{\text{ième}}$ partielle,

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

Or, la série de terme général $\frac{1}{2n(2n-1)}$ est convergente (car ce terme est positif et équivalent à $\frac{1}{4n^2}$) donc la suite $(S_{2n})_n$ converge. Comme, par ailleurs,

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

la suite extraite $(S_{2n+1})_n$ converge vers la même limite. Par conséquent, $(S_n)_n$ converge, ce qui est la définition de la convergence de la série de terme général $\frac{(-1)^n}{n}$.

Proposition 22.13.

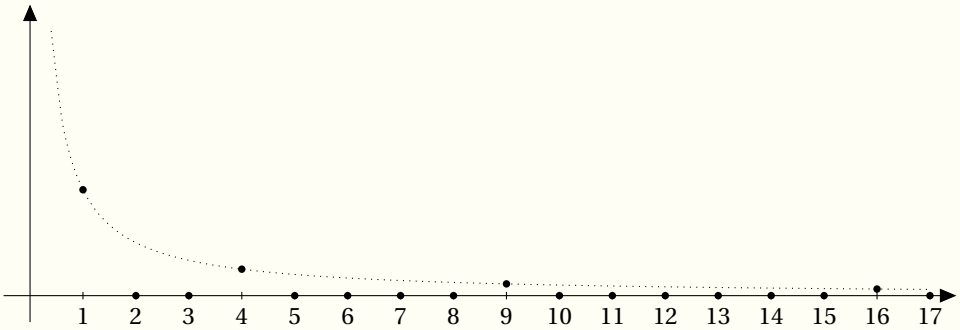
Soit $(u_n)_n \in \mathbb{C}^{\mathbb{N}}$ et $(v_n)_n \in (\mathbb{R}_+)^{\mathbb{N}}$ telles que $u_n = O(v_n)$. Si la série de terme général v_n converge, alors la série de terme général u_n converge absolument.

Démonstration

Il suffit d'écrire qu'il existe $C > 0$ tel qu'à partir d'un certain rang $|u_n| \leq C\,v_n$. Ensuite, on applique le critère de comparaison par inégalité pour obtenir que la série de terme général $|u_n|$ est convergente. ■

Remarque

Il n'y a pas de réciproque simple aux critères de comparaison aux séries de Riemann. Par exemple, on peut facilement construire une suite $(u_n)_n$ positive telle que la série de terme général u_n converge et pourtant $u_n \neq o(\frac{1}{n})$.



Il suffit pour cela de prendre $u_n = 0$ si n n'est pas un carré et $u_n = \frac{1}{n}$ sinon.

▹ La suite $(u_n)_n$ n'est pas négligeable devant $(\frac{1}{n})_n$ à cause des termes d'indices carrés.

▹ Les sommes partielles de la série de terme général u_n sont des sommes partielles de la série des inverses des carrés donc définissent une suite convergente.

3. Applications

3.1. Études asymptotiques de suites

Montrons à travers quelques exemples l'utilisation de la sommation des équivalents pour obtenir des développements asymptotiques de suites.

Exemple

Notons $H_n = \sum_{k=1}^{n} \frac{1}{k}$. Nous avons déjà obtenu par comparaison à une intégrale que cette somme est équivalente à $\ln n$. Pour obtenir le terme suivant du développement, on pose $u_n = H_n - \ln n$. Alors, pour tout $n \in \mathbb{N} \setminus \{0\}$,

$$\begin{aligned} u_{n+1} - u_n &= H_{n+1} - H_n - (\ln(n+1) - \ln(n)) \\ &= \frac{1}{n+1} - \ln\left(1 + \frac{1}{n}\right) = -\frac{1}{2n^2} + o\left(\frac{1}{n^2}\right). \end{aligned}$$

Ainsi, la série de terme général $u_{n+1} - u_n$ est convergente, ce qui par télescopage signifie que (u_n) converge vers un réel noté γ. On a ainsi établi le développement asymptotique à deux termes

$$H_n = \ln(n) + \gamma + o(1).$$

Détaillons désormais la preuve de la formule de Stirling en utilisant à la fois les séries numériques et les intégrales de Wallis.

Proposition 22.14. Formule de Stirling

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

Lemme 22.15.

La suite $\left(\frac{\sqrt{n}.n^n}{n!e^n}\right)_n$ converge vers une limite notée ℓ.

Démonstration

Pour montrer la convergence de cette suite, il suffit d'étudier (grâce à la continuité de l'exponentielle sur $\mathbb{R}$) la suite $\left(u_n = \ln \frac{\sqrt{n}.n^n}{n!e^n}\right)_n$. Or, pour tout $n \in \mathbb{N}$,

$$\begin{aligned} u_{n+1} - u_n &= \left(n + \frac{1}{2}\right) \ln\left(1 + \frac{1}{n}\right) - 1 \\ &= \left(n + \frac{1}{2}\right)\left(\frac{1}{n} - \frac{1}{2n^2} + \frac{1}{3n^3} + o\left(\frac{1}{n^3}\right)\right) - 1 \\ &= \frac{1}{12n^2} + o\left(\frac{1}{n^2}\right). \end{aligned}$$

Ainsi, la série de terme général $u_{n+1} - u_n$ converge ; comme la $(n-1)^{\text{ième}}$ somme partielle de cette série est $u_n - 1$, on en déduit la convergence souhaitée. ■

Lemme 22.16.

Notons, pour tout $n \in \mathbb{N}$, la $n^{\text{ième}}$ intégrale de Wallis

$$I_n = \int_0^{\frac{\pi}{2}} \cos^n t \, \mathrm{d}t \, .$$

La suite $(I_n)_n$ vérifie les propriétés suivantes :
▹ pour tout $n \in \mathbb{N} \setminus \{0\}$, $I_{n+1} = \frac{n}{n+1} I_{n-1}$;
▹ pour tout $p \in \mathbb{N}$, $I_{2p} = \frac{(2p)!}{(2^p p!)^2} \cdot \frac{\pi}{2}$;
▹ pour tout $n \in \mathbb{N} \setminus \{0\}$, $n I_n I_{n-1} = \frac{\pi}{2}$;
▹ $I_n \sim I_{n-1}$.

Démonstration
▹ Les deux premiers points ont d'ores et déjà été établis à l'exemple 3.3. du chapitre précédent.
▹ En utilisant la relation de récurrence, on obtient que, pour tout $n \in \mathbb{N} \setminus \{0\}$,

$$n I_n I_{n-1} = n I_n \frac{n+1}{n} I_{n+1} = (n+1) I_{n+1} I_n .$$

Ainsi, pour tout $n \in \mathbb{N} \setminus \{0\}$, $n I_n I_{n-1} = 1.I_1 I_0 = \frac{\pi}{2}$.
▹ Remarquons tout d'abord que, pour tout $n \in \mathbb{N} \setminus \{0\}$,

$$\forall t \in \left[0, \tfrac{\pi}{2}\right], \qquad \cos^{n+1} t \leq \cos^n t \leq \cos^{n-1} t \, ,$$

donc $I_{n+1} \leq I_n \leq I_{n-1}$ et, avec la relation de récurrence,

$$\frac{n}{n+1} \leq \frac{I_n}{I_{n-1}} \leq 1.$$

Par encadrement, $\frac{I_n}{I_{n-1}} \to 1$, c'est-à-dire $I_n \sim I_{n-1}$. ■

Revenons à la preuve de la proposition.

Démonstration
En combinant les deux derniers résultats du lemme précédent, $[2p I_{2p}^2 \sim \frac{\pi}{2}$. Or, avec l'expression explicite de I_{2p} et le résultat du premier lemme, $2p I_{2p}^2 \to \frac{\pi^2}{\ell^2}$. En comparant ces deux résultats, on obtient $\ell^2 = 2\pi$ et donc la formule de Stirling en reportant dans le premier lemme. ■

3.2. Produit de Cauchy

La convergence passe bien à la somme (par linéarité) ; le résultat suivant indique le résultat correspondant au produit de séries convergentes.

Proposition 22.17.

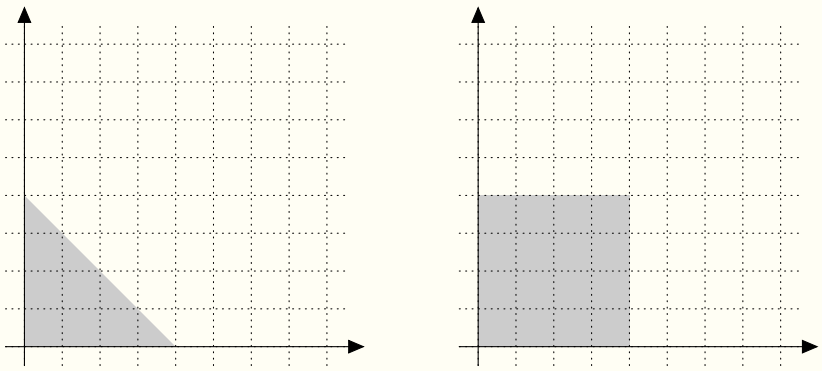
Soit $(a_n)_n$ et $(b_n)_n \in (\mathbb{R}_+)^{\mathbb{N}}$ telles que les séries de termes généraux a_n et b_n sont convergentes. Posons, pour tout n,

$$c_n = \sum_{k=0}^{n} a_k b_{n-k}.$$

Alors, la série de terme général c_n est convergente et sa somme est $\left(\sum_{n=0}^{+\infty} a_n\right)\left(\sum_{n=0}^{+\infty} b_n\right)$.

Démonstration

Notons A_n, B_n et C_n les sommes partielles des séries de termes généraux a_n, b_n et c_n. Les quantités C_n et $A_n B_n$ sont les sommes des termes $a_k b_l$ pour les couples (k,l) dans les ensembles d'équations $k+l \leq n$ et $\max(k,l) \leq n$ représentés ci-dessous :



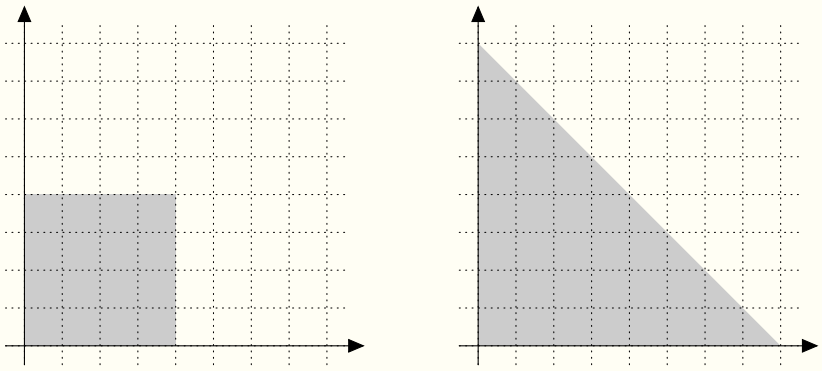
Comme il y a une inclusion entre ces ensembles et que les termes $a_k b_l$ sont positifs, on obtient

$$\forall n, \qquad C_n \leq A_n B_n.$$

Les suites $(A_n)_n$ et $(B_n)_n$ sont convergentes donc majorées ; par conséquent, la suite croissante $(C_n)_n$ est majorée donc convergente. De plus,

$$\sum_{n=0}^{+\infty} c_n \leq \Big(\sum_{n=0}^{+\infty} a_n\Big)\Big(\sum_{n=0}^{+\infty} b_n\Big).$$

Partant cette fois-ci du dessin suivant.



On remarque que, pour tout n, $A_n B_n \leq C_{2n}$; en passant à la limite, on obtient

$$\Big(\sum_{n=0}^{+\infty} a_n\Big)\Big(\sum_{n=0}^{+\infty} b_n\Big) \leq \sum_{n=0}^{+\infty} c_n$$

et donc

$$\sum_{n=0}^{+\infty} c_n = \Big(\sum_{n=0}^{+\infty} a_n\Big)\Big(\sum_{n=0}^{+\infty} b_n\Big).$$

■

Proposition 22.18.

Soit $(a_n)_n$ et $(b_n)_n \in \mathbb{C}^{\mathbb{N}}$ telles que les séries de termes généraux a_n et b_n sont absolument convergentes. Posons, pour tout n,

$$c_n = \sum_{k=0}^{n} a_k b_{n-k}.$$

Alors, la série de terme général c_n est absolument convergente et sa somme est $\left(\sum_{n=0}^{+\infty} a_n\right)\left(\sum_{n=0}^{+\infty} b_n\right)$.

Démonstration

Notons A_n, A'_n, B_n, B'_n, C_n et C'_n les sommes partielles des séries de termes généraux a_n, $|a_n|$, b_n, $|b_n|$, c_n et $|c_n|$. La série de terme général c_n est absolument convergente car

$$\forall n, \qquad |c_n| \leq \sum_{k=0}^{n} |a_k||b_{n-k}|$$

et que le majorant est le terme général d'une série convergente d'après la proposition précédente. Par ailleurs, pour tout n,

$$\begin{aligned} |A_n B_n - C_n| &= \left| \sum_{\substack{k,l \leq n \\ k+l > n}} a_k b_l \right| \\ &\leq \sum_{\substack{k,l \leq n \\ k+l > n}} |a_k||b_l| \\ &= A'_n B'_n - C'_n. \end{aligned}$$

Comme d'après la proposition précédente, $A'_n B'_n - C'_n \to 0$, on en déduit le résultat annoncé. ■

Exemple

Considérons une série absolument convergente de terme général u_n. Alors, le produit de Cauchy des suites $(2^{-n})_n$ et $(u_n)_n$ est

$$v_n = 2^{-n} \sum_{k=0}^{n} 2^k u_k.$$

D'après la proposition précédente, la série de terme général v_n est absolument convergente de somme

$$\sum_{n=0}^{+\infty} v_n = \left(\sum_{n=0}^{+\infty} 2^{-n}\right)\left(\sum_{n=0}^{+\infty} u_n\right) = 2\sum_{n=0}^{+\infty} u_n.$$

FICHE

SYNTHÈSE

Convergence

Soit $(u_n)_n \in \mathbb{C}^{\mathbb{N}}$.

- ▹ La série de terme général u_n est convergente si la suite $\left(S_n = \sum_{k=0}^{n} u_k\right)_n$ converge.
- ▹ La série de terme général u_n est absolument convergente si la série de terme général $|u_n|$ est convergente. Une série absolument convergente est convergente.
- ▹ La série de terme général u_n est semi-convergente si elle est convergente sans être absolument convergente.
- ▹ La série de terme général u_n est grossièrement divergente si la suite $(u_n)_n$ ne converge pas vers 0.

Exemples

▹ La série géométrique de terme général q^n converge pour $|q| < 1$. La somme est alors $\frac{1}{1-q}$.

▹ La série de Riemann de terme général $\frac{1}{n^\alpha}$ converge si, et seulement si, $\alpha > 1$.

En particulier, la série harmonique de terme général $\frac{1}{n}$ diverge ; plus précisément

$$\sum_{k=1}^{n} \frac{1}{k} \sim \ln n.$$

▹ La série harmonique alternée de terme général $\frac{(-1)^n}{n}$ est convergente.

À partir de ces exemples, on peut étudier d'autres séries par les résultats suivants de comparaison (inégalité, équivalence, négligeabilité...).

Comparaison

▹ Soit $(u_n)_n \in (\mathbb{R}_+)^{\mathbb{N}}$ et $(v_n)_n \in (\mathbb{R}_+)^{\mathbb{N}}$ telles que $u_n \leq v_n$ (au moins à partir d'un certain rang).

- Si la série de terme général v_n converge, alors la série de terme général u_n converge.
- Si la série de terme général u_n diverge, alors la série de terme général v_n diverge.

▹ Soit $(u_n)_n \in \mathbb{C}^{\mathbb{N}}$ et $(v_n)_n \in (\mathbb{R}_+)^{\mathbb{N}}$ telles que $u_n = O(v_n)$. Si la série de terme général v_n converge, alors la série de terme général u_n converge absolument.

▹ Soit $f : \mathbb{R}_+ \to \mathbb{R}_+$ décroissante et $a \in \mathbb{R}_+$. Alors, pour tout $n \geq 1$,

$$\int_0^n f(t)\,\mathrm{d}t \geq \sum_{k=1}^{n} f(k) \geq \int_1^{n+1} f(t)\,\mathrm{d}t\,.$$

En particulier, la série de terme général $f(n)$ et la suite $\left(\int_a^n f(t)\,\mathrm{d}t\right)_n$ sont de même nature.

EXERCICES

Vrai ou faux ?

		Vrai	Faux
a)	Si la série de terme général $u_n \in \mathbb{R}$ converge, alors la suite u_n converge.	☐	☐
b)	Si la suite u_n converge vers 0, alors la série de terme général u_n converge.	☐	☐
c)	La série de terme général $n^{-\alpha}$ converge si, et seulement si, $\alpha > 1$.	☐	☐
d)	La somme de deux séries divergentes est divergente.	☐	☐
e)	La somme de deux séries divergentes à termes positifs est divergente.	☐	☐
f)	Si la série de terme général $u_n \in \mathbb{R}$ converge, alors la série de terme général $\lvert u_n \rvert$ converge.	☐	☐
g)	Si la série de terme général $\lvert u_n \rvert \in \mathbb{R}$ converge, alors la série de terme général u_n converge.	☐	☐
h)	Si la série de terme général $u_n \in \mathbb{R}$ converge, alors la série de terme général $(-1)^n u_n$ converge.	☐	☐
i)	Si la série de terme général $u_n \in \mathbb{R}$ converge, alors la série de terme général u_n^3 converge.	☐	☐
j)	Si la série de terme général $u_n \in \mathbb{C}$ converge, alors la série de terme général u_n^2 converge.	☐	☐
k)	Si $u_n = o\left(n^{-1}\right)$, alors la série de terme général u_n converge.	☐	☐
l)	Si $u_n = o\left((-1)^n n^{-1}\right)$, alors la série de terme général u_n converge.	☐	☐
m)	Si $u_n = o\left(n^{-2}\right)$, alors la série de terme général u_n converge.	☐	☐
n)	Si (u_n) décroît vers 0 et si v_n est le terme général d'une série convergente, alors la série de terme général $u_n v_n$ converge.	☐	☐
o)	Si $u_n < v_n$ et si la série de terme général v_n converge, alors la série de terme général u_n converge.	☐	☐
p)	Si $u_n \sim v_n$, alors les séries de termes généraux u_n et v_n sont de même nature.	☐	☐

Exercices d'application du chapitre 22

●○○ **Exercice 1**

Montrer la convergence puis calculer la somme des séries de terme général

1. $\frac{1}{n(4n^2-1)}$
2. $\ln(1+\frac{2}{n(n+3)})$

●○○ **Exercice 2**

Déterminer la nature de la série de terme général u_n où $u_n = \frac{1}{n}$ si n est un carré et $u_n = \frac{1}{n^2}$ sinon.

●○○ **Exercice 3**

Déterminer, selon les valeurs de a, b, $c \in \mathbb{R}$, la nature de la série de terme général

$$u_n = a\sqrt{n} + b\sqrt{n+1} + c\sqrt{n+2}.$$

●●○ **Exercice 4**

Soit, pour tout $n \in \mathbb{N} \setminus \{0\}$, $c(n)$ le nombre de chiffres dans l'écriture décimale de n. Montrer la convergence de la série de terme général $u_n = \frac{c(n)}{n(n+1)}$ puis calculer sa somme.

●●○ **Exercice 5**

Déterminer la nature de la série de terme général $\frac{1}{a^{c(n)}}$ où $a > 0$ et $c(n)$ est le nombre de chiffres dans l'écriture décimale de n.

●○○ **Exercice 6**

1. Soit $k \in \mathbb{N}$. Montrer que la série de terme général $\frac{1}{n} - \frac{1}{n+k}$ converge.
   Notons u_k sa limite.
2. Déterminer un équivalent de u_k.

●●○ **Exercice 7**

Déterminer la somme de la série de terme général $\frac{n-a\lfloor \frac{n}{a} \rfloor}{n(n+1)}$ où $a \in \mathbb{N} \setminus \{0\}$.

●○○ **Exercice 8**

Déterminer un équivalent de la somme $\sum_{k=1}^{n} \frac{1}{k^\alpha}$ pour $\alpha \in]0,1[$.

●○○ **Exercice 9**

Montrer que $\displaystyle\sum_{k=2}^{n} \frac{1}{k \ln k} \sim \ln(\ln n)$.

●●○ **Exercice 10**

Soit $(u_n)_n$ et $(v_n)_n$ deux suites positives telles que les séries associées soient convergentes. Montrer que la série de terme général $\max(u_n, v_n)$ et la série de terme général $\frac{u_n v_n}{u_n + v_n}$ sont convergentes.

●●○ **Exercice 11**

Soit $(u_n)_n \in (\mathbb{R}_+)^{\mathbb{N}}$.

1. Montrer que si la série de terme général u_n converge, alors la série de terme général $\sqrt{u_n u_{n+1}}$ converge également.
2. Étudier la réciproque.

●●○ **Exercice 12**

Soit $(u_n)_n$ une suite de réels positifs. Montrer que la série de terme général u_n converge si, et seulement si, la série de terme général $\frac{u_n}{1+u_n}$ converge.

●●● **Exercice 13**

Soit $(a_n)_n$ une suite d'entiers naturels supérieurs ou égaux à 2 vérifiant

- pour tout n, a_n n'est pas un nombre premier ;
- pour tous $n \neq m$, a_n et a_m sont premiers entre eux.

Étudier la convergence de la série de terme général $\frac{1}{a_n}$.

Exercices d'approfondissement du chapitre 22

●○○ **Exercice A**

Déterminer la nature de la série de terme général u_n défini par

$$u_n = \sum_{k=n+1}^{+\infty} \frac{1}{k^2}.$$

●●○ **Exercice B**

Déterminer la nature de la série de terme général $\frac{1}{\sum_{k=1}^{n} \ln^2 k}$.

●●● **Exercice C**

Soit $(u_n)_n \in (\mathbb{R}_+^*)^{\mathbb{N}}$ strictement décroissante.

Déterminer la nature de la série de terme général $\frac{u_n - u_{n+1}}{u_n}$.

●●○ **Exercice D**

Soit $(u_n)_n$ une suite décroissante telle que la série de terme général u_n converge.

1. Justifier que, pour tout $n \in \mathbb{N}$, $u_n \geq 0$.
2. Montrer que $u_n = o\left(\frac{1}{n}\right)$.
   Indication : on pourra étudier la limite du terme

$$\sum_{k=\lfloor n/2 \rfloor}^{n} u_k.$$

●●○ **Exercice E**

Soit $(u_n)_n$ et $(v_n)_n$ deux suites ne s'annulant pas telles que

- la suite $(v_n)_n$ est le terme général d'une série positive convergente ;
- pour tout $n \in \mathbb{N}$,

$$\left|\frac{u_{n+1}}{u_n}\right| \leq \frac{v_{n+1}}{v_n}.$$

Montrer que la série de terme général u_n est convergente.

EXERCICES

●●● **Exercice F Sommation des équivalents**

Soit $(u_n)_n$ et $(v_n)_n$ deux suites positives équivalentes. Montrer que

- si les séries de termes généraux u_n et v_n divergent, alors $\sum_{k=0}^{n} u_k \sim \sum_{k=0}^{n} v_k$;
- si les séries de termes généraux u_n et v_n convergent, alors $\sum_{k=n}^{\infty} u_k \sim \sum_{k=n}^{\infty} v_k$.

●●● **Exercice G Critère de condensation**

Soit $(u_n) \in \mathbb{R}_+^{\mathbb{N}}$ décroissante. Définissons la suite $(v_n)_n$ par $v_n = 2^n u_{2^n}$. Montrer que les séries de terme général u_n et v_n sont de même nature.

●●○ **Exercice H**

Soit $(u_n)_n \in (\mathbb{R}_+)^{\mathbb{N}}$ et $(v_n)_n$ la suite définie par

$$v_n = \frac{u_n}{\prod_{k=0}^{n}(1+u_k)}.$$

1. Montrer que la série de terme général v_n est convergente.
2. Montrer que, si la série de terme général u_n est divergente, alors la somme de la série de terme général v_n est égale à 1.

●●● **Exercice I**

Soit $(u_n)_n \in (\mathbb{R}_+)^{\mathbb{N}}$ telle que la série de terme général u_n soit convergente. Montrer que

$$\sum_{k=0}^{n} k\, u_k = o(n).$$

●●○ **Exercice J**

Soit $(u_n) \in \mathbb{R}_+^{\mathbb{N}}$. Définissons la suite $(v_n)_n$ par

$$v_n = \frac{1}{n(n+1)} \sum_{k=1}^{n} k\, u_k.$$

Montrer que les séries de terme général u_n et v_n sont de même nature.

●●● **Exercice K Critère des séries alternées**

Soit $(u_n)_n \in (\mathbb{R}_+)^{\mathbb{N}}$ une suite décroissante de limite nulle.

1. Montrer que la série de terme général $(-1)^n u_n$ est convergente.
2. Montrer de plus que le $n^{\text{ième}}$ reste partiel est majoré en valeur absolue par u_{n+1}.
3. Étudier la convergence de la série de terme général $\frac{(-1)^n}{n^\alpha}$ avec $\alpha \in \mathbb{R}$.
4. Déterminer la nature de la série de terme général $\sin(\pi\sqrt{n^2+1})$.

Problèmes du chapitre 22

●●○ **Problème 1**

Attention, afin de simplifier les notations, toutes les suites considérées dans ce problème sont indexées à partir de 1.

Partie A – Sommation au sens de Cesàro

Commençons par quelques définitions

Définition 22.19.

Une suite $(u_n)_{n\geq 1}$ converge vers ℓ au sens de Cesàro si

$$\frac{1}{n}\sum_{k=1}^{n} u_k \to \ell.$$

La série de terme général u_n converge au sens de Cesàro si la suite de ses sommes partielles converge au sens de Cesàro.

On rappelle le résultat suivant.

Proposition 22.20.

Une suite convergente de limite ℓ converge au sens de Cesàro vers ℓ.

1. Montrer que la série de terme général $(-1)^{n-1}$ converge au sens de Cesàro.
2. Déterminer les $\alpha \in \mathbb{R}$ tel que la série de terme général n^α converge au sens de Cesàro.

Partie B – Sommation avec fonction de coupure

Définition 22.21.

Une fonction de coupure est une fonction $\eta : \mathbb{R}_+ \to \mathbb{R}_+$ telle que
- η est bornée,
- $\eta(0) = 1$ et η est continue en 0,
- $\forall x > 1,\ \eta(x) = 0$.

Définition 22.22.

Soit η une fonction de coupure. La série de terme général u_n converge vers ℓ selon η si

$$\sum_{k=1}^{n} u_k \eta\left(\frac{k}{n}\right) \to \ell,$$

ou de manière équivalente

$$\sum_{k=1}^{\infty} u_k \eta\left(\frac{k}{n}\right) \to \ell.$$

1. Déterminer une fonction de coupure η telle que la convergence des séries (au sens habituel) soit la convergence selon η.
2. Montrer que la convergence selon $x \mapsto (1-x)_+$ est la convergence des séries au sens de Cesàro.
3. Soit η une fonction de coupure. Montrer que si une série est absolument convergente, alors elle converge vers la même limite selon η.
4. Soit η une fonction de coupure.
   a. Montrer que
   $$\sum_{k=1}^{\infty}(-1)^{k-1}\eta\left(\frac{k}{n}\right) = \frac{1}{2}\eta\left(\frac{1}{n}\right) + \frac{1}{2}\sum_{k=1}^{\infty}\left(\eta\left(\frac{2k+1}{n}\right) - 2\eta\left(\frac{2k}{n}\right) + \eta\left(\frac{2k-1}{n}\right)\right).$$
   b. Montrer que, si η est de classe $\mathscr{C}^2$, la série de terme général $(-1)^{n-1}$ converge vers $\frac{1}{2}$ selon η.
   c. Le résultat reste-t-il vrai si la fonction η est seulement de classe $\mathscr{C}^1$?

PROBLÈMES

●○○ **Problème 2 Séries de Dirichlet**

Dans ce sujet, on considère une suite $(a_n)_{n\geq 1}$ à valeurs complexes et on étudie la convergence de la série (dite de Dirichlet) dont le terme général est $\frac{a_n}{n^s}$ pour différentes valeurs du paramètre $s \in \mathbb{C}$.

On introduit les deux ensembles suivants

$$\Lambda_a = \{x \in \mathbb{R}, \text{ la série de terme général } \tfrac{a_n}{n^x} \text{ converge absolument}\}$$
$$\Lambda_c = \{x \in \mathbb{R}, \text{ la série de terme général } \tfrac{a_n}{n^x} \text{ converge}\}$$

et leurs bornes inférieures $\sigma_a = \inf \Lambda_a$ et $\sigma_c = \inf \Lambda_c$ qui appartiennent à $\mathbb{R} \cup \{\pm\infty\}$ (avec les conventions que la borne inférieure d'une partie non minorée est $-\infty$ et que la borne inférieure de la partie vide est $+\infty$).

Partie A – Exemples

1. Préciser le module de $\frac{1}{n^s}$ avec $n \in \mathbb{N} \setminus \{0\}$ et $s \in \mathbb{C}$.
2. Calculer σ_c et σ_a dans le cas où la suite $(a_n)_n$ est la suite constante égale à 1.
3. Calculer σ_c et σ_a dans le cas où la suite $(a_n)_n$ est donnée par $a_n = 2^n$.
4. Calculer σ_c et σ_a dans le cas où la suite $(a_n)_n$ est donnée par $a_n = 2^{-n}$.
5. Calculer σ_c et σ_a dans le cas où la suite $(a_n)_n$ est donnée par $a_n = a^n$ où $a \in \mathbb{C} \setminus \mathbb{U}$.
6. Considérons, dans cette question, la suite $(a_n)_n$ telle que a_n est le nombre de diviseurs positifs de n.
   a. Notons $n = \prod_{k=1}^{r} p_k^{\alpha_k}$ la décomposition en facteurs premiers de n. Déterminer a_n.
   Fixons désormais $\varepsilon \in]0,1]$.
   b. Justifier que, pour tout k tel que $p_k \geq 2^{\frac{1}{\varepsilon}}$,
   $$\frac{\alpha_k + 1}{p_k^{\varepsilon \alpha_k}} \leq 1.$$
   c. Étudier les variations de la fonction $f : x \mapsto \frac{x+1}{2^{\varepsilon x}}$ sur $\mathbb{R}_+$.
   d. En déduire qu'il existe C (qui dépend de ε) tel que, pour tout $n \geq 1$, $a_n \leq C n^{\varepsilon}$.
   e. Conclure que $\sigma_c \leq 1$.
   f. Déterminer la valeur de σ_c.

Partie B – Abscisse σ_a de convergence absolue

1. Considérons $s_0 \in \mathbb{C}$ tel que la série de terme général $\frac{a_n}{n^{s_0}}$ converge absolument et $s \in \mathbb{C}$ vérifiant $\mathfrak{Re}(s) > \mathfrak{Re}(s_0)$. Montrer que la série de terme général $\frac{a_n}{n^s}$ converge absolument.
2. Montrer que si $\sigma_a = -\infty$, alors la série de terme général $\frac{a_n}{n^s}$ converge absolument pour tout $s \in \mathbb{C}$.
3. Montrer que si $\sigma_a = +\infty$, alors la série de terme général $\frac{a_n}{n^s}$ ne converge absolument pour aucun complexe s.
4. Supposons $\sigma_a \in \mathbb{R}$ et $s \in \mathbb{C}$. Montrer les implications suivantes :
   - si $\mathfrak{Re}(s) > \sigma_a$, alors la série de terme général $\frac{a_n}{n^s}$ converge absolument ;
   - si $\mathfrak{Re}(s) < \sigma_a$, alors la série de terme général $\frac{a_n}{n^s}$ ne converge pas absolument.

Partie C – Abscisse σ_c de convergence

1. Soit $z \in \mathbb{C}$ tel que $\mathfrak{Re}(z) > 0$ et $u_n = n^{-z}$ pour tout $n \geq 1$.
   a. Justifier que
   $$u_{n+1} - u_n = -\frac{z}{n^{z+1}} + O\left(\frac{1}{n^{z+1}}\right).$$
   b. En déduire que la série de terme général $u_{n+1} - u_n$ converge absolument.
2. Considérons $s_0 \in \mathbb{C}$ tel que la série de terme général $\frac{a_n}{n^{s_0}}$ converge et s tel que $\mathfrak{Re}(s) > \mathfrak{Re}(s_0)$. Montrer que la série de terme général $\frac{a_n}{n^s}$ converge.
   Indication : on pourra exploiter la question précédente pour établir la propriété de Cauchy.
3. Énoncer et justifier des résultats analogues à **B2**, **B3** et **B4** pour l'abscisse σ_c.

Partie D – Inégalités

1. a. Justifier que $\sigma_c \leq \sigma_a$.
   b. Montrer que cette inégalité est optimale.
2. a. Montrer par l'absurde que $\sigma_a \leq \sigma_c + 1$.
   b. En étudiant le cas de la suite $(a_n)_n$ donnée par $a_n = (-1)^n$, montrer que cette inégalité est optimale.

CORRIGÉS

Corrigés des Vrai/Faux

a) Vrai.

b) Faux, avec la série harmonique par exemple.

c) Vrai.

d) Faux avec les suites constantes $u_n = -v_n = 1$.

e) Vrai.

f) Faux, par exemple avec $\frac{(-1)^n}{n}$.

g) Vrai, la convergence absolue entraîne la convergence.

h) Faux avec, par exemple, $\frac{(-1)^n}{n}$.

i) Faux avec, par exemple, la suite définie par $u_{3n+1} = u_{3n+2} = \frac{1}{n^{1/3}}$, $u_{3n+3} = \frac{-2}{n^{1/3}}$.

j) Faux avec, par exemple, $\frac{(-1)^n}{\sqrt{n}}$.

k) Faux avec, par exemple, $\frac{1}{n \ln n}$.

l) Faux avec, par exemple, $\frac{1}{n^{3/4}}$.

m) Vrai.

n) Faux.

o) Faux si ces suites ne sont pas de signe contant.

p) Faux si ces suites ne sont pas de signe contant.

Corrigés des exercices d'application du chapitre 22

Exercice 1

1. Notons $u_n = \frac{1}{n(4n^2-1)}$. On a immédiatement que la série de terme général u_n converge car $u_n \geq 0$ et $u_n \sim \frac{1}{4n^3}$. La décomposition en éléments simples donne

$$u_n = -\frac{1}{n} + \frac{1}{2n+1} + \frac{1}{2n-1}.$$

En notant H_n la $n^{\text{ième}}$ somme partielle de la série harmonique dont un développement asymptotique est $\ln n+\gamma+o(1)$, on obtient

$$\begin{aligned}\sum_{k=1}^{n} u_k &= -\sum_{k=1}^{n}\frac{1}{k}+\sum_{k=1}^{n}\frac{1}{2k+1}+\sum_{k=1}^{n}\frac{1}{2k-1}\\ &= -H_n+H_{2n+1}-1-\sum_{k=1}^{n}\frac{1}{2k}+H_{2n}-\sum_{k=1}^{n}\frac{1}{2k}\\ &= -1-H_n+H_{2n+1}-\frac{1}{2}H_n+H_{2n}-\frac{1}{2}H_n\\ &= H_{2n+1}+H_{2n}-2H_n-1\\ &= \ln(2n(2n+1))-2\ln n-1+o(1)\to \ln 4-1.\end{aligned}$$

La somme de la série de terme général u_n est donc $\ln 4-1$.

2. Notons $u_n=\ln(1+\frac{2}{n(n+3)})=\ln\frac{(n+1)(n+2)}{n(n+3)}=\ln\frac{n+2}{n}-\ln\frac{(n+1)+2}{(n+1)}$. La somme est alors télescopique : pour tout $n\geq 1$,

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

Comme $\ln\frac{(n+1)+2}{(n+1)}\to 0$, la série de terme général u_n est convergente de somme $\ln 3$.

Exercice 2

Majorons la somme partielle au rang $n\geq 1$.

$$\begin{aligned}\sum_{k=1}^{n} u_k &= \sum_{\substack{k=1\\ k \text{ non carré}}}^{n}\frac{1}{k^2}+\sum_{\substack{k=1\\ k \text{ carré}}}^{n}\frac{1}{k}\\ &= \sum_{\substack{k=1\\ k \text{ non carré}}}^{n}\frac{1}{k^2}+\sum_{k=1}^{\lfloor\sqrt{n}\rfloor}\frac{1}{k^2}\\ &\leq 2\sum_{k=1}^{n}\frac{1}{k^2}\leq 2\sum_{k=1}^{+\infty}\frac{1}{k^2}.\end{aligned}$$

La suite des sommes partielles est croissante (car le terme général est positif), majorée donc convergente : la série est convergente.

Exercice 3

Pour tout $n\in\mathbb{N}$,

$$\begin{aligned}u_n &= a\sqrt{n}+b\sqrt{n}(1+n^{-1})^{\frac{1}{2}}+c\sqrt{n}(1+2n^{-1})^{\frac{1}{2}}\\ &= a\sqrt{n}+b\sqrt{n}\left(1+\frac{1}{2n}-\frac{1}{8n^2}+o\left(\frac{1}{n^2}\right)\right)+c\sqrt{n}\left(1+\frac{1}{n}-\frac{1}{2n^2}+o\left(\frac{1}{n^2}\right)\right)\\ &= (a+b+c)\sqrt{n}+\left(\frac{b}{2}+c\right)\frac{1}{\sqrt{n}}+\left(-\frac{b}{8}-\frac{c}{2}\right)\frac{1}{n\sqrt{n}}+o\left(\frac{1}{n\sqrt{n}}\right).\end{aligned}$$

- Si $a+b+c\neq 0$, alors la suite $(u_n)_n$ diverge vers $+\infty$ ou $-\infty$ donc la série diverge grossièrement.
- Sinon,

– si $\frac{b}{2}+c\neq 0$, alors, $u_n\sim\left(\frac{b}{2}+c\right)\frac{1}{\sqrt{n}}$ et la série de terme général $\frac{1}{\sqrt{n}}$ diverge donc la série de terme général u_n diverge ;

CORRIGÉS

– sinon, $u_n = O\left(\frac{1}{n\sqrt{n}}\right)$ et la série de terme général $\frac{1}{n\sqrt{n}}$ converge donc, par comparaison, la série de terme général u_n converge.
En conclusion, la série de terme général u_n converge si, et seulement si, $a+b+c=0$ et $b+2c=0$, c'est-à-dire si $b=-2c$ et $a=c$.

Exercice 4

▹ Remarquons tout d'abord que $\log n \leq c(n) \leq \log n + 1$ donc $c(n) \sim \log n$. Par conséquent, $n^{\frac{3}{2}} u_n \sim \frac{\log n}{\sqrt{n}}$ et donc $u_n = o(n^{-\frac{3}{2}})$: la série de terme général u_n converge par comparaison à une série de Riemann convergente.
▹ Pour tout $n \in \mathbb{N} \setminus \{0\}$,

$$\begin{aligned}
\sum_{k=1}^{10^n} u_k &= \sum_{k=1}^{10^n} \frac{c(k)}{k} - \sum_{k=1}^{10^n} \frac{c(k)}{k+1} \\
&= \sum_{k=1}^{10^n} \frac{c(k)}{k} - \sum_{k=2}^{10^n+1} \frac{c(k-1)}{k} \\
&= \sum_{k=2}^{10^n} \frac{c(k)-c(k-1)}{k} + c(1) - \frac{c(10^n)}{10^n+1} \\
&= \sum_{l=1}^{n} \frac{1}{10^l} + 1 - \frac{n}{10^n+1},
\end{aligned}$$

car la quantité $c(k)-c(k-1)$ vaut 1 si k est une puissance de 10, 0 sinon. Par conséquent, la somme de la série est

$$\lim \sum_{l=1}^{n} \left(\frac{1}{10^l} + 1 - \frac{n}{10^n+1} \right) = \frac{1}{10} \frac{1}{1-\frac{1}{10}} + 1 = \frac{10}{9}.$$

Exercice 5

Comme la série est à termes positifs, soit elle converge, soit elle diverge vers $+\infty$. Pour obtenir la nature de la série, il suffit donc d'étudier la suite des sommes partielles

$$\left(\sum_{k=1}^{10^n-1} \frac{1}{a^{c(k)}} \right)_n.$$

Or, en regroupant par paquets,

$$\begin{aligned}
\sum_{k=1}^{10^n-1} \frac{1}{a^{c(k)}} &= \sum_{l=0}^{n-1} \sum_{k=10^l}^{10^{l+1}-1} \frac{1}{a^{l+1}} \\
&= \sum_{l=0}^{n-1} \frac{9.10^l}{a^{l+1}} = \frac{9}{a} \sum_{l=0}^{n-1} \left(\frac{10}{a} \right)^l.
\end{aligned}$$

Ainsi, la série converge si, et seulement si, $a > 10$ et la somme est alors $\frac{9}{a-10}$.

Exercice 6

1. Le terme général vaut (après réduction au même dénominateur) $\frac{k}{n(n+k)}$. Ce terme est équivalent à $\frac{k}{n^2}$ quand n tend vers $+\infty$ qui est le terme général d'une série de Riemann absolument convergente. Ainsi, la série considérée converge absolument.
2. Calculons une forme plus maniable pour u_k en repassant par les sommes finies. Pour tout $N \geq k$,

$$\begin{aligned}\sum_{n=1}^{N} \frac{1}{n} - \frac{1}{n+k} &= \sum_{n=1}^{N} \frac{1}{n} - \sum_{n=1}^{N} \frac{1}{n+k} \\ &= \sum_{n=1}^{N} \frac{1}{n} - \sum_{n=k+1}^{k+N} \frac{1}{n} \\ &= \sum_{n=1}^{k} \frac{1}{n} - \sum_{n=N+1}^{k+N} \frac{1}{n} \\ &= \sum_{n=1}^{k} \frac{1}{n} - \sum_{n=1}^{k} \frac{1}{N+n}.\end{aligned}$$

Comme les sommes sont finies,

$$\sum_{n=1}^{k} \frac{1}{N+n} \underset{N\to+\infty}{\longrightarrow} 0,$$

donc $u_k = \sum_{n=1}^{k} \frac{1}{n}$. Remarquons que ce terme est la somme harmonique jusqu'au rang k et se trouve donc équivalent à $\ln k$. En conclusion, $u_k \sim \ln k$.

Exercice 7

Remarquons tout d'abord que la série est convergente car le terme général est majoré par $\frac{a}{n^2}$. Alors, pour tout N, la somme partielle jusqu'au rang $Na+a-1$ vaut

$$\begin{aligned}\sum_{k=1}^{Na+a-1} \frac{k-a\left\lfloor \frac{k}{a} \right\rfloor}{k(k+1)} &= \sum_{k=a}^{Na+a-1} \frac{k-a\left\lfloor \frac{k}{a} \right\rfloor}{k(k+1)} + \sum_{k=1}^{a-1} \frac{1}{k+1} \\ &= \sum_{q=1}^{N}\sum_{r=0}^{a-1} \frac{r}{(qa+r)(qa+r+1)} + \sum_{k=1}^{a-1} \frac{1}{k+1} \\ &= \sum_{q=1}^{N}\sum_{r=0}^{a-1} \left(\frac{r}{qa+r} - \frac{r}{qa+r+1}\right) + \sum_{k=1}^{a-1} \frac{1}{k+1} \\ &= \sum_{q=1}^{N}\sum_{r=0}^{a-1} \frac{r}{qa+r} - \sum_{q=1}^{N}\sum_{r=1}^{a} \frac{r-1}{qa+r} + \sum_{k=1}^{a-1} \frac{1}{k+1} \\ &= \sum_{q=1}^{N}\sum_{r=0}^{a-1} \frac{1}{qa+r} - \frac{1}{a}\sum_{q=1}^{N} \frac{1}{q} - \frac{a-1}{a}\sum_{q=1}^{N} \frac{1}{q+1} + \sum_{k=1}^{a-1} \frac{1}{k+1} \\ &= \sum_{n=1}^{Na+a-1} \frac{1}{n} - \sum_{n=1}^{N} \frac{1}{n} - \frac{a-1}{a(N+1)} \\ &= \ln(Na+a-1) + \gamma - \ln(N) - \gamma + o(1) \to \ln a.\end{aligned}$$

La somme de la série est donc $\ln a$.

CORRIGÉS

Exercice 8

La fonction $f : t \to t^{-\alpha}$ est décroissante sur $[1,+\infty[$. D'après le résultat de comparaison entre série et intégrale,

$$f(1)+\int_1^n f(t)\,\mathrm{d}t \geq \sum_{k=1}^n \frac{1}{k^\alpha} \geq \int_1^{n+1} f(t)\,\mathrm{d}t\,.$$

En calculant l'intégrale,

$$1+\frac{1}{1-\alpha}(n^{1-\alpha}-1) \geq \sum_{k=1}^n \frac{1}{k^\alpha} \geq \frac{1}{1-\alpha}((n+1)^{1-\alpha}-1).$$

Les deux membres sont équivalents à $\frac{1}{1-\alpha}n^{1-\alpha}$, donc, par encadrement,

$$\sum_{k=1}^n \frac{1}{k^\alpha} \sim \frac{1}{1-\alpha}n^{1-\alpha}.$$

Exercice 9

La fonction $f : t \mapsto \frac{1}{t\ln t}$ est décroissante sur $[2,+\infty[$. D'après le résultat de comparaison entre série et intégrale,

$$f(2)+\int_2^n f(t)\,\mathrm{d}t \geq \sum_{k=2}^n \frac{1}{k\ln k} \geq \int_2^{n+1} f(t)\,\mathrm{d}t\,.$$

Or, $\int_a^b f(t)\,\mathrm{d}t = \ln(\ln b)-\ln(\ln a)$ (car $f = \frac{\ln'}{\ln}$). Par conséquent, chacun des membres de l'inégalité est équivalent à $\ln(\ln n)$ donc la somme considérée est équivalente à $\ln(\ln n)$.

Exercice 10

▹ Pour tout $n \in \mathbb{N}$, $0 \leq \max(u_n, v_n) \leq u_n + v_n$. La série de terme général $u_n + v_n$ est convergente comme somme de séries convergentes. Par comparaison des séries à termes positifs, la série de terme général $\max(u_n, v_n)$ est convergente.

▹ Pour tout $n \in \mathbb{N}$,

$$0 \leq \frac{u_n v_n}{u_n + v_n} \leq u_n \frac{u_n + v_n}{u_n + v_n} = u_n,$$

donc, par comparaison des séries à termes positifs, la série de terme général $\frac{u_n v_n}{u_n+v_n}$ est convergente.

Exercice 11

1. Remarquons que, pour tout n,

$$0 \leq \sqrt{u_n u_{n+1}} \leq \frac{1}{2}(u_n + u_{n+1}).$$

Comme les séries de termes généraux u_n et u_{n+1} convergent, la série de terme général $\frac{1}{2}(u_n + u_{n+1})$ converge. Par inégalité, la série de terme général $\sqrt{u_n u_{n+1}}$ converge.

2. La réciproque est fausse : on peut facilement construire un contre-exemple en prenant $u_n = 0$ si n est impair, $u_n = \frac{1}{n}$ si n est impair.

▹ Le terme $\sqrt{u_n u_{n+1}}$ est nul pour tout n donc la série de terme général $\sqrt{u_n u_{n+1}}$ est convergente.

▹ Les sommes partielles de la série de terme général u_n sont égales à la moitié des sommes partielles de la série harmonique. Par conséquent, la série de terme général u_n diverge.

Remarque

En revanche, si la suite $(u_n)_n$ est décroissante, la réciproque devient vraie en utilisant l'inégalité

$$u_n = \sqrt{u_n u_n} \leq \sqrt{u_{n-1} u_n}.$$

Exercice 12

($\Rightarrow$) Comme la série de terme général u_n converge, la suite $(u_n)_n$ tend vers 0 et donc $\frac{1}{1+u_n} \to 1$, c'est-à-dire $\frac{u_n}{1+u_n} \sim u_n$. Comme les séries sont à termes positifs, cela entraîne que la série de terme général $\frac{u_n}{1+u_n}$ converge.

($\Leftarrow$) Supposons que la série de terme général $v_n = \frac{u_n}{1+u_n}$ converge. Alors, la suite $(v_n)_n$ tend vers 0 et, au moins à partir d'un certain rang, $u_n = \frac{v_n}{1-v_n}$. Comme précédemment, $u_n \sim v_n$ et cette équivalence entre termes généraux positifs de séries dont une est convergente entraîne que la série de terme général u_n converge.

Exercice 13

Posons, pour tout $n \in \mathbb{N}$, p_n le plus petit facteur premier de a_n. D'après le deuxième point de la définition, pour tous $n \neq m$, $p_n \neq p_m$.

Par ailleurs, comme a_n n'est pas premier, $a_n \geq p_n^2$ (soit la valuation p_n-adique est supérieure ou égale à 1, soit a_n admet un facteur premier plus grand). Ainsi, $a_n \geq p_n^2$.

Combinant ces deux résultats, on obtient, pour tout $n \in \mathbb{N}$,

$$\sum_{k=0}^{n} \frac{1}{a_k} \leq \sum_{k=0}^{n} \frac{1}{p_k^2} \leq \sum_{l=1}^{\infty} \frac{1}{l^2}.$$

En conclusion, la suite des sommes partielles de la série de terme général $\frac{1}{a_n} \geq 0$ est majorée : la série converge.

Corrigés des exercices d'approfondissement du chapitre 22

Exercice A

Comme la fonction $f : t \mapsto t^{-2}$ est décroissante, le résultat de comparaison entre série et intégrale donne

$$\int_n^N \frac{\mathrm{d}t}{t^2} \geq \sum_{k=n+1}^{N} \frac{1}{k^2} \geq \int_{n+1}^{N+1} \frac{\mathrm{d}t}{t^2}$$

donc

$$\frac{1}{n} - \frac{1}{N} \geq \sum_{k=n+1}^{N} \frac{1}{k^2} \geq \frac{1}{n+1} - \frac{1}{N+1}.$$

En passant à la limite en $N \to +\infty$, on obtient $\frac{1}{n} \geq u_n \geq \frac{1}{n+1}$ donc $u_n \sim \frac{1}{n}$. Comme la série harmonique est divergente et que ces séries sont à termes équivalents et positifs, la série de terme général u_n diverge.

Exercice B

Commençons par remarquer qu'avec la monotonie de $t \mapsto \ln^2 t$, pour $k \geq 2$,

$$\int_{k-1}^{k} \ln^2 t \,\mathrm{d}t \leq \ln^2 k \leq \int_{k}^{k+1} \ln^2 t \,\mathrm{d}t\,.$$

Ainsi,

$$\int_{1}^{n} \ln^2 t \,\mathrm{d}t \leq \sum_{k=1}^{n} \ln^2 k \leq \int_{2}^{n+1} \ln^2 t \,\mathrm{d}t\,.$$

Or, par intégration par parties,

$$\begin{aligned}\int_{1}^{x} \ln^2 t \,\mathrm{d}t &= [\ln t(t\ln t - t)]_1^x - \int_1^x \frac{1}{t}(t\ln t - t)\,\mathrm{d}t \\ &= x\ln x(\ln x - 1) - x\ln x + x - 1 + x - 1 \sim x\ln^2 x.\end{aligned}$$

Ainsi, $\sum_{k=1}^{n} \ln^2 k \sim n\ln^2 n$.

La série de terme général $\frac{1}{\sum_{k=1}^{n} \ln^2 k}$ est donc de même nature que la série de terme général $\frac{1}{n\ln^2 n}$ dont on montre qu'elle est convergente par une nouvelle comparaison série-intégrale.

Exercice C

Il s'agit une nouvelle fois d'une comparaison série-intégrale : pour tout $n \in \mathbb{N}$,

$$\frac{u_n - u_{n+1}}{u_n} = \int_{u_{n+1}}^{u_n} \frac{\mathrm{d}t}{u_n} \leq \int_{u_{n+1}}^{u_n} \frac{\mathrm{d}t}{t}.$$

Par conséquent, pour tout n,

$$\sum_{k=0}^{n} \frac{u_k - u_{k+1}}{u_k} \leq \int_{u_{n+1}}^{u_0} \frac{\mathrm{d}t}{t} = \ln(u_0) - \ln(u_{n+1}).$$

La suite $(u_n)_n$ est décroissante minorée (par 0 notamment) donc converge vers une limite $\ell \geq 0$.

▹ Si $\ell > 0$, le terme de droite dans la dernière égalité est majoré, donc la somme partielle de la série considérée est majorée et converge (car le terme général est positif).

▹ Si $\ell = 0$, remarquons que

$$\begin{aligned}\frac{u_n - u_{n+1}}{u_n} &= \frac{u_n - u_{n+1}}{u_{n+1}}\left(\frac{u_{n+1}}{u_n}\right) \\ &= \frac{u_n - u_{n+1}}{u_{n+1}}\left(1 - \frac{u_n - u_{n+1}}{u_n}\right) \\ &\sim \frac{u_n - u_{n+1}}{u_{n+1}}.\end{aligned}$$

Les séries de termes généraux positifs $\frac{u_n - u_{n+1}}{u_n}$ et $\frac{u_n - u_{n+1}}{u_{n+1}}$ sont donc de même nature. Pour conclure, on écrit, pour tout $n \in \mathbb{N}$,

$$\frac{u_n - u_{n+1}}{u_{n+1}} = \int_{u_{n+1}}^{u_n} \frac{\mathrm{d}t}{u_{n+1}} \geq \int_{u_{n+1}}^{u_n} \frac{\mathrm{d}t}{t}.$$

Par conséquent, pour tout n,

$$\sum_{k=0}^{n} \frac{u_k - u_{k+1}}{u_{k+1}} \geq \int_{u_{n+1}}^{u_0} \frac{\mathrm{d}t}{t} = \ln(u_0) - \ln(u_{n+1}).$$

Comme le minorant diverge vers $+\infty$, les deux séries considérées sont divergentes.

En conclusion, la série de terme général $\frac{u_n - u_{n+1}}{u_n}$ converge si, et seulement si, $u_n \not\to 0$.

Exercice D

1. Comme la série de terme général u_n converge, $u_n \to 0$. Par décroissance de la suite $(u_n)_n$, $u_n \geq 0$ pour tout $n \in \mathbb{N}$.
2. Remarquons tout d'abord que la somme de l'indication tend vers 0 (comme différence de deux sommes partielles). Par ailleurs, par monotonie de la suite $(u_n)_n$,

$$\sum_{k=\lfloor n/2 \rfloor}^{n} u_k \geq \sum_{k=\lfloor n/2 \rfloor}^{n} u_n \geq (n - \lfloor n/2 \rfloor) u_n.$$

Par conséquent $(n - \lfloor n/2 \rfloor) u_n \to 0$; or, $n - \lfloor n/2 \rfloor \sim \frac{n}{2}$ donc $n u_n \to 0$.

Exercice E

Par une récurrence immédiate, $|u_n| \leq \frac{|u_0|}{v_0} v_n$; on conclut avec le résultat de comparaison entre séries à termes positifs.

Exercice F

▹ Montrons l'équivalence des sommes partielles dans le cas de séries divergentes. Soit $\varepsilon > 0$. Il existe un rang N tel que pour tout $n \geq N$, $|u_n - v_n| \leq \varepsilon v_n$. Ainsi, pour $n \geq N$,

$$\begin{aligned}\left|\sum_{k=0}^{n} u_k - \sum_{k=0}^{n} v_k\right| &\leq \sum_{k=0}^{n} |u_k - v_k| \\ &\leq \sum_{k=0}^{N-1} |u_k - v_k| + \sum_{k=N}^{n} |u_k - v_k| \\ &\leq \sum_{k=0}^{N-1} |u_k - v_k| + \varepsilon \sum_{k=0}^{n} v_k.\end{aligned}$$

Comme $\sum_{k=0}^{N-1} |u_k - v_k|$ est une constante et que la série de terme général v_n diverge, il existe $N' > 0$ tel que, pour $n \geq N'$,

$$\sum_{k=0}^{N-1} |u_k - v_k| \leq \varepsilon \sum_{k=0}^{n} v_k.$$

En conclusion, pour $n \geq \max\{N, N'\}$,

$$\left|\sum_{k=0}^{n} u_k - \sum_{k=0}^{n} v_k\right| \leq 2\varepsilon \sum_{k=0}^{n} v_k.$$

▹ Montrons l'équivalence des restes partiels dans le cas de séries convergentes. Soit $\varepsilon > 0$. Il existe un rang N tel que pour tout $n \geq N$, $|u_n - v_n| \leq \varepsilon v_n$. Ainsi, pour $p \geq n \geq N$,

$$\begin{aligned}\left|\sum_{k=n}^{p} u_k - \sum_{k=n}^{p} v_k\right| &\leq \sum_{k=n}^{p} |u_k - v_k| \\ &\leq \varepsilon \sum_{k=n}^{p} v_k.\end{aligned}$$

Ainsi, en faisant tendre p vers $+\infty$,

$$\left|\sum_{k=n}^{+\infty} u_k - \sum_{k=n}^{+\infty} v_k\right| \leq \varepsilon \sum_{k=n}^{+\infty} v_k.$$

En conclusion, $\sum_{k=n}^{\infty} u_k \sim \sum_{k=n}^{\infty} v_k$.

Exercice G

D'après la décroissance de la suite (u_n), on a, pour tout n,

$$\frac{1}{2} v_{n+1} = 2^n u_{2^{n+1}} \leq \sum_{k=2^n}^{2^{n+1}-1} u_k \leq 2^n u_{2^n} = v_n.$$

Ainsi, pour tout n,

$$\sum_{k=1}^{2^{n+1}-1} u_k \leq \sum_{k=0}^{n} v_k \leq 2 \sum_{k=1}^{2^{n+1}-1} u_k.$$

On en déduit aisément que les deux séries sont de même nature.

Exercice H

1. L'idée de la solution qui suit est de calculer les sommes partielles V_n de la série de terme général v_n. Les premiers termes sont

$$\begin{aligned} V_1 &= \frac{u_1}{1+u_1} = 1 - \frac{1}{1+u_1}, \\ V_2 &= V_1 + v_2 = 1 - \frac{1}{1+u_1} + \frac{u_2}{(1+u_1)(1+u_2)} \\ &= 1 - \frac{1}{(1+u_1)(1+u_2)}. \end{aligned}$$

Montrons par récurrence sur $n \in \mathbb{N} \setminus \{0\}$ que $V_n = 1 - \frac{1}{\prod_{k=1}^{n}(1+u_k)}$.

- L'initialisation est obtenue précédemment (c'est grâce à ces premiers calculs que l'on a obtenu la forme de l'hypothèse de récurrence).
- Soit $n \in \mathbb{N} \setminus \{0\}$ tel que $V_n = 1 - \frac{1}{\prod_{k=1}^{n}(1+u_k)}$. Alors

$$\begin{aligned} V_{n+1} = V_n + v_{n+1} &= 1 - \frac{1}{\prod_{k=1}^{n}(1+u_k)} + \frac{u_{n+1}}{\prod_{k=0}^{n+1}(1+u_k)} \\ &= 1 - \frac{(1+u_{n+1}) - u_{n+1}}{\prod_{k=0}^{n+1}(1+u_k)} \\ &= 1 - \frac{1}{\prod_{k=0}^{n+1}(1+u_k)}. \end{aligned}$$

Avec cette récurrence, nous avons obtenu que la série de terme général v_n positif est majorée par 1 donc convergente.

2. Reprenons le résultat démontré dans la récurrence de la question précédente. Ainsi, pour tout $n \in \mathbb{N} \setminus \{0\}$, on obtient l'encadrement

$$1 \geq V_n \geq 1 - \frac{1}{1 + \sum_{k=1}^{n} u_k},$$

puisque la suite $(u_n)_n$ est positive. Les deux termes extrêmes de cette inégalité tendent vers 1 (car la série de terme général u_n diverge vers $+\infty$) donc, par le lemme d'encadrement, $\lim V_n = 1$. Ainsi, si la série de terme général u_n est divergente, alors la somme de la série de terme général v_n est égale à 1.

Exercice I

Notons pour, tout $n \in \mathbb{N}$, $R_n = \sum_{k=n}^{+\infty} u_k$. Alors

$$\begin{aligned}
\frac{1}{n}\sum_{k=0}^{n} k\,u_k &= \frac{1}{n}\sum_{k=0}^{n} k(R_k - R_{k+1}) \\
&= \frac{1}{n}\sum_{k=0}^{n} kR_k - \frac{1}{n}\sum_{k=0}^{n} kR_{k+1} \\
&= \frac{1}{n}\sum_{k=0}^{n} kR_k - \frac{1}{n}\sum_{k=1}^{n+1}(k-1)R_k \\
&= \frac{1}{n}\sum_{k=1}^{n} R_k - R_{n+1}.
\end{aligned}$$

Or, la suite $(R_n)_n$ tend vers 0 (c'est la suite des restes d'une série convergente) donc $\frac{1}{n}\sum_{k=1}^{n} R_k \to 0$ (par exemple avec le théorème de Cesàro). En conclusion,

$$\frac{1}{n}\sum_{k=0}^{n} k\,u_k = o(1).$$

Exercice J

Notons, pour tout $n \in \mathbb{N}\setminus\{0\}$, $S_n = \sum_{k=1}^{n} k\,u_k$. Alors, pour tout $n \geq 1$,

$$\begin{aligned}
\sum_{k=1}^{n} v_k &= \sum_{k=1}^{n} \frac{1}{k} S_k - \sum_{k=1}^{n} \frac{1}{k+1} S_k \\
&= \sum_{k=2}^{n} \frac{1}{k}(S_k - S_{k-1}) + u_1 - \frac{1}{n+1} S_n \\
&= \sum_{k=1}^{n} u_k - n\,v_n.
\end{aligned}$$

- Si la série de terme général u_n converge, alors la série de terme général v_n converge car elle est à termes positifs et

$$\forall n, \qquad \sum_{k=1}^{n} v_k \leq \sum_{k=1}^{n} u_k \leq \sum_{k=1}^{\infty} u_k.$$

- Si la série de terme général u_n diverge, alors soit $(n\,v_n)$ diverge vers $+\infty$, soit $\left(\sum_{k=1}^{n} v_k\right)$ diverge vers $+\infty$. Dans les deux cas, la série de terme général v_n diverge.

Exercice K

1. Notons $(S_n)_n$ la suite des sommes partielles de la série de terme général $(-1)^n u_n$ et montrons que les suites $(S_{2n})_n$ et $(S_{2n+1})_n$ sont adjacentes.
- Pour tout $n \in \mathbb{N}$, $S_{2n+2} - S_{2n} = u_{2n+2} - u_{2n+1} < 0$. La suite $(S_{2n})_n$ est décroissante.
- Pour tout $n \in \mathbb{N}$, $S_{2n+3} - S_{2n+1} = u_{2n+2} - u_{2n+3} > 0$. La suite $(S_{2n+1})_n$ est croissante.
- $S_{2n+1} - S_{2n} = u_{2n+1} \to 0$.

Les suites $(S_{2n})_n$ et $(S_{2n+1})_n$ sont adjacentes donc convergent vers la même limite et donc $(S_n)_n$ converge.
2. Notons S la somme de cette série. Alors, les monotonies des suites $(S_{2n})_n$ et $(S_{2n+1})_n$ donnent

$$|S - S_n| \leq |S_{n+1} - S_n| = u_{n+1},$$

d'où la majoration annoncée sur le reste partiel.
3. Pour $\alpha \leq 0$, il y a divergence grossière ; pour $\alpha > 1$, il y a convergence absolue. Enfin, pour $\alpha \in]0,1]$, on peut appliquer le résultat de la première question et la série est (semi-)convergente.
4. Commençons par étudier le terme général

$$\begin{aligned}\sin(\pi\sqrt{n^2+1}) = \sin(n\pi\sqrt{1+n^{-2}}) &= \sin\Big(n\pi + \frac{\pi}{2n} + o\big(\frac{1}{n^2}\big)\Big)\\ &= (-1)^n \frac{\pi}{2n} + o\big(\frac{1}{n^2}\big).\end{aligned}$$

La série de terme général $(-1)^n \frac{\pi}{2n}$ est convergente d'après la première question ; une série dont le terme général est $o(\frac{1}{n^2})$ est absolument convergente par comparaison. En conclusion, la série de terme général $\sin(\pi\sqrt{n^2+1})$ est convergente.

Corrigés des problèmes du chapitre 22

Problème 1

Partie A – Sommation au sens de Cesàro

1. La suite des sommes partielles de la série de terme général est majorée par 1 donc converge vers 0 au sens de Cesàro.
2. ▹ Pour $\alpha < -1$, la série converge donc converge au sens de Cesàro.
▹ Pour $\alpha > -1$, la série diverge et une comparaison série-intégrale entraîne que la somme partielle S_n est équivalente à $\frac{1}{\alpha+1} n^{\alpha+1}$. En recommençant une comparaison série-intégrale, on obtient que

$$\sum_{k=1}^{n} S_k \sim \frac{1}{\alpha+1} \sum_{k=1}^{n} k^{\alpha+1} \sim \frac{1}{(\alpha+1)(\alpha+2)} n^{\alpha+2},$$

donc il n'y a pas convergence au sens de Cesàro.
▹ Pour $\alpha = -1$, la série diverge et une comparaison série-intégrale entraîne que la somme partielle S_n est équivalente à $\ln n$ En recommençant une comparaison série-intégrale, on obtient que

$$\sum_{k=1}^{n} S_k \sim \sum_{k=1}^{n} \ln k \sim n \ln n,$$

donc il n'y a pas convergence au sens de Cesàro.

Partie B – Sommation avec fonction de coupure

1. Il suffit de prendre pour η la fonction indicatrice de l'intervalle $[0,1]$. On obtient alors que

$$\sum_{k=1}^{n} u_k \eta\left(\frac{k}{n}\right) = \sum_{k=1}^{n} u_k$$

et donc le résultat voulu.
2. Avec cette fonction de coupure, on a, pour tout n,

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

Or,

$$\sum_{k=1}^{n-1}\sum_{l=1}^{k} u_l = \sum_{l=1}^{n-1}(n-l)u_l$$

En conclusion,

$$\sum_{k=1}^{n} u_k \eta\left(\frac{k}{n}\right) = \frac{n-1}{n}\left(\frac{1}{n-1}\sum_{k=1}^{n-1}\left(\sum_{l=1}^{k} u_l\right)\right)$$

et donc la convergence selon $x \mapsto (1-x)_+$ est la convergence des séries au sens de Cesàro.
3. Considérons $(u_n)_n$ telle que la série de terme général u_n soit absolument convergente. Soit $\varepsilon > 0$ et $\nu > 0$ tels que

$$\forall x \leq \nu, \qquad |\eta(x) - 1| \leq \varepsilon.$$

Soit N tel que $\sum_{k=N+1}^{\infty} |u_k| \leq \varepsilon$. Alors, pour tout $n \geq \max(N, \frac{N}{\nu})$,

$$\begin{aligned}
\left|\sum_{k=1}^{\infty} u_k \eta\left(\frac{k}{n}\right) - \sum_{k=1}^{\infty} u_k\right| &\leq \sum_{k=1}^{\infty} |u_k| \left|\eta\left(\frac{k}{n}\right) - 1\right| \\
&\leq \sum_{k=1}^{N} |u_k| \left|\eta\left(\frac{k}{n}\right) - 1\right| + \sum_{k=N+1}^{\infty} |u_k| \left|\eta\left(\frac{k}{n}\right) - 1\right| \\
&\leq \sum_{k=1}^{N} |u_k| \varepsilon + (1 + \|\eta\|_\infty) \sum_{k=N+1}^{\infty} |u_k| \\
&\leq (\sum_{k=1}^{N} |u_k| + 1 + \|\eta\|_\infty)\varepsilon.
\end{aligned}$$

Par conséquent, $\sum_{k=1}^{\infty} u_k \eta\left(\frac{k}{n}\right) \to \sum_{k=1}^{\infty} u_k$.
4. a. Partons du terme de droite noté D

$$\begin{aligned}
D &= \frac{1}{2}\eta\left(\frac{1}{n}\right) + \frac{1}{2}\sum_{k=1}^{\infty}\left(\eta\left(\frac{2k+1}{n}\right) - 2\eta\left(\frac{2k}{n}\right) + \eta\left(\frac{2k-1}{n}\right)\right) \\
&= \frac{1}{2}\eta\left(\frac{1}{n}\right) + \frac{1}{2}\sum_{k=1}^{\infty}\left(\eta\left(\frac{2k+1}{n}\right) - 2\eta\left(\frac{2k}{n}\right)\right) + \frac{1}{2}\sum_{k=1}^{\infty}\left(-\eta\left(\frac{2k}{n}\right) + \eta\left(\frac{2k-1}{n}\right)\right) \\
&= 1.\frac{1}{2}\sum_{k=1}^{\infty}(-1)^{k-1}\eta\left(\frac{k}{n}\right).
\end{aligned}$$

b. Remarquons tout d'abord que $\frac{1}{2}\eta\left(\frac{1}{n}\right) \to \frac{1}{2}$ par continuité de η en 0. Par ailleurs, d'après la formule de Taylor-Young à l'ordre 2 en $\frac{2k}{n}$:

$$\eta\left(\frac{2k+1}{n}\right) - 2\eta\left(\frac{2k}{n}\right) + \eta\left(\frac{2k-1}{n}\right) = \frac{1}{n^2}\eta''\left(\frac{2k}{n}\right) + o(\frac{1}{n^2}).$$

Par ailleurs, la somme considérée comporte au plus n termes donc

$$\frac{1}{2}\sum_{k=1}^{\infty}\left(\eta\left(\frac{2k+1}{n}\right)-2\eta\left(\frac{2k}{n}\right)+\eta\left(\frac{2k-1}{n}\right)\right)=o(\frac{1}{n}).$$

En conclusion,

$$\sum_{k=1}^{\infty}(-1)^{k-1}\eta\left(\frac{k}{n}\right)\to\frac{1}{2}.$$

c. Oui, avec la propriété des accroissements finis sur chaque intervalle $[(k-1)/n, k/n]$ et $[k/n,(k+1)/n]$, il existe des réels c_k et $\tilde{c}_k$ dans ces intervalles tels que la somme de droite se réécrive

$$\frac{1}{n}\sum_{k=1}^{\infty}\left(\eta'(2c_k)-\eta'(2\tilde{c}_k)\right)$$

donc soit de limite $\int_0^1(\eta'(2t)-\eta'(2t))\,\mathrm{d}t=0$ avec le résultat sur les sommes de Riemann.

Problème 2

Partie A – Exemples

1. Notons $z=x+iy$ avec $x,\,y\in\mathbb{R}$. Alors, $\frac{1}{n^s}=\frac{1}{n^x}e^{-iy\ln n}$ donc $|\frac{1}{n^s}|=\frac{1}{n^x}$.
2. La série de terme général $\frac{1}{n^x}$ est à termes positifs donc elle converge si, et seulement si, elle converge absolument : $\sigma_c=\sigma_a$.
Par ailleurs, il s'agit d'une série de Riemann donc elle converge si, et seulement si, $x>1$. Par conséquent, on obtient $\sigma_c=\inf]1,+\infty[=1$.
3. La série de terme général $\frac{2^n}{n^x}$ est à termes positifs donc elle converge si, et seulement si, elle converge absolument : $\sigma_c=\sigma_a$.
Or, $\frac{2^n}{n^x}=e^{n\ln 2-x\ln n}\to+\infty$. La série de terme général $\frac{2^n}{n^x}$ diverge grossièrement pour tout $x\in\mathbb{R}$ donc on trouve $\sigma_c=\inf\varnothing=+\infty$.
4. Encore une fois, la série de terme général $\frac{2^{-n}}{n^x}$ est à termes positifs donc elle converge si, et seulement si, elle converge absolument : $\sigma_c=\sigma_a$.
Or, $n^2\frac{2^{-n}}{n^x}=e^{-n\ln 2+(2-x)\ln n}\to 0$. Ainsi, $\frac{2^{-n}}{n^x}=o(\frac{1}{n^2})$. Par comparaison d'une série à termes positifs avec une série de Riemann, la série de terme général $\frac{2^{-n}}{n^x}$ converge pour tout $x\in\mathbb{R}$. En conclusion, on obtient $\sigma_c=\inf\mathbb{R}=-\infty$.
5. ▹ Si $|a|>1$, on procède comme dans la question **3** et on obtient que la série de terme général $\frac{a^n}{n^x}$ diverge grossièrement pour tout $x\in\mathbb{R}$ donc $\sigma_c=\sigma_a=+\infty$.
▹ Si $|a|<1$, on procède comme dans la question **4** et on obtient que la série de terme général $\frac{a^n}{n^x}$ converge absolument pour tout $x\in\mathbb{R}$ donc $\sigma_c=\sigma_a=-\infty$.
6. a. Un entier d divise n si, et seulement s'il est de la forme $\prod_{k=1}^{r}p_k^{\beta_k}$ avec, pour tout k, $\beta_k\in[\![0,\alpha_k]\!]$. Ainsi,

$$a_n=|[\![0,\alpha_1]\!]\times\ldots\times[\![0,\alpha_p]\!]|=\prod_{k=1}^{p}(\alpha_k+1).$$

b. L'hypothèse se réécrit $p_k^{\varepsilon}\geq 2$ (car $x\mapsto x^{\varepsilon}$ croissante sur $\mathbb{R}_+$) et donc $p_k^{\varepsilon\alpha_k}\geq 2^{\alpha_k}$. Il suffit alors de vérifier par récurrence sur n que $2^n\geq n+1$. C'est vrai pour $n=0$ et si $2^n\geq n+1$, alors

$$2^{n+1}=2\times 2^n\geq 2(n+1)=n+2+n\geq n+2.$$

En conclusion, $p_k^{\varepsilon\alpha_k}\geq\alpha_k+1$.
c. Cette fonction f est dérivable comme produit et composition de fonctions dérivables. De plus, pour tout $x>0$, $f'(x)=(1-\varepsilon\ln(2)(x+1))2^{-\varepsilon x}$.
Ainsi, f est croissante jusqu'en $x_0=\frac{1-\varepsilon\ln 2}{\varepsilon\ln 2}$ et on notera y_0 sa valeur en ce point, puis décroissante.

d. Considérons $I_\varepsilon = \{k \le p,\ p_k \ge 2^{1/\varepsilon}\}$ et $J_\varepsilon = \{k \le p,\ p_k < 2^{1/\varepsilon}\}$.

$$\begin{aligned}\frac{a_n}{n^\varepsilon} &= \Big(\prod_{k\in I_\varepsilon} \frac{\alpha_k+1}{p_k^{\varepsilon\alpha_k}}\Big)\Big(\prod_{k\in J_\varepsilon} \frac{\alpha_k+1}{p_k^{\varepsilon\alpha_k}}\Big)\\ &\le \Big(\prod_{k\in J_\varepsilon} \frac{\alpha_k+1}{p_k^{\varepsilon\alpha_k}}\Big) \le \Big(\prod_{k\in J_\varepsilon} \frac{\alpha_k+1}{2^{\varepsilon\alpha_k}}\Big) \le y_0^{|J_\varepsilon|}\end{aligned}$$

en utilisant tour à tour la question **a** pour exprimer a_n, **b** pour gérer les termes d'indices dans I_ε et le maximum de f obtenu à la question **c**. Comme le nombre de nombres premiers inférieurs à $2^{1/\varepsilon}$ est majoré, on obtient un majorant C de la quantité $\frac{a_n}{n^\varepsilon}$.
e. Pour tout $x > 1$, il existe $\varepsilon > 0$ tel que $x \ge 1 + 2\varepsilon$ Alors, $0 \le \frac{a_n}{n^x} \le \frac{C}{n^{1+\varepsilon}}$. Par comparaison d'une série de terme général positif à une série de Riemann convergente, la série de terme général $\frac{a_n}{n^x}$ converge. Ainsi, $]1,+\infty[\subset \Lambda_c$ et donc, en passant à la borne inférieure, $\sigma_c \le 1$.
f. La série de terme général $\frac{a_n}{n}$ diverge car elle est supérieure à la série de terme général $\frac{2}{n}$ qui diverge. Par suite, pour tout $x < 1$, la série de terme général $\frac{a_n}{n^x}$ diverge (car $\frac{a_n}{n^x} \ge \frac{a_n}{n}$). En conclusion, $\Lambda_c =]1,+\infty[$ et donc, $\sigma_c = 1$.

Partie B – Abscisse σ_a de convergence absolue

1. Écrivons $\left|\frac{a_n}{n^s}\right| = \left|\frac{a_n}{n^{s_0}}\right| \times \left|\frac{1}{n^{s-s_0}}\right| = \left|\frac{a_n}{n^{s_0}}\right| \times \frac{1}{n^{\Re\mathfrak{e}(s-s_0)}}$. Comme $\Re\mathfrak{e}(s-s_0) > 0$, $\frac{1}{n^{\Re\mathfrak{e}(s-s_0)}} \to 0$ et donc, $\left|\frac{a_n}{n^s}\right| = o\left(\left|\frac{a_n}{n^{s_0}}\right|\right)$. Par comparaison avec le terme général d'une série absolument convergente, la série de terme général $\frac{a_n}{n^s}$ converge absolument.
2. Soit $s \in \mathbb{C}$. Par définition de σ_a comme borne inférieure (autrement dit $\Re\mathfrak{e}(s)$ n'est pas un minorant de Λ_a), il existe $x < \Re\mathfrak{e}(s)$ tel que la série de terme général $\frac{a_n}{n^x}$ converge absolument. D'après la question **1** avec $s_0 = x$, la série de terme général $\frac{a_n}{n^s}$ converge absolument.
3. Par définition de σ_a comme borne inférieure, $\Lambda_a = \varnothing$. Or, si la série de terme général $\frac{a_n}{n^s}$ converge absolument pour un $s \in \mathbb{C}$, il en est de même pour la série de terme général $\frac{a_n}{n^{\Re\mathfrak{e}(s)}}$ (car les termes généraux ont même module) ce qui entraîne $\Re\mathfrak{e}(s) \in \Lambda_a$: contradiction.
4. • si $\Re\mathfrak{e}(s) > \sigma_a$, $\Re\mathfrak{e}(s)$ n'est pas un minorant de Λ_a donc il existe $x \in \Lambda_a$ tel que $x < \Re\mathfrak{e}(s)$. On conclut à nouveau avec la question **1** avec $s_0 = x$: la série de terme général $\frac{a_n}{n^s}$ converge absolument ;
• raisonnons par l'absurde en supposant qu'il existe s tel que $\Re\mathfrak{e}(s) < \sigma_a$ et la série de terme général $\frac{a_n}{n^s}$ converge absolument. Alors, pour tout $x \in]\Re\mathfrak{e}(s), \sigma_a[$, la série de terme général $\frac{a_n}{n^x}$ converge absolument d'après la question **1** et donc, $x \in \Lambda_a$: cela contredit la définition de σ_a comme borne inférieure de Λ_a.

Partie C – Abscisse σ_c de convergence

1. a.

$$\begin{aligned}u_{n+1} - u_n &= \frac{1}{n^z}\Big(\Big(1+\frac{1}{n}\Big)^{-z} - 1\Big) = \frac{1}{n^z}\Big(-z\frac{1}{n} + o(\frac{1}{n})\Big)\\ &= -\frac{z}{n^{z+1}} + O(\frac{1}{n^{z+1}}).\end{aligned}$$

b. Les deux termes de droite sont les termes généraux de séries absolument convergentes car $\Re\mathfrak{e}(z+1) > 1$ donc la série de terme général $u_{n+1} - u_n$ converge absolument.
2. Notons S_n la somme partielle de la série de terme général $\frac{a_n}{n^{s_0}}$ et $u_n = \frac{1}{n^{s-s_0}}$. Pour tous $m \ge n$,

$$\begin{aligned}\sum_{k=n}^{m} \frac{a_n}{n^s} &= \sum_{k=n}^{m} (S_k - S_{k-1})u_k = \sum_{k=n}^{m} S_k u_k - \sum_{k=n-1}^{m-1} S_k u_{k+1}\\ &= S_m u_{m+1} - S_{n-1} u_n + \sum_{k=n}^{m} S_k(u_k - u_{k+1}).\end{aligned}$$

CORRIGÉS

Soit $\varepsilon > 0$, N tel que $|u_k| \leq \varepsilon$ pour tout $k \geq N$, N' tel que pour tout $m \geq n \geq N'$,

$$\sum_{k=n}^{m} |u_k - u_{k+1}| \leq \varepsilon$$

(qui existe d'après la question **1b**. Notons enfin A un majorant de la suite $(S_n)_n$ (qui converge par hypothèse). Alors, combinant ces inégalités, pour $m \geq n \geq \max(N, N')$,

$$\left| \sum_{k=n}^{m} \frac{a_n}{n^s} \right| \leq 3A\varepsilon.$$

La suite des sommes partielles de la série de terme général $\frac{a_n}{n^s}$ est de Cauchy donc converge.
3. On obtient directement en remplaçant la question **B1** par la question **C2** :
▹ si $\sigma_c = -\infty$, alors la série de terme général $\frac{a_n}{n^s}$ converge pour tout $s \in \mathbb{C}$.
▹ si $\sigma_c = +\infty$, alors la série de terme général $\frac{a_n}{n^s}$ ne converge pour aucun $s \in \mathbb{C}$.
▹ si $\sigma_c \in \mathbb{R}$ et $s \in \mathbb{C}$.
- si $\mathfrak{Re}(s) > \sigma_c$, alors la série de terme général $\frac{a_n}{n^s}$ converge ;
- si $\mathfrak{Re}(s) < \sigma_c$, alors la série de terme général $\frac{a_n}{n^s}$ ne converge pas.

Partie D – Inégalités

1. a. On remarque que $\Lambda_a \subset \Lambda_c$ car la convergence absolue entraîne la convergence. En passant à la borne inférieure, on obtient que $\sigma_c \leq \sigma_a$.
b. Cette inégalité est optimale puisqu'il s'agit d'une égalité dès que la suite $(a_n)_n$ est à termes réels positifs.
2. a. Raisonnons par l'absurde et supposons $\sigma_c + 1 < \sigma_a$. Soit $x \in]\sigma_c + 1, \sigma_a[$ et $\varepsilon > 0$ tel que $x - 1 - \varepsilon > \sigma_c$. La série de terme général $\frac{a_n}{n^{x-1-\varepsilon}}$ converge donc $\frac{a_n}{n^{x-1-\varepsilon}} \to 0$. Par conséquent

$$\left| \frac{a_n}{n^x} \right| = \left| \frac{a_n}{n^{x-1-\varepsilon}} \right| \times \frac{1}{n^{1+\varepsilon}} = o\left(\frac{1}{n^{1+\varepsilon}}\right).$$

La série de terme général $\frac{a_n}{n^x}$ est absolument convergente donc $x \in \Lambda_a$ et par conséquent $\sigma_a \leq x$: contradiction.
En conclusion, $\sigma_a \leq \sigma_c + 1$.
b. ▹ Pour cette suite, la convergence absolue correspond à une série de Riemann ; par conséquent, $\Lambda_a =]1, +\infty[$ et $\sigma_a = 1$.
▹ La série de terme général $\frac{(-1)^n}{n^x}$ diverge grossièrement pour $x \leq 0$ et converge pour $x > 0$ d'après le critère spécial des séries alternées. Ainsi, $\Lambda_c =]0, +\infty[$ et $\sigma_c = 0$.
Dans ce cas $\sigma_a = \sigma_c + 1$ donc l'inégalité est optimale.

Chapitre 23

Dénombrement

1. Ensembles finis — **2.** Exemples de dénombrements usuels

Objectifs et compétences du programme

- Calculer le cardinal d'un ensemble fini.
- Exploiter une hypothèse de cardinalité.
- Maîtriser les manipulations des coefficients binomiaux.

Johann Peter Gustav Lejeune Dirichlet, mathématicien allemand (1805-1859).
Fort connu pour son importante activité en théorie des nombres et en analyse harmonique, il développe le théorème de Dirichlet qui assure la convergence de la série de Fourier d'une fonction continue par morceaux. On lui doit à cette occasion la définition formelle de fonction. En arithmétique, il a laissé son nom à un théorème de progression arithmétique (qui assure que pour des entiers m et n premiers entre eux, il existe une infinité de nombres premiers de la forme $km+n$ avec k entier) et a ouvert un champ gigantesque désormais appelé théorie algébrique des nombres. C'est pour des considérations arithmétiques qu'il devient l'un des premiers mathématiciens à exploiter le principe des tiroirs.

COURS

Dans ce chapitre, on s'intéresse aux ensembles finis et au dénombrement de leurs éléments. Dans un premier temps, le nombre d'éléments, ou cardinal, d'un ensemble est défini rigoureusement puis sert à établir des bijections entre ensembles. Dans un second temps, on examine différents dénombrements classiques pour illustrer la variété des méthodes de calcul.

1. Ensembles finis

Les démonstrations de la première sous-section sont hors programme.

1.1. Définition, parties finies de $\mathbb{N}$

Définition 23.1.

▹ Un ensemble E est fini s'il est vide ou s'il existe un entier $n \in \mathbb{N} \setminus \{0\}$ et une bijection $\varphi : [\![1, n]\!] \to E$.
▹ Un ensemble qui n'est pas fini est infini.

Remarque
Cette définition correspond à une idée très intuitive : la bijection correspond à une façon de compter les éléments, la bijection φ de la définition est alors une « énumération » de ceux-ci.

Proposition 23.2.

Toute partie de l'ensemble $[\![1, n]\!]$ pour $n \in \mathbb{N} \setminus \{0\}$ est finie.

Démonstration
Soit A une partie de $[\![1, n]\!]$.
▹ Si A est l'ensemble vide, alors A est finie.
▹ Sinon, on ordonne les éléments de A en posant $A_1 = A$ et $x_1 = \min A_1$ puis, pour tout $k \in \mathbb{N} \setminus \{0\}$

$$A_{k+1} = \begin{cases} A_k & \text{si } A_k = \varnothing \\ A_k \setminus \{x_k\} & \text{sinon} \end{cases} \qquad x_{k+1} = \begin{cases} x_k & \text{si } A_{k+1} = \varnothing \\ \min A_{k+1} & \text{sinon.} \end{cases}$$

L'ensemble $\{x_k, k \in \mathbb{N} \setminus \{0\}\}$ est non vide majoré donc admet un plus grand élément. Comme la suite $(x_k)_{k \in \mathbb{N} \setminus \{0\}}$ est, par construction, strictement croissante ou stationnaire à partir d'un certain rang, elle stationne à cette plus grande valeur. Considérons $p = \min\{k \in \mathbb{N} \setminus \{0\},\ x_k = x_{k+1}\}$. L'application

$$\begin{cases} [\![1, p]\!] & \to & A \\ k & \mapsto & x_k \end{cases}$$

est une bijection par construction de la suite finie et strictement croissante $(x_k)_{k \in [1,p]}$. ■

Corollaire 23.3.

Toute partie d'un ensemble fini est finie.

Démonstration
Le résultat est évident pour l'ensemble vide. Soit F une partie non vide d'un ensemble fini E. Par définition, il existe $n \in \mathbb{N} \setminus \{0\}$ et une bijection $\varphi : E \to [\![1, n]\!]$. $\varphi(F)$ est une partie de $[\![1, n]\!]$ donc est finie : il existe $p \in \mathbb{N} \setminus \{0\}$ et une bijection $\psi : \varphi(F) \to [\![1, p]\!]$.
Par construction, $\psi \circ \varphi : F \to [\![1, p]\!]$ est une bijection et donc F est finie. ■

1.2. Cardinaux

Proposition 23.4.

Soit p, $q \in \mathbb{N} \setminus \{0\}$. Alors, il existe une injection de $[\![1, p]\!]$ dans $[\![1, q]\!]$ si, et seulement si, $p \leq q$.

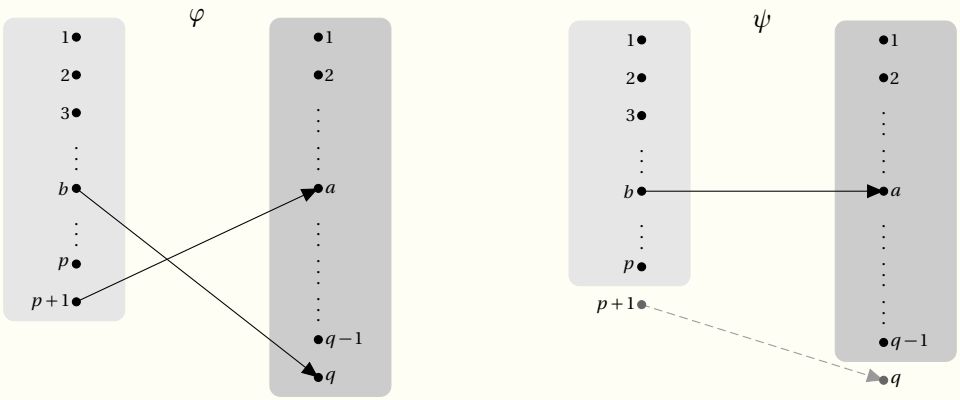
Démonstration
Le sens retour est immédiat en considérant l'application injective

$$\left\{ \begin{array}{ccc} [\![1, p]\!] & \to & [\![1, q]\!] \\ n & \mapsto & n \end{array} \right.$$

Montrons le sens direct par récurrence en posant, pour tout $p \in \mathbb{N} \setminus \{0\}$, $\mathscr{P}(p)$:« pour tout $q \in \mathbb{N} \setminus \{0\}$, s'il existe une injection de $[\![1, p]\!]$ dans $[\![1, q]\!]$, alors $p \leq q$ ».
▹ $\mathscr{P}(1)$ est vrai car pour tout $q \in \mathbb{N} \setminus \{0\}$, $q \geq 1$.
▹ Soit $p \in \mathbb{N} \setminus \{0\}$. Supposons $\mathscr{P}(p)$ vrai et montrons $\mathscr{P}(p+1)$.
Soit $q \in \mathbb{N} \setminus \{0\}$ et $\varphi : [\![1, p+1]\!] \to [\![1, q]\!]$ une injection.
∘ Si $\varphi(p+1) = q$ (ou si q n'admet pas d'antécédent), alors la restriction de φ entre $[\![1, p]\!]$ et $[\![1, q-1]\!]$ est une injection et par hypothèse de récurrence, $p \leq q-1$ et donc $p+1 \leq q$.
∘ Si $\varphi(p+1) = a < q$ et définissons $\psi : [\![1, p]\!] \to [\![1, q-1]\!]$ par $\psi(b) = a$ et

$$\forall n \in [\![1, p]\!] \setminus \{b\}, \quad \psi(n) = \varphi(n)$$

où b désigne l'unique antécédent de q par φ.



On vérifie aisément que ψ est injective et on conclut dans ce cas également avec l'hypothèse de récurrence. ■

Proposition 23.5.

Soit p, $q \in \mathbb{N} \setminus \{0\}$. Alors, il existe une surjection de $[\![1,p]\!]$ dans $[\![1,q]\!]$ si, et seulement si, $p \geq q$.

Démonstration

Le sens retour est immédiat en considérant l'application surjective

$$\left\{\begin{array}{ccccc} [\![1,p]\!] & \to & [\![1,q]\!] & & \\ n & \mapsto & n & \text{si } n \leq q, \\ & & q & \text{sinon.} \end{array}\right.$$

Pour le sens direct, ramenons-nous à la première propriété.

Supposons qu'il existe une surjection $\varphi : [\![1,p]\!] \to [\![1,q]\!]$ et posons pour tout $n \leq q$

$$\psi(n) = \min \varphi^{-1}(\{n\}).$$

Le plus petit élément est bien défini car l'ensemble considéré est non vide (surjectivité de φ). On vérifie alors que ψ est une application de $[\![1,q]\!]$ dans $[\![1,p]\!]$ qui est injective car $\varphi \circ \psi$ est l'identité de $[\![1,q]\!]$. On conclut avec le premier point de la proposition. ■

En combinant ces deux derniers résultats, on obtient la caractérisation suivante pour l'existence de bijections entre ensembles finis.

Corollaire 23.6.

Soit p, $q \in \mathbb{N} \setminus \{0\}$. Alors, il existe une bijection de $[\![1,p]\!]$ dans $[\![1,q]\!]$ si, et seulement si, $p = q$.

Théorème 23.7.

Pour tout ensemble fini non vide E, il existe un unique $n \in \mathbb{N} \setminus \{0\}$ tel qu'il existe une bijection entre E et $[\![1,n]\!]$.
Cet entier n est appelé cardinal de E et noté $|E|$, $\mathrm{card}(E)$ ou, plus rarement $\#E$.

Notation

Par convention, l'ensemble vide est de cardinal 0.

Démonstration

L'existence provient de la définition d'ensemble fini.

Considérons deux entiers n et n' tels qu'il existe des bijections $\varphi : E \to [\![1,n]\!]$, $\varphi' : E \to [\![1,n']\!]$. Alors, l'application $\varphi' \circ \varphi^{-1}$ réalise une bijection entre $[\![1,n]\!]$ et $[\![1,n']\!]$ d'où $n = n'$ et on obtient l'unicité. ■

Les propositions précédentes se traduisent alors pour des ensembles finis quelconques.

Proposition 23.8.

Soit E et F deux ensembles de cardinal p et q.
▹ Il existe une injection de E dans F si, et seulement si, $p \leq q$.
▹ Il existe une surjection de E dans F si, et seulement si, $p \geq q$.
▹ Il existe une bijection de E dans F si, et seulement si, $p = q$.

1.3. Opérations sur les ensembles finis

Proposition 23.9.

Soit E et F deux ensembles finis. Alors, $E \cup F$ est fini et $|E \cup F| = |E| + |F| - |E \cap F|$.

Démonstration

Procédons en deux temps.

▹ Si E et F sont disjoints, considérons des bijections $\varphi : E \to [\![1, p]\!]$ et $\psi : F \to [\![1, q]\!]$. On conclut en considérant la bijection

$$\left\{ \begin{array}{ccll} E \cup F & \to & [\![1, p+q]\!] & \\ x & \to & \varphi(x) & \text{si } x \in E \\ & & p + \psi(x) & \text{si } x \in F. \end{array} \right.$$

▹ Les ensembles $E \cap F$, $E \setminus (E \cap F)$ et $F \setminus (E \cap F)$ sont deux à deux disjoints et de réunion $E \cup F$. D'après le point précédent,

$$|E \cup F| = |E \cap F| + |E \setminus (E \cap F)| + |F \setminus (E \cap F)|.$$

En remarquant que $E \cap F$ et $E \setminus (E \cap F)$ sont disjoints de réunion E (et on procède de même pour F),

$$|E| = |E \cap F| + |E \setminus (E \cap F)|,$$
$$|F| = |E \cap F| + |F \setminus (E \cap F)|.$$

On obtient ainsi la formule cherchée en combinant les trois égalités. ■

Remarque

Cette formule est intuitive : lorsque l'on compte les éléments de $E \cup F$ en comptant ceux de E puis ceux de F, on a compté deux fois les éléments de $E \cap F$.

Exemple

Dénombrons les entiers à, au plus, quatre chiffres qui ne sont divisibles ni par 3 ni par 5. Notons, pour tout $p \in \mathbb{N} \setminus \{0\}$, E_p l'ensemble des entiers à, au plus, quatre chiffres qui sont divisibles par p. Avec cette notation, le nombre recherché est

$$10^4 - |E_3 \cup E_5| = 10^4 - |E_3| - |E_5| + |E_{15}|,$$

car $E_3 \cap E_5 = E_{15}$. Or,

$$\begin{aligned} |E_p| &= \left| \{ n \in [\![0, 10^4]\!],\ \exists k \in \mathbb{N},\ n = kp \} \right| \\ &= \left| \{ k \in \mathbb{N},\ 0 \leq k \leq \tfrac{10^4}{p} \} \right| \\ &= \left| [\![0, \lfloor \tfrac{10^4}{p} \rfloor]\!] \right| = \lfloor \tfrac{10^4}{p} \rfloor + 1. \end{aligned}$$

En reportant, on obtient 5 332 entiers à, au plus, quatre chiffres divisibles ni par 3 ni par 5.

La démonstration précédente entraîne en particulier le corollaire suivant.

Corollaire 23.10.

Soit E un ensemble fini et F une partie de E. Alors, le cardinal du complémentaire de F est $|E| - |F|$ donc $|F| \leq |E|$ avec égalité si, et seulement si, $F = E$.

Conseils méthodologiques

Pour prouver l'égalité de deux ensembles finis, il suffit désormais d'établir une inclusion et de prouver l'égalité des cardinaux.

Ce premier corollaire entraîne par une récurrence immédiate la version suivante.

Corollaire 23.11.

Soit $E_1, \dots, E_n$ des ensembles finis. L'ensemble $\cup_{i=1}^n E_i$ est fini et

$$\left|\bigcup_{i=1}^n E_i\right| \leq \sum_{i=1}^n |E_i|.$$

De plus, il y a égalité si les ensembles $E_1, \dots, E_n$ sont deux à deux disjoints.

En fait, on peut généraliser la formule de la proposition 23.9 pour n ensembles et obtenir la formule du crible de Poincaré.

Proposition 23.12. Formule du crible

Soit $(E_i)_{i\in I}$ une famille finie d'ensembles finis. Alors,

$$\left|\bigcup_{i\in I} E_i\right| = \sum_{k=1}^{|I|} (-1)^{k-1} \sum_{J\in\mathscr{P}_k(I)} \left|\bigcap_{i\in J} E_i\right|.$$

Démonstration

Montrons le résultat par récurrence sur $n = |I|$.

▹ Si $n = 1$, le résultat est immédiat puisque la formule se réécrit $|E_1| = |E_1|$. Remarquons que l'on a également déjà établi la formule dans le cas où $n = 2$.

▹ Soit $n \geq 1$ tel que la formule est établie pour toute famille de n ensembles finis et considérons $(E_i)_{i\leq n+1}$ une famille de $n+1$ ensembles finis (on a choisi sans perte de généralité $I = [\![1, n+1]\!]$). Alors,

$$\begin{aligned}
\left|\bigcup_{i\leq n+1} E_i\right| &= \left|E_{n+1} \cup \bigcup_{i\leq n} E_i\right| \\
&= |E_{n+1}| + \left|\bigcup_{i\leq n} E_i\right| - \left|\bigcup_{i\leq n} E_i \cap E_{n+1}\right| \\
&= |E_{n+1}| + \sum_{k=1}^n (-1)^{k-1} \sum_{J\in\mathscr{P}_k([\![1,n]\!])} \left|\bigcap_{i\in J} E_i\right| - \sum_{k=1}^n (-1)^{k-1} \sum_{J\in\mathscr{P}_k([\![1,n]\!])} \left|\bigcap_{i\in J} E_i \cap E_{n+1}\right| \\
&= |E_{n+1}| + \sum_{k=1}^n (-1)^{k-1} \sum_{J\in\mathscr{P}_k([\![1,n]\!])} \left|\bigcap_{i\in J} E_i\right| + \sum_{k=1}^n (-1)^{k+1-1} \sum_{J\in\mathscr{P}_k([\![1,n]\!])} \left|\bigcap_{i\in J\cup\{n+1\}} E_i\right|
\end{aligned}$$

ce qui est bien la formule recherchée. ■

Exemple

Voici la version de la formule du crible pour $n = 3$ ensembles finis E, F et G,

$$|E \cup F \cup G| = |E| + |F| + |G| - |E \cap F| - |E \cap G| - |F \cap G| + |E \cap F \cap G|.$$

La situation est beaucoup plus simple avec le produit cartésien (c'est-à-dire pour dénombrer des n-uplets ou n-listes)

Proposition 23.13.

Soit $E_1, \ldots, E_n$ des ensembles finis. $E_1 \times \ldots \times E_n$ est fini et

$$|E_1 \times \ldots \times E_n| = \prod_{i=1}^{n} |E_i|.$$

Démonstration

Il suffit de démontrer le résultat pour deux ensembles E et F de cardinal p et q ; considérons deux bijections $\varphi : E \to [\![1, p]\!]$ et $\psi : F \to [\![1, q]\!]$ puis définissons

$$\Phi : \left\{ \begin{array}{ccc} E \times F & \to & [\![1, pq]\!] \\ (k, l) & \mapsto & (\varphi(k) - 1)q + \psi(l). \end{array} \right.$$

On conclut en montrant que Φ est une bijection : pour tout entier $n \in [\![1, pq]\!]$, il existe un unique couple d'entiers $(a, b) \in [\![0, p-1]\!] \times [\![1, q]\!]$ tel que $n = aq + b$ (division euclidienne avec reste non nul) ; alors n admet comme unique antécédent par Φ le couple $(\varphi^{-1}(a+1), \psi^{-1}(b))$. ■

Remarque

Ce résultat est très intuitif : pour dénombrer les n-uplets, on raisonne composante par composante : il y a $|E_1|$ possibilités pour la première composante qui est un élément de E_1 ; ensuite, il y a $|E_2|$ possibilités pour la deuxième composante qui est un élément de E_2 et ainsi de suite... Au final, il y a bien $|E_1| \times \ldots \times |E_n|$ n-uplets de $E_1 \times \ldots \times E_n$.

Dénombrer E^n revient à dénombrer les choix avec remise de n éléments dans l'ensemble E.

1.4. Applications entre ensembles finis

Proposition 23.14.

L'image directe d'un ensemble fini par une injection est un ensemble fini de même cardinal.

Démonstration

Si $f : E \to F$ est une injection, alors l'application induite par f entre E et $f(E)$ est une bijection donc $|f(E)| = |E|$. ■

On en déduit le théorème suivant

Théorème 23.15.

Soit E et F deux ensembles finis de même cardinal et $f : E \to F$ une application. Il y a équivalence entre les propriétés suivantes

- f est injective ;
- f est surjective ;
- f est bijective.

Démonstration

Il suffit de montrer l'équivalence entre les deux premières propriétés.

▹ Si f est injective, alors $|f(E)| = |E| = |F|$ or $f(E) \subset F$ donc $f(E) = F$ et f est surjective.

▹ Si f est surjective, alors on définit une application $g : F \to E$ telle que $f \circ g = \mathrm{Id}_F$ et donc g est injective. D'après le premier point, g est aussi surjective donc bijective. Par conséquent, f est bijective en tant que bijection réciproque de g. ■

Conseils méthodologiques

Pour prouver qu'une application entre deux ensembles finis est bijective, il suffit désormais d'établir l'injectivité (ou de manière équivalente la surjectivité) et de prouver l'égalité des cardinaux.

Remarque

Cette propriété est fausse lorsque l'on supprime l'hypothèse de cardinal fini. Par exemple, l'application dérivation $D : P \mapsto P'$ est surjective de $\mathbb{R}[X]$ dans $\mathbb{R}[X]$ mais pas injective.

2. Exemples de dénombrements usuels

2.1. Principes de dénombrement

Conseils méthodologiques

La section précédente permet de dégager deux méthodes essentielles de dénombrement :
- mettre l'ensemble à dénombrer en bijection avec un ensemble connu ;
- considérer une partition de l'ensemble à dénombrer.

Détaillons et formalisons deux résultats classiques de dénombrement.

Proposition 23.16. Principe des bergers

Soit E un ensemble fini et $f : E \to F$ une application telle que

$$\exists p \in \mathbb{N} \setminus \{0\}, \quad \forall y \in F, \qquad |f^{-1}(\{y\})| = p.$$

Alors, F est un ensemble fini et $|E| = p|F|$.

Démonstration

L'ensemble E est la réunion disjointe des ensembles $f^{-1}(\{y\})$ lorsque y décrit F, d'où

$$|E| = \sum_{y \in F} |f^{-1}(\{y\})| = p|F|.$$

■

Remarque

Une version plus intuitive (et qui justifie le nom) serait : un berger compte les pattes de ses brebis, divise le résultat par 4 et obtient le nombre de brebis.

En procédant ainsi, il applique le résultat précédent avec $p = 4$ et f l'application qui à une patte associe la brebis correspondante.

Proposition 23.17. Principe des tiroirs de Dirichlet

Soit E et F deux ensembles finis de cardinal n et p respectivement, et $f : E \to F$. Supposons qu'il existe un entier $k \in \mathbb{N} \setminus \{0\}$ tel que $n > kp$.
Alors, il existe un élément $y \in F$ tel que $|f^{-1}(\{y\})| > k$.

Démonstration
Par l'absurde, si pour tout $y \in F$, $|f^{-1}(\{y\})| \leq k$. L'ensemble E est la réunion disjointe des ensembles $f^{-1}(\{y\})$ lorsque y décrit F, d'où

$$|E| = \sum_{y \in F} |f^{-1}(\{y\})| \leq kp,$$

ce qui contredit l'hypothèse de la proposition. ■

Remarque
Une version plus intuitive (et qui justifie le nom) serait (pour $k = 1$) : si on range n chaussettes dans p tiroirs et que $n > p$, alors il existe au moins un tiroir qui contient au moins deux chaussettes.

Pour utiliser ce résultat, on cherche souvent les « bons tiroirs », c'est-à-dire le découpage adapté à notre problème.

Exemple

Considérons 5 points du plan à coordonnées entières et montrons que parmi tous les segments ayant ces points pour extrémités, l'un au moins admet pour milieu un point à coordonnées entières.
Introduisons les ensembles A des points à abscisse paire et ordonnée paire, B des points à abscisse paire et ordonnée impaire, C des points à abscisse impaire et ordonnée paire, D des points à abscisse impaire et ordonnée impaire. Au moins l'un de ces quatre ensembles contient deux points et le milieu du segment des deux points concernés est à coordonnées entières.

2.2. Parties d'un ensemble fini (coefficients binomiaux)

Proposition 23.18.

Soit E un ensemble fini de cardinal n. L'ensemble $\mathscr{P}(E)$ est fini et de cardinal 2^n.

Démonstration
Il suffit de vérifier que $\mathscr{P}(E)$ est en bijection avec $\{0,1\}^n$: l'image d'une partie est la suite des valeurs de sa fonction indicatrice sur une énumération des éléments de E (ce qui caractérise bien la partie). ■

Définition 23.19.

Le coefficient binomial $\binom{n}{p}$ est le nombre de parties à p éléments (quelquefois appelées p-combinaisons) d'un ensemble de cardinal n.

Remarque
▹ Il existe une seule partie à n éléments (respectivement sans élément) dans un ensemble E de cardinal n : E (respectivement $\varnothing$). Ainsi, $\binom{n}{n} = \binom{n}{0} = 1$.

▹ Il n'y a pas de partie à $p > n$ éléments dans un ensemble E de cardinal n donc $\binom{n}{p} = 0$ pour tout $p > n$.

Remarque

Par le dénombrement on retrouve la formule du binôme de Newton

$$\forall a, b \in \mathbb{C}, \qquad (a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}.$$

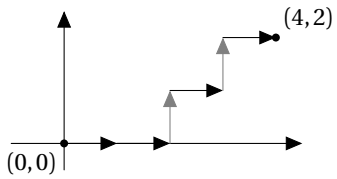
En effet, dans le développement du produit $(a+b)\times(a+b)\times\ldots\times(a+b)$, il y a autant de termes $a^k b^{n-k}$ qu'il y a de choix des k facteurs où l'on prend a (et donc des $n-k$ où l'on prend b), c'est-à-dire autant que d'ensembles à k éléments dans un ensemble à n éléments.

Exemple

L'identité $\sum_{p=0}^{n} \binom{n}{p} = 2^n$ (qui se montre facilement avec la formule du binôme ou avec un calcul par récurrence) peut se voir en considérant la partition de l'ensemble des parties d'un ensemble de cardinal n selon le cardinal des parties.

Exemple

Dénombrons les chemins joignant les points $(0,0)$ et $(m,n) \in \mathbb{N}^2$ n'utilisant que des déplacements de vecteurs $(1,0)$ (horizontaux) et $(0,1)$ (verticaux).



Un tel chemin est composé de $m+n$ déplacements et m parmi ceux-ci sont horizontaux : il y a donc autant de chemins que de choix de parties de cardinal m dans l'ensemble des $m+n$ déplacements, c'est-à-dire $\binom{m+n}{m}$.

Proposition 23.20.

Pour tous $p \leq n$,

$$\binom{n}{p} = \binom{n}{n-p}.$$

Démonstration

L'application $A \mapsto \overline{A}$ est une bijection de l'ensemble des parties de $[\![1,n]\!]$ à p éléments dans l'ensemble des parties de $[\![1,n]\!]$ à $n-p$ éléments. ■

Proposition 23.21. Triangle de Pascal

Pour tous $p \leq n$,

$$\binom{n+1}{p+1} = \binom{n}{p} + \binom{n}{p+1}.$$

Démonstration
Démontrons le résultat en considérant la partition de l'ensemble des $\binom{n+1}{p+1}$ parties de $[\![1, n+1]\!]$ de cardinal $p+1$ selon l'appartenance de l'élément $n+1$ à ces parties.

- Il y a $\binom{n}{p}$ parties de cardinal p dans $[\![1, n]\!]$ donc autant de parties à $p+1$ éléments dans $[\![1, n+1]\!]$ qui contiennent $n+1$.
- Il y a $\binom{n}{p+1}$ parties de cardinal $p+1$ dans $[\![1, n]\!]$ donc autant de parties à $p+1$ éléments dans $[\![1, n+1]\!]$ qui ne contiennent pas $n+1$.

■

Proposition 23.22. Formule de Vandermonde

Pour tous $a, b, n \in \mathbb{N}$,
$$\sum_{k=0}^{n} \binom{a}{k}\binom{b}{n-k} = \binom{a+b}{n}.$$
En particulier, pour $a = b = n$,
$$\sum_{k=0}^{n} \binom{n}{k}^2 = \binom{2n}{n}.$$

Démonstration
Soit A et B deux parties disjointes de cardinal a et b respectivement. Partitionnons l'ensemble des $\binom{a+b}{n}$ parties de $A \cup B$ de cardinal n selon le cardinal de leur intersection avec A. Il y a $\binom{a}{k}$ parties de A de cardinal k et $\binom{b}{n-k}$ parties de B de cardinal $n-k$ donc $\binom{a}{k}\binom{b}{n-k}$ parties de $A \cup B$ dont l'intersection avec A est de cardinal k. Ainsi
$$\sum_{k=0}^{n} \binom{a}{k}\binom{b}{n-k} = \binom{a+b}{n}.$$

■

2.3. Applications entre ensembles finis

Proposition 23.23.

Soit E et F deux ensembles finis de cardinal p et n. Alors, l'ensemble F^E des applications de E dans F est fini et $|F^E| = n^p$.

Démonstration
Il suffit de montrer le résultat pour $E = [\![1, p]\!]$. Procédons par récurrence sur $p \in \mathbb{N} \setminus \{0\}$.

▹ Pour $p = 1$, il y a n valeurs possibles pour l'image de l'unique élément de l'ensemble de départ : d'où n applications.

▹ Soit $p \in \mathbb{N} \setminus \{0\}$. Supposons le résultat établi au rang p et montrons-le au rang $p+1$. Une application de $F^{[1,p+1]}$ est uniquement déterminée par la valeur $f(p+1)$ et l'application $f_{|[1,p]} \in F^{[1,p]}$. En d'autres termes, $F^{[1,p+1]}$ est en bijection avec $F \times F^{[1,p]}$ donc
$$|F^{[1,p+1]}| = |F|.|F^{[1,p]}| = n^{p+1}.$$

■

On montre également par récurrence le résultat suivant.

Proposition 23.24. Nombre d'arrangements

Soit E et F deux ensembles finis de cardinal p et n. Alors, l'ensemble des injections de E dans F est fini et son cardinal est le nombre d'arrangements de p éléments parmi n défini par

$$\begin{cases} 0 & \text{si } p > n, \\ \frac{n!}{(n-p)!} & \text{sinon.} \end{cases}$$

Démonstration

Il suffit de montrer le résultat pour $E = [\![1,p]\!]$ et $F = [\![1,n]\!]$. Procédons par récurrence sur $p \in \mathbb{N} \setminus \{0\}$. Notons A_n^p le nombre d'injections de $[\![1,p]\!]$ dans $[\![1,n]\!]$.

▹ Pour $p = 1$, il y a n valeurs possibles pour l'image de l'unique élément de l'ensemble de départ : d'où n applications toutes injectives.

▹ Soit $p \in \mathbb{N} \setminus \{0\}$. Supposons le résultat établi au rang p et montrons-le au rang $p+1$. Une application injective de $[\![1,n]\!]^{[1,p+1]}$ (pour $n \geq p$) est uniquement déterminée par la valeur $f(p+1)$ et l'application injective $f_{|[1,p]} \in [\![1,n]\!] \setminus \{f(p+1)\}^{[1,p]}$. Par conséquent

$$A_n^{p+1} = nA_{n-1}^p = \frac{n!}{(n-1-p)!} = \frac{n!}{(n-(p+1))!}.$$

■

En utilisant la caractérisation des bijections entre deux ensembles de même cardinal, on en déduit

Corollaire 23.25.

Le nombre de bijections entre deux ensembles finis de cardinal n est $n!$.

En guise d'application, fournissons une preuve plus combinatoire de la définition des coefficients binomiaux avec les factorielles.

Proposition 23.26.

Soit $n \in \mathbb{N}$ et $p \in [\![0,n]\!]$. Alors,

$$\binom{n}{p} = \frac{n!}{p!(n-p)!}.$$

Démonstration

Partitionnons l'ensemble des $n!$ bijections σ de $[\![1,n]\!]$ dans $[\![1,n]\!]$ selon la valeur de $A = \sigma([\![1,p]\!])$.

- Il y a $\binom{n}{p}$ choix pour la partie A à p éléments.
- Une fois A fixée, une bijection est déterminée par une bijection de $[\![1,p]\!]$ dans A (il y en a $p!$) et une bijection de $[\![p+1,n]\!]$ dans $\overline{A}$ (il y en a $(n-p)!$).

Ainsi, $n! = \binom{n}{p} p!(n-p)!$. ■

Exemple

Dénombrons les applications croissantes de $[\![1,p]\!]$ dans $[\![1,n]\!]$ pour $p \leq n$.

▹ Commençons par dénombrer les applications strictement croissantes de $[\![1,p]\!]$ dans $[\![1,n]\!]$: une telle application est uniquement déterminée par les p images des entiers de $[\![1,p]\!]$; il y en a donc $\binom{n}{p}$.

▹ Ramenons-nous au cas précédent. À toute application croissante f de $[\![1,p]\!]$ dans $[\![1,n]\!]$, on associe l'application $\tilde{f} : [\![1,p]\!] \to [\![1,n+p-1]\!]$ définie par

$$\forall i \in [\![1,p]\!], \qquad \tilde{f}(i) = f(i) + i - 1.$$

On vérifie aisément que $\tilde{f}$ est strictement croissante puisque

$$\forall i \in [\![2,p]\!], \qquad \tilde{f}(i) - \tilde{f}(i-1) = f(i) - f(i-1) + 1 \geq 1.$$

Réciproquement, pour toute application $g : [\![1,p]\!] \to [\![1,n+p-1]\!]$ strictement croissante, il existe $f : [\![1,p]\!] \to [\![1,n]\!]$ croissante telle que $g = \tilde{f}$: l'application f vérifiant

$$\forall i \in [\![1,p]\!], \qquad f(i) = g(i) + 1 - i,$$

est bien définie (car $g(i) \geq i$ pour tout i) et croissante car

$$\forall i \in [\![2,p]\!], \qquad f(i) - f(i-1) = g(i) - g(i-1) - 1 \geq 0.$$

Ainsi, nous avons exhibé une bijection entre l'ensemble des applications croissantes de $[\![1,p]\!]$ dans $[\![1,n]\!]$ et les applications strictement croissantes de $[\![1,p]\!]$ dans $[\![1,n+p-1]\!]$: le cardinal recherché est

$$\binom{n+p-1}{p}.$$

FICHE

SYNTHÈSE

Dans ce chapitre on a défini le cardinal des ensembles finis comme le nombre d'éléments d'une partie. Ensuite pour le manipuler, on a montré les propriétés suivantes.

Opération sur les ensembles finis

Soit E et F des ensembles finis.
▹ $|E \cup F| = |E| + |F| - |E \cap F|$.
▹ $|E \times F| = |E|.|F|$.
▹ $|F^E| = |F|^{|E|}$.

L'intérêt des calculs de cardinal repose sur les simplifications suivantes dans la démonstration d'égalité entre ensembles ou de propriétés de bijection.

Utilisation du cardinal fini

▹ Une partie F d'un ensemble E fini est égale à E si, et seulement si, $|E| = |F|$.
▹ Soit E et F deux ensembles finis de même cardinal et $f : E \longrightarrow F$ une application. Il y a équivalence entre les propriétés suivantes
- f est injective ;
- f est surjective ;
- f est bijective.

En pratique, le calcul d'un cardinal repose sur deux méthodes essentielles :
- l'une consiste à mettre l'ensemble à dénombrer en bijection avec un ensemble connu ;
- l'autre à considérer une partition « simple » de l'ensemble à dénombrer.

Dans les deux cas, il faut connaître un certain nombre de résultat de dénombrements usuels.

Dénombrements usuels

▹ Soit E un ensemble de cardinal n. Il y a
- n^p p-uplets d'éléments de E ;
- $\binom{n}{p}$ parties de E de cardinal p ;
- 2^n parties de E.

▹ Soit E un ensemble de cardinal p et F un ensemble de cardinal n. Il y a
- n^p applications de E dans F,
- $\frac{n!}{(n-p)!}$ applications injectives de E dans F.

EXERCICES

Vrai ou faux ?

		Vrai	Faux
a)	Pour tous $a \leq b \in \mathbb{N}$, $\lvert [\![a,b]\!] \rvert = b-a$.	☐	☐
b)	Si f est une injection entre $[\![1,n]\!]$ et $[\![1790, n+1789]\!]$, alors f est surjective.	☐	☐
c)	Soit E et F deux ensembles finis non vides. Si toutes les surjections de E dans F et de F dans E sont des bijections, alors E et F ont le même cardinal.	☐	☐
d)	Il y a 50 nombres pairs dans l'intervalle $[0,100]$.	☐	☐
e)	Pour tout $x \in [1,+\infty[$, il y a $\lfloor x \rfloor$ entiers dans l'intervalle $[1,x]$.	☐	☐
f)	Soit $n, m \in \mathbb{N} \setminus \{0\}$. Alors, $\lvert \{k \in [\![1,n]\!],\ m \mid k\} \rvert = \left\lfloor \frac{n}{m} \right\rfloor$.	☐	☐
g)	Soit E et F finis. Alors, $\lvert E \times F \rvert = \lvert E \rvert . \lvert F \rvert$.	☐	☐
h)	Soit E et F finis. Alors, le nombre d'applications de E dans F est $\lvert E \rvert^{\lvert F \rvert}$.	☐	☐
i)	Soit $p \leq n$. Le nombre d'injections de $[\![1,p]\!]$ dans $[\![1,n]\!]$ est $\frac{n!}{(n-p)!}$.	☐	☐
j)	Il y a $(p+1)^{n+1}$ applications impaires de $[\![-n,n]\!]$ dans $[\![-p,p]\!]$.	☐	☐
k)	Il y a 2^{n-1} parties de cardinal pair dans un ensemble de cardinal n.	☐	☐
l)	Pour tous $k, n \in \mathbb{N}$, $$\binom{n}{k} + \binom{n+1}{k} = \binom{n+1}{k+1}.$$	☐	☐
m)	Pour tous $a, b, r \in \mathbb{N}$, $$\sum_{k=0}^{r} \binom{a}{k}\binom{b}{r-k} = \binom{a+b}{r}.$$	☐	☐

Exercices d'application du chapitre 23

●○○ **Exercice 1**

Dénombrer les triplets (i,j,k) tels que $i+j+k=151$ et $i,j,k \in [\![1,75]\!]$.

●○○ **Exercice 2**

Calculer le cardinal de la différence symétrique $E \Delta F = (\overline{E} \cap F) \cup (E \cap \overline{F})$ de deux ensembles finis E et F.

●○○ **Exercice 3**

Soit $n \in \mathbb{N} \setminus \{0\}$.

1. Montrer par un argument de dénombrement que, pour tous $1 \leq p \leq n$,
$$p\binom{n}{p} = n\binom{n-1}{p-1}.$$
2. Montrer par un argument de dénombrement que, pour tous $a \leq b \leq n$,
$$\binom{n}{b}\binom{b}{a} = \binom{n}{a}\binom{n-a}{b-a}.$$

●○○ **Exercice 4**

1. Déterminer le nombre d'applications $f : [\![1,12]\!] \to [\![1,12]\!]$ réalisant simultanément les trois conditions suivantes
   - l'image d'un entier pair est paire ;
   - l'image d'un entier divisible par 3 est divisible par 3 ;
   - f est bijective.
2. Recommencer en supprimant la dernière condition.

●○○ **Exercice 5**

Soit $n \in \mathbb{N} \setminus \{0\}$ et $p \in [\![1,n]\!]$, E un ensemble de cardinal n et X une partie de E de cardinal p. Dénombrer les parties Y de E telles que $|X \cap Y| = 1$.

●○○ **Exercice 6**

Soit $n \in \mathbb{N} \setminus \{0\}$ et E un ensemble de cardinal n. Dénombrer les couples $(X,Y) \in \mathscr{P}(E)^2$ tels que $X \cap Y = \varnothing$.

●○○ **Exercice 7**

Soit $n \in \mathbb{N} \setminus \{0\}$ et E un ensemble de cardinal n. Dénombrer les couples $(X,Y) \in \mathscr{P}(E)^2$ tels que $X \subset Y$.

●○○ **Exercice 8**

Soit $n \in \mathbb{N} \setminus \{0\}$. Montrer que
$$\sum_{\substack{r,s,t \in [\![0,n]\!] \\ r+s+t=n}} \binom{n}{r}\binom{n}{s}\binom{n}{t} = \binom{3n}{n}.$$

●●● **Exercice 9**

Soit $n \in \mathbb{N} \setminus \{0\}$.

1. Déterminer le nombre de solutions dans $(\mathbb{N} \setminus \{0\})^k$ de l'équation $x_1 + \ldots + x_k = n$. On pourra introduire l'application qui a une solution $(x_1, \ldots, x_k)$ associe l'ensemble de $[\![1, n-1]\!]$ défini par
$$\{x_1, x_1 + x_2, \ldots, x_1 + \ldots + x_{k-1}\}.$$
2. En déduire le nombre de solutions dans $(\mathbb{N} \setminus \{0\})^k$ de l'équation $x_1 + \ldots + x_k \leq n$.

●●○ **Exercice 10**

Soit $n \in \mathbb{N} \setminus \{0\}$. Un chemin de Delannoy aboutissant en (m,n) est une suite de points $M_0, \ldots, M_N$ tels que $M_0(0,0)$, $M_N(m,n)$ et, pour tout $k \in [\![1,N]\!]$, le vecteur de M_{k-1} à M_k est soit $e_1 : (1,0)$, soit $e_2 : (0,1)$, soit $e_3 : (1,1)$. Montrer que le nombre de chemins de Delannoy aboutissant en (m,n) est
$$\sum_{k=0}^{\min(m,n)} \binom{m+n-k}{k}\binom{m+n-2k}{n-k}.$$

●●○ **Exercice 11**

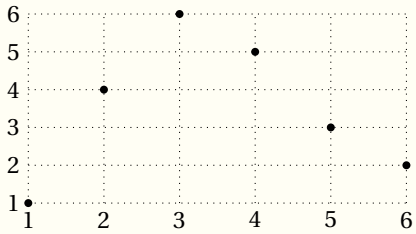
Soit $n, k \in \mathbb{N} \setminus \{0\}$. Le nombre de Stirling de deuxième espèce $S_{n,k}$ est le nombre de partitions de $[\![1, n]\!]$ en k parties.

1. Calculer $S_{n,1}$, $S_{n,2}$ et $S_{n,n-1}$.
2. Montrer que pour tout $n \in \mathbb{N} \setminus \{0\}$ et tout $k \in \mathbb{N} \setminus \{0\}$, $S_{n,k} = S_{n-1,k-1} + kS_{n-1,k}$.

●○○ **Exercice 12**

Soit $n \in \mathbb{N} \setminus \{0\}$. Dénombrer les permutations $\sigma \in S_{2n}$ vérifiant à la fois que $\sigma(1) < \sigma(2) < \ldots < \sigma(n)$ et que $\sigma(2n) < \sigma(2n-1) < \ldots < \sigma(n)$.

Voici la représentation d'une telle permutation :



●●○ **Exercice 13**

Soit $n \in \mathbb{N} \setminus \{0\}$.

1. Déterminer le nombre de permutations circulaires de S_n.
2. Déterminer le nombre de cycles de longueur $k \leq n$ de S_n.
3. Déterminer le nombre de permutations $\sigma \in S_n$ telles que $\sigma^2 = \mathrm{Id}$.

●○○ **Exercice 14**

Dénombrer les permutations de S_{20} dont la décomposition en cycle de supports disjoints contient trois 4-cycles, deux 3-cycles et deux points fixes.

●●○ **Exercice 15**

Pour tout $n \in \mathbb{N} \setminus \{0\}$, notons c_n le nombre de permutations $\sigma \in S_n$ connectées, c'est-à-dire telles que

$$\forall k \in [\![1, n-1]\!], \qquad \sigma([\![1, k]\!]) \neq [\![1, k]\!].$$

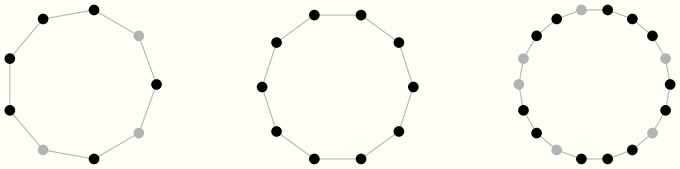
Montrer que, pour tout $n \in \mathbb{N} \setminus \{0\}$,

$$n! = \sum_{k=1}^{n} c_k (n-k)!.$$

●●● **Exercice 16**

Soit un polygone régulier à n côtés. Déterminer le cardinal du plus grand ensemble de sommets tels trois d'entre eux ne forment pas un triangle équilatéral.

Voici des illustrations pour $n = 9$, 10 et 18.



EXERCICES

Exercices d'approfondissement du chapitre 23

●○○ **Exercice A**

Montrer que, pour tout $n \geq 2$,

$$\sum_{X \in \mathscr{P}([1,n])} \sum_{k \in X} k = n(n+1)2^{n-2}.$$

●●○ **Exercice B**

Soit $f : \mathbb{N} \to \mathbb{N}$ telle que, pour $n \in \mathbb{N}$, $f(n)$ est le nombre de fois que le chiffre 1 apparaît dans la liste des entiers compris entre 0 et n. Par exemple, $f(12) = 5$ car il y a 5 chiffres 1 dans l'énumération 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12.

1. Déterminer $f\left(10^k - 1\right)$ pour $k \in \mathbb{N}$.
2. Montrer qu'il existe un nombre fini de $n \in \mathbb{N} \setminus \{0\}$ tels que $n \geq f(n)$.

●○○ **Exercice C**

Soit A une partie de $[\![1, 100]\!]$ de cardinal 55. Montrer qu'il existe $n \in A$ tel que $n + 9 \in A$.

●●○ **Exercice D**

Soit $n,\ k \in \mathbb{N} \setminus \{0\}$. Soit $u_{n,k}$ le nombre de parties de $[\![1, n]\!]$ de cardinal k qui ne contiennent pas deux entiers consécutifs. Montrer que $u_{n,k} = \binom{n-k+1}{k}$.

On pourra commencer par établir une relation de récurrence.

●●○ **Exercice E**

Soit E un ensemble de cardinal n muni d'une relation d'équivalence $\mathscr{R}$, k le nombre de classes d'équivalence et m le cardinal de l'ensemble des couples $(x, y) \in E^2$ tels que $x \mathscr{R} y$.

Montrer l'inégalité $n^2 \leq km$.

●●● **Exercice F**

Soit $A_1, \ldots, A_m$ des parties d'un ensemble S de cardinal $n \geq 2$. On suppose que, pour tous $x, y \in S$ avec $x \neq y$, il existe $i \in \{1, \ldots, m\}$ tel que $x \in A_i$ et $y \notin A_i$, ou $y \in A_i$ et $x \notin A_i$. Montrer que $m \geq \log_2(n)$.

Problèmes du chapitre 23

●●○ **Problème 1**

Partie A – Dénombrement des surjections

Notons, pour tous $n,\ p \geq 1$, $S_{n,p}$ le nombre de surjections de $[\![1, n]\!]$ dans $[\![1, p]\!]$.

1. a. Calculer $S_{n,p}$ pour $p > n$, puis pour les cas particuliers $p \in \{1, 2, n\}$.
   b. Calculer $S_{p+1,p}$.
2. a. Montrer que $\binom{p}{k}\binom{k}{q} = \binom{p}{q}\binom{p-q}{k-q}$ pour $q \leq k \leq p$.
   b. En déduire que, si $0 \leq q < p$, alors $\sum_{k=q}^{p} (-1)^k \binom{p}{k}\binom{k}{q} = 0$.
3. Supposons désormais $p \leq n$.
   a. Démontrer que $p^n = \sum_{q=1}^{p} \binom{p}{q} S_{n,q}$.
   b. En déduire que $S_{n,p} = (-1)^p \sum_{k=1}^{p} (-1)^k \binom{p}{k} k^n$.

c. Soit $p \geq 2$. Déduire de la question précédente que $S_{n,p} = p(S_{n-1,p} + S_{n-1,p-1})$.

d. Montrer que $S_{p+2,p} = \frac{p(3p+1)}{24}(p+2)!$.

Partie B – Dénombrement des dérangements

Un dérangement de $[\![1,n]\!]$ est une bijection de $[\![1,n]\!]$ dans $[\![1,n]\!]$ sans point fixe. Notons d_n le nombre de dérangements de $[\![1,n]\!]$. Par convention, $d_0 = 1$.

1. Démontrer que $n! = \sum_{k=0}^{n} \binom{n}{k} d_k$.
2. En déduire que $d_n = (-1)^n \sum_{k=0}^{n} (-1)^k \binom{n}{k} k!$.

●○○ **Problème 2**

Définition 23.27.

Soit E un ensemble de cardinal $n \geq 1$.
▹ Une chaîne de E est un ensemble de parties de E totalement ordonné pour l'inclusion.
▹ Une antichaîne de E est un ensemble de parties de E qui ne sont pas comparables deux à deux pour l'inclusion.
▹ Une chaîne (respectivement une antichaîne) est dite maximale si elle n'est pas strictement incluse dans une chaîne (respectivement une antichaîne).

Partie A – Inégalité de Bollobas-Lubell-Yamamoto-Meshalkin

1. Donner toutes les chaînes maximales de l'ensemble $\{a, b, c\}$.
2. a. Montrer que le nombre de chaînes maximales de E est $n!$.
   b. En déduire le nombre de chaînes maximales contenant une partie $A \subset E$ de cardinal k.
3. Soit A et B deux parties d'une antichaîne de E. Montrer que A et B n'appartiennent pas à une même chaîne de E.
4. Considérons une antichaîne de E et notons a_k le nombre de parties de cardinal k de cette antichaîne. Montrer que
$$\sum_{k=0}^{n} a_k k!(n-k)! \leq n!.$$

Partie B – Théorème de Sperner

L'objectif de cette partie consiste à déterminer C_n, le plus grand cardinal d'une antichaîne maximale de E (toujours de cardinal n).

1. Montrer que $C_n \geq \binom{n}{\lfloor \frac{n}{2} \rfloor}$.
2. a. Justifier que, pour tout $k \in [\![1,n]\!]$, $\binom{n}{k} \leq \binom{n}{\lfloor \frac{n}{2} \rfloor}$.
   b. En utilisant la partie A, montrer que $C_n \leq \binom{n}{\lfloor \frac{n}{2} \rfloor}$.
3. Conclure.

●●● **Problème 3 Systèmes de Steiner**

Définition 23.28.

Un système de Steiner de paramètres $(p,q,n) \in \mathbb{N}^3$ est une partie $S(p,q,n) \subset \mathscr{P}(E)$ qui vérifie les propriétés suivantes :

(i) E est un ensemble de cardinal n.

(ii) Tout élément de $S(p,q,n)$ est une partie de E de cardinal p.

(iii) Toute partie de E de cardinal q est incluse dans une unique partie de $S(p,q,n)$.

PROBLÈMES

Partie A – Exemples

1. Montrer qu'il n'existe pas de système de Steiner pour $p < q < n$.
2. Déterminer un système de Steiner $S(q, q, n)$ pour chaque $q \leq n$ puis calculer $|S(q, q, n)|$.
3. Déterminer une condition nécessaire et suffisante sur p et n pour l'existence d'un système de Steiner $S(p, 1, n)$. Préciser, lorsque cette condition est vérifiée, le cardinal d'un tel système de Steiner.

Partie B – Cardinal d'un système de Steiner

On suppose qu'il existe un système de Steiner $S(p, q, n)$ d'un ensemble E.

1. Soit $x \in E$ fixé.
   a. Déterminer le nombre de parties de E de cardinal q contenant x.
   b. Déterminer le nombre de parties de E de cardinal q contenant x et incluses dans une partie de cardinal p fixée.
   c. En déduire le nombre $r(p, q, n)$ de parties du système de Steiner $S(p, q, n)$ contenant x.
2. Montrer que $|S(p, q, n)| = \frac{n}{p} r(p, q, n)$.

Partie C – Demoiselles de Kirkman

Quinze demoiselles se promènent chaque jour en cinq rangs de 3. Déterminer le nombre maximal de jours de promenade tel que toute paire de demoiselles ne se retrouve jamais deux fois dans le même rang.
Indication : on pourra commencer par calculer $|S(3, 2, 15)|$.

CORRIGÉS

Corrigés des Vrai/Faux

a) Faux, la cardinal est $b-a+1$.

b) Vrai car les ensembles ont le même cardinal.

c) Vrai.

d) Faux, il y en a 51 car ce sont les $2p$ avec $p \in [\![0,50]\!]$.

e) Vrai.

f) Vrai.

g) Vrai.

h) Faux, le cardinal est $|F|^{|E|}$.

i) Vrai.

j) Faux, la cardinal est $(2p+1)^n$: une telle application envoie 0 par 0 et est déterminée par l'image des entiers $1, \ldots, n$ dans $[\![-p,p]\!]$.

k) Vrai.

l) Faux.

m) Vrai, c'est la formule de Vandermonde.

Corrigés des exercices d'application du chapitre 23

Exercice 1

Il suffit de dénombrer les couples (i,j) tels que $i \in [\![1,75]\!]$ et $j \in [\![76-i,75]\!]$; le nombre recherché est alors

$$\sum_{i=1}^{75} |[\![76-i,75]\!]| = \sum_{i=1}^{75} i = \frac{75.76}{2} = 2850.$$

Exercice 2

Comme les parties $E\Delta F$ et $E \cap F$ sont disjointes de réunion $E \cup F$, $|E\Delta F| + |E \cap F| = |E \cup F|$. Donc, avec la formule pour la réunion de deux ensembles finis, $|E\Delta F| = |E| + |F| - 2|E \cap F|$.

Exercice 3

1. Dénombrons de deux manières différentes l'ensemble des couples (k,A) où $k \in [\![1,n]\!]$, A une partie de $[\![1,n]\!]$ de cardinal p et $k \in A$.

- Il y a $\binom{n}{p}$ possibilités pour la partie A puis, une fois A fixée, p possibilités pour l'élément $k \in A$.
- Il y a n possibilités pour k puis, une fois k fixé, $\binom{n-1}{p-1}$ possibilités pour la partie $A \setminus \{k\}$ (de cardinal $p-1$ incluse dans l'ensemble $[\![1, n]\!] \setminus \{k\}$).

2. Dénombrons de deux façons distinctes les couples $(A, B) \in \mathscr{P}([\![1, n]\!])^2$ tels que $|A| = a$, $|B| = b$ et $A \subset B$.
- Il y a $\binom{n}{b}$ parties B de cardinal b puis une fois B fixée, $\binom{b}{a}$ parties A de B de cardinal a : il y a donc $\binom{n}{b}\binom{b}{a}$ tels couples.
- Il y a $\binom{n}{a}$ parties A de cardinal a puis une fois A fixée, $\binom{n-a}{b-a}$ parties $B \setminus A$ de cardinal $b-a$ ne rencontrant pas A : il y a donc $\binom{n}{a}\binom{n-a}{b-a}$ tels couples.

Exercice 4

1. Avec les conditions de l'énoncé, une telle application est déterminée par des bijections des ensembles $\{3, 9\}$, $\{6, 12\}$, $\{2, 4, 8, 10\}$ et $\{1, 5, 7, 11\}$ dans eux-mêmes. Il y a donc $2!2!4!4! = 2304$ telles applications.
2. Avec les conditions de l'énoncé, une telle application est déterminée par des applications des ensembles $\{3, 9\}$, $\{6, 12\}$, $\{2, 4, 8, 10\}$ et $\{1, 5, 7, 11\}$ dans eux-mêmes. Il y a donc $2^2 2^2 4^4 4^4 = 1048576$ telles applications.

Exercice 5

Une telle partie Y est uniquement déterminée par
- l'unique élément de X qu'elle contient : p choix ;
- l'ensemble de ses autres éléments qui est une partie du complémentaire de X : 2^{n-p} choix.

Le nombre de parties satisfaisant la condition est donc $p2^{n-p}$.

Exercice 6

Partitionnons cet ensemble de couples (X, Y) selon le cardinal k de X :
- il y a $\binom{n}{k}$ choix pour la partie X ;
- une fois X fixée, il y a 2^{n-k} choix pour la partie Y ne rencontrant pas X (nombre de parties du complémentaire de X qui est de cardinal $n-k$).

Le cardinal recherché est donc
$$\sum_{k=0}^{n} \binom{n}{k} 2^{n-k} = 3^n.$$

Remarque

Le résultat obtenu est relativement naturel : en effet, choisir de telles parties X et Y revient à décider, pour chaque élément de E s'il appartient seulement à X, seulement à Y ou à aucun des deux : il y a donc 3 choix pour chacun des n éléments d'où le dénombrement en 3^n.

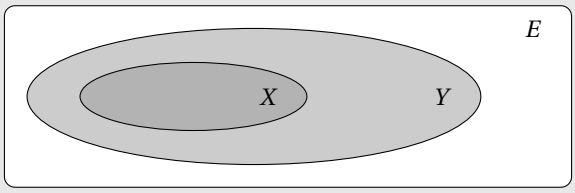
Exercice 7

Partitionnons cet ensemble de couples (X, Y) selon le cardinal k de X :
- il y a $\binom{n}{k}$ choix pour la partie X ;
- une fois X fixée, il y a 2^{n-k} choix pour la partie Y contenant X puisqu'il y a autant de choix pour $Y \setminus X$ dans le complémentaire de X qui est de cardinal $n-k$.

Le cardinal recherché est
$$\sum_{k=0}^{n} \binom{n}{k} 2^{n-k} = 3^n.$$

Remarque

Le résultat obtenu est relativement naturel : en effet, choisir de telles parties X et Y revient à décider, pour chaque élément de E s'il appartient à X et Y, seulement à Y ou à aucun des deux : il y a donc 3 choix pour chacun des n éléments d'où le dénombrement en 3^n.



On peut, de même, établir que l'ensemble des triplets $(X,Y,Z) \in \mathscr{P}(E)^2$ tels que $X \subset Y$ et $Y \subset Z$ est de cardinal 4^n.

Exercice 8

Soit A, B, C trois parties de cardinal n formant une partition de $[\![1,3n]\!]$. Le résultat est le dénombrement des parties de cardinal n de $[\![1,3n]\!]$ selon le nombre d'éléments de A, de B et de C.

Exercice 9

1. Notons $\mathscr{S}_n$ l'ensemble des solutions de l'équation. L'application

$$\varphi : \left\{ \begin{array}{ccc} \mathscr{S}_n & \to & \mathscr{P}_{k-1}([\![1,n-1]\!]) \\ (x_1,\ldots,x_k) & \mapsto & \{x_1, x_1+x_2, \ldots, x_1+\ldots+x_{k-1}\} \end{array} \right.$$

est une bijection. En effet, pour tous $y_1 < y_2 < \ldots < y_{k-1}$, $\varphi(x_1,\ldots,x_k) = \{y_1, y_2, \ldots, y_{k-1}\}$ si, et seulement si, $x_1 = y_1$, $x_i = y_i - y_{i-1}$ pour tout $i \in [\![2,k-1]\!]$ et $x_k = n - y_k$. Une partie à $k-1$ éléments admet donc un unique antécédent par φ.
En conclusion, les deux ensembles sont de même cardinal donc $|\mathscr{S}_n| = |\mathscr{P}_{k-1}([\![1,n-1]\!])| = \binom{n-1}{k-1}$.
2. Les solutions de l'équation sont dans l'un des ensembles disjoints $\mathscr{S}_k, \mathscr{S}_{k+1}, \ldots, \mathscr{S}_n$. Le cardinal recherché est

$$\sum_{i=k}^{n} \binom{i-1}{k-1} = \binom{n}{k}.$$

Exercice 10

Partitionnons l'ensemble des chemins de Delannoy selon le nombre de vecteurs e_3 apparaissant dans le chemin. Il y a entre 0 et $\min(m,n)$ tels vecteurs (car l'abscisse et l'ordonnée du point final est supérieur à ce nombre). Si un chemin de Delannoy aboutissant en (m,n) contient k vecteurs e_3, alors il contient également $m-k$ vecteurs e_1 (dénombrement des déplacements horizontaux) et $n-k$ vecteurs e_2 (dénombrement des déplacements verticaux), soit au total $n+m-k$ vecteurs. Il reste à choisir parmi ce $(n+m-k)$-uplet de vecteurs la position des k vecteurs e_3 puis des $n-k$ vecteurs e_2 parmi les places restantes. En conclusion,

$$\sum_{k=0}^{\min(m,n)} \binom{m+n-k}{k} \binom{m+n-2k}{n-k}.$$

Exercice 11

1. $S_{n,1}=1$ (une seule permutation en une partie), $S_{n,2}=\frac{1}{2}(2^n-2)=2^{n-1}-1$ (n'importe quelle partie non vide et différente de $[\![1,n]\!]$ définit avec son complémentaire une partition à deux parties) et $S_{n,n-1}=\binom{n}{2}$ (une partition à $n-1$ parties est composée de $n-2$ singletons et d'une paire ; il y a $\binom{n}{2}$ choix pour cette paire).
2. Partitionnons l'ensemble des partitions de $[\![1,n]\!]$ en k parties selon que n y apparaît dans un singleton ou non.
• Il y a $S_{n-1,k-1}$ partitions de $[\![1,n-1]\!]$ en $k-1$ parties donc $S_{n-1,k-1}$ partitions de $[\![1,n]\!]$ en k parties dont le singleton $\{n\}$.
• Il y a $S_{n-1,k}$ partitions de $[\![1,n-1]\!]$ en k parties donc $kS_{n-1,k}$ partitions de $[\![1,n]\!]$ en k parties en ajoutant n à l'une des k parties formant une partition de $[\![1,n-1]\!]$.
Par conséquent, $S_{n,k}=S_{n-1,k-1}+kS_{n-1,k}$.

Exercice 12

Remarquons tout d'abord que pour une telle permutation, $\sigma(n)$ est l'élément maximal donc $2n$. Choisissons l'ensemble des images des $n-1$ premiers éléments : il y a $\binom{2n-1}{n-1}$ possibilités pour cette partie. Comme les images sont ordonnées (σ est croissante sur $[\![1,n-1]\!]$), on a ainsi fixé les images des n premiers éléments. Mais on a aussi l'ensemble des images des entiers de $n+1$ à $2n$ (comme complémentaire des éléments déjà choisis) et, encore avec la monotonie, l'image de chaque entier est fixée. Par conséquent, une telle permutation est uniquement déterminée par l'image des $n-1$ premiers entiers dans $[\![1,2n-1]\!]$: le cardinal recherché est $\binom{2n-1}{n-1}$.

Exercice 13

1. Une permutation circulaire s'écrit $(n\ x_1\ \dots\ x_{n-1})$ avec $k\mapsto x_k$ une permutation de l'ensemble $[\![1,n-1]\!]$: il y a donc $|S_{n-1}|=(n-1)!$ permutations circulaires dans S_n.
2. Un cycle de longueur k est une permutation circulaire des k éléments de son support : il y a $\binom{n}{k}$ choix pour le support puis $(k-1)!$ permutations circulaires de celui-ci. En conclusion, il y a $(k-1)!\binom{n}{k}$ cycles de longueur k dans S_n.
3. Une permutation σ vérifie cette propriété si, et seulement si, sa décomposition en cycles à supports disjoints ne contient que des transpositions et en nombre inférieur à $\lfloor\frac{n}{2}\rfloor$. Le nombre de k-uplets de transpositions à supports disjoints de S_n est en choisissant les éléments de chaque transposition tour à tour

$$\binom{n}{2}\binom{n-2}{2}\binom{n-4}{2}\cdots\binom{n-2(k-1)}{2}=\frac{n!}{(n-2k)!2^k}.$$

Ainsi, en supprimant l'ordre, il y a $\frac{n!}{k!(n-2k)!2^k}$ permutations dont la décomposition en cycles à supports disjoints est composée de k transpositions. Le résultat recherché est donc

$$\sum_{k=0}^{\lfloor\frac{n}{2}\rfloor}\frac{n!}{k!(n-2k)!2^k}.$$

Exercice 14

Choisissons les trois cycles de longueur 4 à supports disjoints sans ordre soit $\frac{1}{3!}3!\binom{20}{4}3!\binom{16}{4}3!\binom{12}{4}$ possibilités.
Une fois ces cycles choisis, choisissons les deux cycles de longueur 3 à supports disjoints (et disjoints des supports des cycles précédents), soit $\frac{1}{2}2!\binom{8}{3}2!\binom{5}{3}$ possibilités.
Les deux derniers éléments étant fixes, leurs images sont déterminées à ce stade.
Le résultat recherché est donc $3!^2\binom{20}{4}\binom{16}{4}\binom{12}{4}2\binom{8}{3}\binom{5}{3}$.

Exercice 15

Partitionnons S_n selon l'indice $k_\sigma = \min\{k \in [\![1,n]\!],\ \sigma([\![1,k]\!]) = [\![1,k]\!]\}$.

Une permutation σ vérifie $k_\sigma = k$ si, et seulement si, elle se décompose en une permutation connectée de $[\![1,k]\!]$ et une permutation (quelconque) de $[\![k+1,n]\!]$: il y a donc $c_k(n-k)!$ telles permutations ; le résultat annoncé suit.

Exercice 16

▹ Si $n \neq 0[3]$, alors il n'y a jamais de triangle équilatéral à sommets communs avec le polygone régulier. Le nombre recherché est alors n.

▹ Si $n = 3p$ avec $p \in \mathbb{N} \setminus \{0\}$, alors on partitionne les sommets du polygone en p ensembles de trois sommets formant chacun les sommets d'un triangle équilatéral. La partie recherchée comporte au plus 2 sommets dans chaque ensemble : le nombre recherché est donc $2p = \frac{2}{3}n$.

Corrigés des exercices d'approfondissement du chapitre 23

Exercice A

Calculons cette somme en échangeant les deux sommes (finies) :

$$\begin{aligned}\sum_{X \in \mathscr{P}([1,n])} \sum_{k \in X} k &= \sum_{k=1}^{n} \sum_{\substack{X \in \mathscr{P}([\![1,n]\!]) \\ k \in X}} k \\ &= \sum_{k=1}^{n} k \sum_{\substack{X \in \mathscr{P}([\![1,n]\!]) \\ k \in X}} 1 \\ &= \sum_{k=1}^{n} k |\{X \in \mathscr{P}([\![1,n]\!]),\ k \in X\}|.\end{aligned}$$

Comme il y a autant de parties de $[\![1,n]\!]$ contenant k que de parties de $[\![1,n]\!] \setminus \{k\}$, on obtient

$$\sum_{X \in \mathscr{P}([1,n])} \sum_{k \in X} k = \sum_{k=1}^{n} k 2^{n-1} = n(n+1)2^{n-2}.$$

Exercice B

1. Notons pour simplifier $u_k = f\left(10^k - 1\right)$ pour $k \in \mathbb{N}$. Pour tout $k \geq 1$, partitionnons les 1 présents lorsque l'on écrit tous les entiers de 1 à $10^k - 1$:

- il y a 10^{k-1} chiffres 1 en première position dans des nombres à k chiffres,
- il y a $10u_{k-1}$ chiffres 1 dans les autres positions (le 10 correspond aux possibilités pour le premier chiffre).

Ainsi, $u_k = 10^{k-1} + 10u_{k-1}$. Comme $u_0 = 1$, cette relation de récurrence donne $u_k = k10^{k-1}$ pour tout $k \geq 1$ (par exemple, en utilisant un télescopage depuis la relation $\frac{u_k}{10^k} - \frac{u_{k-1}}{10^{k-1}} = \frac{1}{10}$).

2. Remarquons que la fonction f est croissante. Soit $n \in \mathbb{N} \setminus \{0\}$ et $k \in \mathbb{N}$ tel que $10^k - 1 \geq n \geq 10^{k-1}$ (k est le nombre de chiffres de n). Alors,

$$k10^{k-1} = u_k \geq f(n) \geq u_{k-1} = (k-1)10^{k-2}.$$

Si $n \geq f(n)$, alors en particulier $10^k \geq (k-1)10^{k-2}$ donc $k \leq 101$. Par conséquent, aucun entier n supérieur à 10^{101} ne peut vérifier l'hypothèse $n \geq f(n)$: il n'y a qu'un nombre fini d'entiers vérifiant cette condition.

Exercice C

Raisonnons par l'absurde en supposant qu'il existe une partie A ne vérifiant pas la conclusion de l'énoncé. Alors, les parties A et $B = \{n+9,\ n \in A \cap [\![1,91]\!]\}$ sont disjointes dans $[\![1,100]\!]$ d'où la contradiction suivante sur le cardinal de $A \cup B$

$$\begin{aligned} 100 = |[\![1,100]\!]| &\geq |A \cup B| = |A| + |B| \\ &\geq |A| + |A \cap [\![1,91]\!]| \\ &\geq 2|A| - |A \cap [\![92,100]\!]| \\ &\geq 2 \times 55 - 9 = 101. \end{aligned}$$

Remarque

Remarquons que le résultat n'est plus vrai pour la partie de cardinal 54 définie par

$$[\![1,9]\!] \cup [\![19,27]\!] \cup [\![37,45]\!] \cup [\![55,63]\!] \cup [\![73,81]\!] \cup [\![91,99]\!].$$

Exercice D

Partitionnons l'ensemble des parties de $[\![1,n]\!]$ de cardinal k qui ne contiennent pas deux entiers consécutifs selon que n y appartient ou pas :

- il y a $u_{n-1,k}$ parties de $[\![1,n]\!]$ de cardinal k qui ne contiennent ni n, ni deux entiers consécutifs ;
- il y a $u_{n-2,k-1}$ parties de $[\![1,n]\!]$ de cardinal k qui ne contiennent pas deux entiers consécutifs mais qui contiennent n (et donc pas $n-1$) ;

Ainsi, pour tous n, k, $u_{n,k} = u_{n-1,k} + u_{n-2,k-1}$. Montrons alors, par récurrence sur $n \in \mathbb{N} \setminus \{0\}$, la propriété $\mathscr{P}(n)$: « pour tout k, $u_{n,k} = \binom{n-k+1}{k}$ ».

- Pour $n = 1$, la propriété est immédiate car $u_{1,0} = u_{1,1} = 1$ et $u_{1,k} = 0$ pour $k \geq 2$.
- Pour $n = 2$, la propriété est immédiate car $u_{2,0} = u_{2,1} = 1$ et $u_{2,k} = 0$ pour $k \geq 2$.
- Soit $n \geq 2$ tel que $\mathscr{P}(j)$ soit vérifiée pour tout $j \leq n$. Alors, $u_{n+1,0} = 1 = \binom{n+1-0+1}{0}$ et, pour tout $k \geq 1$,

$$u_{n+1,k} = u_{n,k} + u_{n-1,k-1} = \binom{n-k+1}{k} + \binom{n-k+1}{k-1} = \binom{n+1-k+1}{k},$$

d'après la formule de Pascal.

Exercice E

Notons $C_1, \ldots, C_k$ les différentes classes d'équivalence pour $\mathscr{R}$ et $\overline{x}$ la classe d'un élément $x \in E$. Alors, en partitionnant l'ensemble des couples $(x,y) \in E^2$ tels que $x \mathscr{R} y$ selon leur première composante, on obtient

$$\begin{aligned} km &= k \sum_{x \in E} |\overline{x}| = k \sum_{i=1}^{k} \sum_{x \in C_i} |\overline{x}| \\ &= k \sum_{i=1}^{k} \sum_{x \in C_i} |C_i| = \Big(\sum_{i=1}^{k} 1^2\Big)\Big(\sum_{i=1}^{k} |C_i|^2\Big). \end{aligned}$$

Il reste alors à appliquer l'inégalité de Cauchy-Schwarz pour obtenir

$$km \geq \Big(\sum_{i=1}^{k} 1 \times |C_i|\Big)^2 = n^2$$

puisque les classes d'équivalence C_i forment une partition de E.

Exercice F

On pose $I = \{1, \dots, m\}$ et on considère l'application $f : S \longrightarrow \mathscr{P}(I)$ définie pour $x \in S$ par

$$f(x) = \{i \in I,\ x \in A_i\}.$$

Soit $x, y \in S$, avec $x \neq y$. D'après l'énoncé, il existe donc $i \in I$ tel que $i \in f(x)$ et $i \notin f(y)$, ou tel que $i \in f(y)$ et $i \notin f(x)$, ce qui assure que, dans tous les cas, on a $f(x) \neq f(y)$.
L'application f est donc injective, d'où $|S| \leq |\mathscr{P}(I)|$, ou encore $n \leq 2^m$ et la conclusion en découle.

Corrigés des problèmes du chapitre 23

Problème 1

Partie A – Dénombrement des surjections

1. a. ▹ Pour $p > n$, il n'y a pas de surjection donc $S_{n,p} = 0$. Si $p = n$, une surjection entre ces deux ensembles de même cardinal est bijection d'où $S_{n,n} = n!$.
▹ Si $p = 1$, il y une seule application (la fonction constante) et elle est surjective : $S_{n,1} = 1$.
▹ Si $p = 2$, une application de $[\![1, n]\!]$ dans $[\![1, 2]\!]$ est uniquement déterminée par l'ensemble des antécédents de 1. Toutes les applications ainsi associées à des parties de $[\![1, n]\!]$ sont surjectives sauf celles associées à $\varnothing$ et à $[\![1, n]\!]$; donc $S_{n,2} = 2^n - 2$.
b. Pour une application de $S_{p+1,p}$, tous les éléments de $[\![1, p]\!]$ (sauf un) admettent un unique antécédent (qui en admet deux). Commençons par choisir le cas particulier : il y a p choix à l'arrivée et $\binom{p+1}{2}$ choix pour ses antécédents ; il reste alors une permutation des $p-1$ autres éléments soit $(p-1)!$ choix. Par conséquent, $S_{p+1,p} = p.\binom{p+1}{2}.(p-1)! = \frac{1}{2}p(p+1)!$.
2. a. Proposons deux preuves de ce résultat.
▹ Par le calcul avec les factorielles,

$$\begin{aligned}\binom{p}{k}\binom{k}{q} &= \frac{p!}{k!(p-k)!}\frac{k!}{q!(q-k)!} = \frac{p!}{q!(p-q)!}\frac{(p-q)!}{(p-k)!(k-q)!}\\ &= \binom{p}{q}\binom{p-q}{k-q}.\end{aligned}$$

▹ Par interprétation combinatoire ; soit C un ensemble de cardinal p. $\binom{p}{k}\binom{k}{q}$ est le nombre de façon de choisir des parties A et B telles que $A \subset B \subset C$, A de cardinal q et B de cardinal k. On peut aussi dénombrer ces couples d'ensembles en cherchant A de cardinal q dans C puis $B \setminus A$ de cardinal $k-q$ dans $C \setminus A$ qui est de cardinal $p-q$: il y en a donc $\binom{p}{q}\binom{p-q}{k-q}$.
b. D'après la question précédente,

$$\begin{aligned}\sum_{k=q}^{p}(-1)^k\binom{p}{k}\binom{k}{q} &= \sum_{k=q}^{p}(-1)^k\binom{p}{q}\binom{p-q}{k-q}\\ &= \binom{p}{q}(-1)^q\sum_{k=0}^{p-q}(-1)^k\binom{p-q}{k}\\ &= \binom{p}{q}(-1)^q(1-1)^{p-q} = 0.\end{aligned}$$

3. a. Partitionnons l'ensemble des p^n applications selon le cardinal de l'image. Il y a $\binom{p}{q}$ choix pour un ensemble de cardinal dans $[\![1, p]\!]$ puis $S_{n,q}$ applications qui admettent cette image. Par conséquent,

$$p^n = \sum_{q=1}^{p}\binom{p}{q}S_{n,q}.$$

b. Utilisons les deux dernières questions :

$$\begin{aligned}(-1)^p\sum_{k=1}^{p}(-1)^k\binom{p}{k}k^n &= (-1)^p\sum_{k=1}^{p}(-1)^k\binom{p}{k}\sum_{q=1}^{k}\binom{k}{q}S_{n,q}\\ &= (-1)^p\sum_{q=1}^{p}\Big[\sum_{k=q}^{p}(-1)^k\binom{p}{k}\binom{k}{q}\Big]S_{n,q} = S_{n,p}.\end{aligned}$$

En effet, le terme entre crochets est nul si $q \neq p$.
c. Avec la formule du triangle de Pascal, le second membre devient

$$\begin{aligned}p(S_{n-1,p}+S_{n-1,p-1}) &= (-1)^p\sum_{k=1}^{p}(-1)^k p\binom{p-1}{k-1}k^{n-1}\\ &= (-1)^p\sum_{k=1}^{p}(-1)^k\binom{p}{k}k^n = S_{n,p},\end{aligned}$$

car $p\binom{p-1}{k-1} = k\binom{p}{k}$.
d. En notant $u_p = S_{p+1,p-1}$ et en utilisant l'expression de $S_{p+1,p}$ trouvée précédemment (pour $n = p+2$), la relation devient

$$u_{p+1} = S_{p+2,p} = p(S_{p+1,p}+S_{p+1,p-1}) = p\Big(\frac{1}{2}p(p+1)! + u_p\Big).$$

Une simple récurrence permet de démontrer la formule proposée.

Partie B – Dénombrement des dérangements

1. On partitionne les $n!$ permutations selon le cardinal de leur support. Il y a $\binom{n}{k}$ façons de choisir un support de cardinal k puis d_k façon de permuter sans point fixe ces k éléments. Par conséquent,

$$n! = \sum_{k=0}^{n}\binom{n}{k}d_k.$$

2. On procède exactement comme à la partie A.

Problème 2

Partie A – Inégalité de Bollobas-Lubell-Yamamoto-Meshalkin

1. Une chaîne maximale de l'ensemble $\{a,b,c\}$ contient des ensembles de cardinal $0,1,2,3$. On obtient six chaînes :

- $\varnothing, \{a\}, \{a,b\}, \{a,b,c\}$
- $\varnothing, \{a\}, \{a,c\}, \{a,b,c\}$
- $\varnothing, \{b\}, \{a,b\}, \{a,b,c\}$
- $\varnothing, \{b\}, \{b,c\}, \{a,b,c\}$
- $\varnothing, \{c\}, \{a,c\}, \{a,b,c\}$
- $\varnothing, \{c\}, \{b,c\}, \{a,b,c\}$

2. a. Une chaîne maximale de E est une suite strictement croissante d'ensembles de cardinal entre 0 et n : elle est uniquement déterminée par la suite des éléments rajoutés pour passer d'un ensemble au suivant, c'est-à-dire par la bijection de $[\![1,n]\!]$ dans E qui à l'entier k associe l'unique élément appartenant au $(k+1)^{\text{ième}}$ ensemble et pas au $k^{\text{ième}}$. Or, il y a $n!$ telles bijections.
b. Soit une partie $A \subset E$ de cardinal k. Une chaîne maximale contenant A est composée d'une chaîne maximale de A (il y en a $k!$) et d'une chaîne obtenue en ajoutant A à chaque partie d'une chaîne maximale du complémentaire de A (il y en a $(n-k)!$). En conclusion, il y a donc $k!(n-k)!$ chaînes maximales contenant A.

3. Si A et B appartiennent à une même chaîne, ces deux parties sont comparables pour l'inclusion donc n'appartiennent pas à la même antichaîne. Le résultat est démontré par contraposition.
4. Chacune des parties d'une antichaîne de E appartient à $k!(n-k)!$ chaînes et ces chaînes sont deux à deux distinctes d'où

$$\sum_{k=0}^{n} a_k k!(n-k)! \leq n!.$$

Partie B – Théorème de Sperner

1. La famille des parties de cardinal $\lfloor \frac{n}{2} \rfloor$ forme une antichaîne ; or, cette famille est de cardinal $\binom{n}{\lfloor \frac{n}{2} \rfloor}$. En conclusion, $C_n \geq \binom{n}{\lfloor \frac{n}{2} \rfloor}$.
2. a. Pour tout $k \in [\![1, n]\!]$,

$$\frac{\binom{n}{k+1}}{\binom{n}{k}} = \frac{n-k}{k+1}.$$

Or, $\frac{n-k}{k+1} \geq 1$ si, et seulement si, $k \leq \frac{n-1}{2}$. Ainsi, la suite $\left(\binom{n}{k}\right)_k$ est croissante jusqu'à $k \lfloor \frac{n}{2} \rfloor$, puis décroissante ; d'où

$$\forall k \in [\![0, n]\!], \qquad \binom{n}{k} \leq \binom{n}{\lfloor \frac{n}{2} \rfloor}.$$

b. D'après la partie A, pour une antichaîne maximale

$$\sum_{k=0}^{n} \frac{a_k}{\binom{n}{k}} \leq 1.$$

Or,

$$\sum_{k=0}^{n} \frac{a_k}{\binom{n}{k}} \geq \frac{1}{\binom{n}{\lfloor \frac{n}{2} \rfloor}} \sum_{k=0}^{n} a_k \geq \frac{C_n}{\binom{n}{\lfloor \frac{n}{2} \rfloor}}.$$

En conclusion, $C_n \leq \binom{n}{\lfloor \frac{n}{2} \rfloor}$.
3. En combinant les deux inégalités, on obtient $C_n = \binom{n}{\lfloor \frac{n}{2} \rfloor}$.

Problème 3

Partie A – Exemples

1. S'il existe un système de Steiner avec $p < q < n$, alors, d'après (iii), toute partie à q éléments de E est incluse dans une partie de cardinal $p < q$ (un élément du système de Steiner) ce qui est impossible.
2. Dans un système de Steiner $S(q, q, n))$ pour $q \leq n$, toute partie à q éléments de E est incluse dans l'une des parties des parties à q éléments du système de Steiner (propriété (iii)) donc égale à l'une de ces parties. Par conséquent, le système de Steiner $S(q, q, n))$ est l'ensemble de toutes les parties à q éléments de E et $|S(q, q, n)| = \binom{n}{q}$.
3. ▹ D'après la propriété (iii), tout singleton de E appartient à une unique partie à p éléments d'un système de Steiner $S(p, 1, n)$; par conséquent les parties de cardinal p constituant ce système de Steiner forment une partition de E et donc p divise n.
▹ Réciproquement, si p divise n, on peut partitionner E en ensembles de cardinal p : les ensembles formant cette partition sont les éléments d'un système de Steiner $S(p, 1, n)$.
Enfin, lorsque p divise n, $|S(p, 1, n)| = \frac{n}{p}$.

Partie B – Cardinal d'un système de Steiner

1. a. Pour déterminer une partie de E de cardinal q contenant x, il faut et il suffit de choisir les $q-1$ éléments de E distincts de x : il y a donc $\binom{n-1}{q-1}$ telles parties.
b. Si $a \notin A$, ce nombre est 0. Sinon, pour déterminer une partie de E de cardinal q incluse dans une partie A de cardinal p, contenant x, il faut et il suffit de choisir les $q-1$ éléments de A distincts de x : il y a donc $\binom{p-1}{q-1}$ telles parties.
c. Partitionnons l'ensemble des parties à q éléments contenant x selon la partie du système de Steiner $S(p,q,n)$ dans laquelle elle est incluse (c'est effectivement une partition en vertu de (iii)). D'après la question précédente, chaque partie du système de Steiner contenant x contient $\binom{p-1}{q-1}$ parties à q éléments contenant x. Ainsi,

$$\binom{n-1}{q-1} = \binom{p-1}{q-1}.r(p,q,n)$$

d'où $r(p,q,n) = \frac{\binom{n-1}{q-1}}{\binom{p-1}{q-1}}$.
2. Dénombrons que deux façons distinctes l'ensemble $\mathfrak{E} = \{(x,A),\ x \in E,\ A \in S(p,q,n),\ x \in A\}$.
• D'une part, on peut partitionner l'ensemble $\mathfrak{E}$ selon le premier élément du couple (i.e. l'élément $x \in E$). Pour chaque $x \in E$, il existe $r(p,q,n)$ parties A du système de Steiner telles que $x \in A$. Le cardinal de $\mathfrak{E}$ est donc $n\,r(p,q,n)$.
• D'autre part, on peut partitionner l'ensemble $\mathfrak{E}$ selon le second élément du couple (c'est-à-dire l'élément du système de Steiner $A \in S(p,q,n)$). Chaque $A \in S(p,q,n)$ contient exactement p éléments de E. Le cardinal de $\mathfrak{E}$ est donc $p|S(p,q,n)|$.
En conclusion, $|S(p,q,n)| = \frac{n}{p}\,r(p,q,n)$.

Partie C – Demoiselles de Kirkman

D'après la partie B, $|S(3,2,15)| = 35$: il y a donc 35 triplets de demoiselles tels que deux demoiselles quelconques ne se retrouvent pas deux fois dans le même rang.

Lors de chaque promenade, il y a cinq rangs ; donc il y aura au plus 7 jours possibles de promenade.

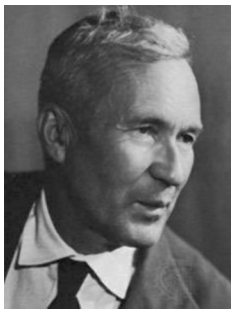
Chapitre 24

Probabilités sur un univers fini

1. Espace probabilisé fini — **2.** Conditionnement — **3.** Indépendance

Objectifs et compétences du programme

- Trouver le formalisme probabiliste adapté à une situation de modélisation.
- Utiliser les opérateurs ensemblistes sur les événements.
- Identifier et exploiter une propriété d'indépendance.
- Calculer la probabilité d'événements simples.



Andreï Nikolaïevitch Kolmogorov, mathématicien russe (1903-1987). Bien qu'ayant traversé les périodes difficiles de la science soviétique soumise au pouvoir politique, il a contribué à de nombreux domaines. Il propose en 1932 dans un ouvrage une axiomatisation complète de la théorie des probabilités ; et dès l'année suivante, il publie des résultats de topologie ! Dans les années 40, il décrit les phénomènes de cascades turbulents. Pendant la décennie suivante, il révolutionne l'étude des systèmes dynamiques avec ce que l'on appelle désormais la théorie KAM (Kolmogorov-Arnol'd-Moser) ; plus tard encore, il fonde la théorie de la complexité algorithmique (complexité de Kolmogorov)... Difficile de trouver des mathématiciens avec des recherches recouvrant un si grand spectre thématique et des contributions à chaque fois fondamentale !

COURS

Ce chapitre introductif a pour objectif de fournir le cadre formel aux calculs de probabilités d'événements. On introduit les notions d'espace probabilisé et d'événements, puis on utilise les différentes opérations ensemblistes (intersection, réunion, complémentaire...) pour les manipuler. Le conditionnement permet ensuite d'étudier des situations plus compliquées que d'éventuelles propriétés d'indépendance simplifient.

1. Espace probabilisé fini

1.1. Définition

Définition 24.1.

Un espace probabilisable fini est un couple $(\Omega, \mathscr{P}(\Omega))$ où Ω est un ensemble fini appelé univers.
Les éléments de Ω sont appelés résultats ou réalisations, un élément de $\mathscr{P}(\Omega)$ est appelé événement ou événement observable.

Remarque

Il peut sembler étrange de considérer un couple $(\Omega, \mathscr{P}(\Omega))$ quand la seconde composante est entièrement déterminée par la première. Ce point de la définition change lorsque l'on généralise et que Ω n'est plus un ensemble fini : la seconde composante est alors une partie de $\mathscr{P}(\Omega)$ vérifiant certaines propriétés de stabilité.

On justifiera à la fin de cette partie l'importance de cette définition et les différentes appellations. Toutefois, afin de ne pas mélanger le travail de modélisation d'un phénomène aléatoire et celui de raisonnement dans la théorie des probabilités, on préfère séparer la définition abstraite de son interprétation. Pour l'instant, fixons le vocabulaire utilisé.

Définition 24.2.

Soit $(\Omega, \mathscr{P}(\Omega))$ un espace probabilisable, A, $B \subset \Omega$.
▹ L'événement A et B est défini par l'intersection $A \cap B$.
▹ L'événement A ou B est défini par la réunion $A \cup B$.
▹ L'événement contraire de A est le complémentaire $\overline{A}$.
▹ Les événements A et B sont incompatibles s'ils sont disjoints.
▹ L'événement Ω est appelé certain et l'événement $\varnothing$ impossible.

Définition 24.3.

Soit $(\Omega, \mathscr{P}(\Omega))$ un espace probabilisable. Un système complet d'événements est une partition de Ω.

1.2. Probabilité sur un univers fini

Définition 24.4.

Une probabilité sur l'espace probabilisable fini $(\Omega, \mathscr{P}(\Omega))$ est une application $\mathbf{P} : \mathscr{P}(\Omega) \to [0,1]$ telle que
- $\mathbf{P}(\Omega) = 1$,
- pour toutes parties A, $B \subset \Omega$ disjointes, $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B)$.

Le triplet $(\Omega, \mathscr{P}(\Omega), \mathbf{P})$ est appelé espace probabilisé.

La deuxième propriété entraîne par une récurrence immédiate le résultat suivant.

Proposition 24.5. Additivité

Soit $\mathbf{P}$ une probabilité sur l'espace probabilisable $(\Omega, \mathscr{P}(\Omega))$. Alors, pour toute famille $A_1, \dots, A_n$ d'événements deux à deux disjoints,

$$\mathbf{P}\Big(\bigcup_{k=1}^{n} A_k\Big) = \sum_{k=1}^{n} \mathbf{P}(A_k).$$

Proposition 24.6.

Soit $(\Omega, \mathscr{P}(\Omega), \mathbf{P})$ un espace probabilisé fini. Alors, pour tout $A \subset \Omega$,

$$\mathbf{P}(A) = \sum_{x \in A} \mathbf{P}(\{x\}).$$

Démonstration

Il suffit de remarquer que A est la réunion disjointe des singletons de ses éléments. ■

Remarque

En fait, on comprend avec cette propriété qu'une probabilité sur un espace probabilisable est uniquement déterminée par les images des singletons (à condition que la somme de ces images soit égale à 1).

Définition 24.7.

Soit $(\Omega, \mathscr{A}, \mathbf{P})$ un espace probabilisé.
▹ Un événement $A \in \mathscr{A}$ est négligeable si $\mathbf{P}(A) = 0$.
▹ Un événement $A \in \mathscr{A}$ est presque sûr (en abrégé p.s.) si $\mathbf{P}(A) = 1$.

Avec l'additivité, on obtient immédiatement que :

Proposition 24.8.

Toute réunion finie d'événements négligeables est négligeable et toute intersection finie d'événements presque sûr est presque sûre.

Proposition 24.9.

Soit $(\Omega, \mathscr{P}(\Omega), \mathbf{P})$ un espace probabilisé fini. Alors,
▹ pour tout $A \subset \Omega$, $\mathbf{P}(\overline{A}) = 1 - \mathbf{P}(A)$;
▹ pour tous A, $B \subset \Omega$ tels que $A \subset B$, $\mathbf{P}(A) \leq \mathbf{P}(B)$;
▹ pour tous A, $B \subset \Omega$, $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B)$.

Démonstration

▹ Soit $A \subset \Omega$. Les événements A et $\overline{A}$ sont disjoints donc

$$\mathbf{P}(A) + \mathbf{P}(\overline{A}) = \mathbf{P}(A \cup \overline{A}) = \mathbf{P}(\Omega) = 1.$$

▹ Soit A, $B \subset \Omega$ tels que $A \subset B$. Alors, A et $B \cap \overline{A}$ sont disjoints de réunion B donc

$$\mathbf{P}(B) = \mathbf{P}(A) + \mathbf{P}(B \cap \overline{A}) \geq \mathbf{P}(A).$$

▹ Soit A, $B \subset \Omega$. Alors, comme $A \cap \overline{B}$ et B sont disjoints de réunion $A \cup B$,

$$\mathbf{P}(A \cup B) = \mathbf{P}(B) + \mathbf{P}(A \cap \overline{B}).$$

Par ailleurs, $A \cap \overline{B}$ et $A \cap B$ sont disjoints et de réunion A donc

$$\mathbf{P}(A) = \mathbf{P}(A \cap B) + \mathbf{P}(A \cap \overline{B}).$$

Le résultat suit en effectuant la différence de ces deux dernières égalités. ■

La proximité de cette propriété avec les résultats analogues pour la fonction cardinal se comprend bien avec le résultat suivant.

Proposition 24.10.

Soit $(\Omega, \mathscr{P}(\Omega))$ un espace probabilisable fini. Alors, l'application

$$\mathbf{P} : \left\{ \begin{array}{ccc} \mathscr{P}(\Omega) & \longrightarrow & [0,1] \\ A & \longmapsto & \dfrac{|A|}{|\Omega|} \end{array} \right.$$

est une probabilité appelée probabilité uniforme.

Démonstration

Vérifions les deux propriétés de probabilité.
▹ $\mathbf{P}(\Omega) = \frac{|\Omega|}{|\Omega|} = 1$.
▹ Soit A et $B \subset \Omega$ disjoints. Alors,

$$\mathbf{P}(A \cup B) = \frac{|A \cup B|}{|\Omega|} = \frac{|A| + |B|}{|\Omega|} = \mathbf{P}(A) + \mathbf{P}(B).$$

■

Remarque

La probabilité uniforme est bien souvent la loi par défaut lorsque l'on ne dispose pas d'autres informations sur le problème considéré.

La règle pour la probabilité d'une réunion de deux événements peut se généraliser à un nombre quelconque d'événements.

Proposition 24.11. Formule du crible de Poincaré

Soit $(\Omega, \mathscr{P}(\Omega), \mathbf{P})$ un espace probabilisé fini et $A_1, \ldots, A_n \in \mathscr{P}(\Omega)$ des événements. Alors,

$$\mathbf{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{\substack{I \subset [\![1,n]\!] \\ I \neq \varnothing}} (-1)^{-1+|I|}\, \mathbf{P}\left(\bigcap_{i \in I} A_i\right).$$

Démonstration

Procédons par récurrence sur l'entier $n \geq 2$.

▹ Pour $n = 2$, la formule est déjà établie.

▹ Soit $n \geq 2$. Supposons le résultat acquis au rang n et montrons-le au rang $n+1$ avec des événements $A_1, \ldots, A_{n+1}$. Notons $\widehat{A}_n$ la réunion de $A_1, \ldots, A_n$. Alors, en utilisant tour à tour la relation pour $n = 2$ et l'hypothèse de récurrence pour $\mathbf{P}(\widehat{A}_n)$, on obtient

$$\begin{aligned}
\mathbf{P}\left(\widehat{A}_n \cup A_{n+1}\right) &= \mathbf{P}(\widehat{A}_n) + \mathbf{P}(A_{n+1}) - \mathbf{P}(\widehat{A}_n \cap A_{n+1}) \\
&= \mathbf{P}(\widehat{A}_n) + \mathbf{P}(A_{n+1}) - \mathbf{P}\left(\bigcup_{i=1}^{n} A_i \cap A_{n+1}\right) \\
&= \sum_{\substack{I \subset [\![1,n]\!] \\ I \neq \varnothing}} (-1)^{-1+|I|}\, \mathbf{P}\left(\bigcap_{i \in I} A_i\right) + \mathbf{P}(A_{n+1}) - \sum_{\substack{I \subset [\![1,n]\!] \\ I \neq \varnothing}} (-1)^{-1+|I|}\, \mathbf{P}\left(\bigcap_{i \in I} A_i \cap A_{n+1}\right) \\
&= \sum_{\substack{I \subset [\![1,n+1]\!] \\ I \neq \varnothing, n+1 \notin I}} (-1)^{-1+|I|}\, \mathbf{P}\left(\bigcap_{i \in I} A_i\right) + \mathbf{P}(A_{n+1}) + \sum_{\substack{I \subset [\![1,n+1]\!] \\ I \neq \{n+1\}, n+1 \in I}} (-1)^{-1+|I|}\, \mathbf{P}\left(\bigcap_{i \in I} A_i\right) \\
&= \sum_{\substack{I \subset [\![1,n+1]\!] \\ I \neq \varnothing}} (-1)^{-1+|I|}\, \mathbf{P}\left(\bigcap_{i \in I} A_i\right).
\end{aligned}$$

■

1.3. Modélisation

Nous venons d'introduire le formalisme permettant d'étudier des expériences aléatoires. Un travail complet consiste à modéliser l'expérience aléatoire en définissant l'espace probabilisé correspondant puis à oublier la situation concrète pour faire des calculs reposant sur la théorie des probabilités (qui s'applique comme n'importe quelle théorie mathématique). Ce fonctionnement en deux temps n'est pas propre aux probabilités mais il y est davantage perceptible. Par exemple, on pourrait très bien imaginer un problème de mécanique du solide : dans un premier temps, on établit les équations différentielles du mouvement ; dans un second temps, on les étudie avec la théorie développée dans les premiers chapitres du livre.

Dans cette partie, on donne quelques exemples de modélisations : Ω est l'ensemble des résultats de l'expérience, un événement est un ensemble de résultats, $\mathbf{P}$ mesure la notion de probabilité. Dans la suite du chapitre, on développera la théorie pour effectuer des calculs.

Exemple

Le jeu de Pile ou Face consiste à lancer une pièce et à regarder la partie visible de la pièce (Pile ou Face) une fois celle-ci retombée.
Avec une pièce équilibrée usuelle, l'univers des résultats qui s'impose est $\Omega = \{F, P\}$ (avec des notations évidentes) et la probabilité $\mathbf{P}$ associée vérifie $\mathbf{P}(\{F\}) = \mathbf{P}(\{P\}) = \frac{1}{2}$.
Si on veut modéliser une pièce biaisée, il suffit de changer la probabilité $\mathbf{P}$; si on veut considérer la possibilité d'une pièce retombant sur la tranche ou ne retombant pas, il faudrait modifier l'univers pour introduire un troisième résultat possible.

Exemple

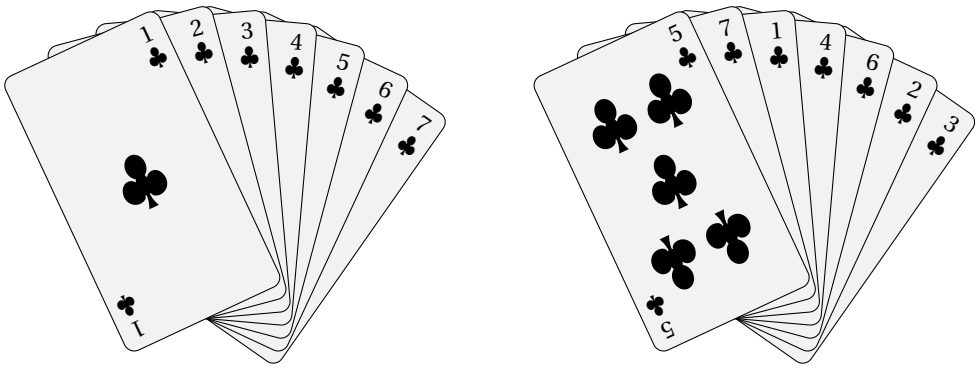
▹ Pour modéliser le lancer d'un dé usuel, l'univers qui s'impose est $\Omega = [\![1,6]\!]$ et la probabilité est uniforme.
▹ Pour modéliser le lancer de deux dés usuels, l'univers qui s'impose est $\Omega = [\![1,6]\!]^2$ et la probabilité est uniforme. Dans ce cadre, l'événement « la somme des deux nombres obtenus est un multiple de 3 » est

$$A = \{(1,2),(1,5),(2,1),(2,4),(3,3),(3,6),(4,2),(4,5),(5,1),(5,4),(6,3),(6,6)\}$$

et sa probabilité est $\frac{12}{36} = \frac{1}{3}$.

Exemple

Modélisons le battage d'un jeu de n cartes. Voici une illustration d'un jeu de $n = 7$ cartes avant et après le battage.



En notant les cartes de 1 à n, l'univers naturel est alors $\Omega = S_n$: pour chaque résultat σ, $\sigma(k)$ est la position de la carte numérotée k.
Dans l'exemple illustré ci-dessus, $\sigma(1) = 3$ car la carte initialement en première position est, après mélange, en troisième position ; plus généralement,

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 6 & 7 & 4 & 1 & 5 & 2 \end{pmatrix}.$$

Si la personne qui manipule le jeu a suffisamment de dextérité, on peut considérer Ω muni de la probabilité uniforme ; sinon, il faudra savoir quels sont les défauts du mélangeur pour trouver la probabilité adaptée.

Exemple

Considérons une situation plus compliquée ; déterminons la probabilité d'obtenir au moins une fois face en deux lancers d'une pièce équilibrée et montrons l'importance de la modélisation initiale du problème.
▹ Naturellement l'espace probabilisé est défini par $\Omega = \{P, F\}^2$ muni de la probabilité uniforme. Avec cette modélisation, l'événement étudié est $\{(P,F),(F,P),(F,F)\}$ et sa probabilité est $\frac{3}{4}$.
▹ Toutefois, on peut supposer que si le premier résultat est Face, on ne va pas relancer la pièce : l'univers serait $\Omega = \{F, PF, PP\}$; l'événement étudié est $\{F, PF\}$.

- avec la probabilité uniforme, cet événement a pour probabilité $\frac{2}{3}$;
- avec la probabilité telle que $\mathbf{P}(\{F\}) = \frac{1}{2}$ et $\mathbf{P}(\{PP\}) = \mathbf{P}(\{PF\}) = \frac{1}{4}$, cet événement a pour probabilité $\frac{3}{4}$.

Aucun de ces deux résultats n'est mathématiquement faux, ils répondent juste à des questions différentes car la modélisation est différente.

Exemple

Le jeu de Monty Hall est un jeu télévisé : un candidat se trouve aux côtés du présentateur et devant trois portes fermées. Derrière l'une de ces portes se trouve une voiture de sport, derrière les deux autres, une chèvre. Le candidat choisit une porte sans l'ouvrir. Le présentateur ouvre alors l'une des deux autres portes, porte derrière laquelle il n'y a pas la voiture. Enfin, le candidat choisit s'il ouvre la porte initialement choisie ou l'autre qui reste close. Calculons la probabilité de gagner (la voiture) en changeant de portes à la dernière étape du jeu.
Le jeu avec cette stratégie pour le candidat se résume en la succession des trois actions suivantes

- le candidat choisit une porte parmi les trois possibles,
- le présentateur ouvre l'une des deux portes restantes qui cache une chèvre,
- le candidat change de porte.

La porte qui cache la voiture est notée 1, les deux autres 2 et 3. On modélise un résultat du jeu par un triplet (a,b,c) où a est le numéro de la porte choisie initialement, b celui de la porte ouverte par le présentateur, c celui de celle finalement choisie (cette porte cache la voiture si $c=1$). Ainsi, l'univers est

$$\Omega = \{(1,2,3),(1,3,2),(2,3,1),(3,2,1)\}.$$

Pour définir une probabilité $\mathbf{P}$ sur cet univers qui soit adaptée à notre modélisation, il convient de supposer que le choix initial du candidat est uniforme donc

$$\mathbf{P}(\{(1,2,3),(1,3,2)\}) = \mathbf{P}(\{(2,3,1)\}) = \mathbf{P}(\{(3,2,1)\}) = \frac{1}{3}.$$

Sans hypothèse sur la façon dont le présentateur ouvre la porte dans le cas où le candidat a initialement choisi la bonne porte, il est difficile de préciser $\mathbf{P}(\{(1,2,3)\})$ et $\mathbf{P}(\{(1,3,2)\})$ mais cela ne perturbe pas le calcul.
Ainsi, la probabilité de gain du candidat qui change de porte est $\mathbf{P}(\{(2,3,1),(3,2,1)\}) = \frac{2}{3}$.

2. Conditionnement

2.1. Probabilité conditionnelle

Proposition 24.12.

Soit $(\Omega, \mathscr{P}(\Omega), \mathbf{P})$ un espace probabilisé fini et $A \subset \Omega$ de probabilité non nulle. L'application

$$\mathbf{P}_A : \left\{ \begin{array}{ccc} \mathscr{P}(\Omega) & \longrightarrow & [0,1] \\ B & \longmapsto & \frac{\mathbf{P}(A\cap B)}{\mathbf{P}(A)} \end{array} \right.$$

est une probabilité appelée probabilité conditionnellement à A (ou probabilité sachant A).
On note également $\mathbf{P}(B|A) = \mathbf{P}_A(B)$ pour tout événement $B \subset \Omega$.

Démonstration

Vérifions les deux propriétés de probabilité.

▹ $\mathbf{P}_A(\Omega) = \frac{\mathbf{P}(A\cap\Omega)}{\mathbf{P}(A)} = \frac{\mathbf{P}(A)}{\mathbf{P}(A)} = 1.$

▹ Soit B et $C \subset \Omega$ disjoints. Alors, $A \cap B$ et $A \cap C$ sont disjoints et

$$\begin{aligned}\mathbf{P}_A(B \cup C) &= \frac{\mathbf{P}(A \cap (B \cup C))}{\mathbf{P}(A)} \\ &= \frac{\mathbf{P}((A \cap B) \cup (A \cap C))}{\mathbf{P}(A)} \\ &= \frac{\mathbf{P}(A \cap B) + \mathbf{P}(A \cap C)}{\mathbf{P}(A)} = \mathbf{P}_A(B) + \mathbf{P}_A(C).\end{aligned}$$

■

Remarque

La définition de probabilité conditionnelle correspond à un calcul intuitif de fréquences lors de n réalisations de l'expérience (et heuristiquement une probabilité est une « limite » de fréquence de réalisation de l'événement quand n tend vers $+\infty$). Notons f_A, $f_{A\cap B}$ les fréquences de réalisations des événements A et $A \cap B$ en n expériences, n_A et $n_{A\cap B}$ les nombres de réalisations de ces événements. Alors, la fréquence de B sachant A est le nombre de réalisations de B parmi les expériences qui réalisent A :

$$\frac{n_{A\cap B}}{n_A} = \frac{n_{A\cap B}}{n}\frac{n}{n_A} = \frac{f_{A\cap B}}{f_A}.$$

Remarque

Soit $(\Omega, \mathscr{P}(\Omega), \mathbf{P})$ un espace probabilisé fini. Pour tout $A \subset \Omega$, $\mathbf{P}(A|\Omega) = \mathbf{P}(A)$.

Proposition 24.13. Formule des probabilités composées

Soit $(\Omega, \mathscr{P}(\Omega), \mathbf{P})$ un espace probabilisé fini et $(A_k)_{k\leq n}$ une famille d'événements vérifiant

$$\mathbf{P}\Big(\bigcap_{k=1}^{n} A_k\Big) > 0.$$

Alors,

$$\mathbf{P}\Big(\bigcap_{k=1}^{n} A_k\Big) = \mathbf{P}\Big(A_n \Big| \bigcap_{k=1}^{n-1} A_k\Big)\mathbf{P}\Big(A_{n-1} \Big| \bigcap_{k=1}^{n-2} A_k\Big)\dots\mathbf{P}(A_2|A_1)\mathbf{P}(A_1).$$

Remarque

L'hypothèse $\mathbf{P}(\cap_{k=1}^{n} A_k) > 0$ entraîne $\mathbf{P}(\cap_{k=1}^{j} A_k) > 0$ pour tout $j \leq n$ puisque l'on dispose de l'inclusion

$$\bigcap_{k=1}^{n} A_k \subset \bigcap_{k=1}^{j} A_k$$

et de la croissance de $\mathbf{P}$.

Démonstration

Montrons le résultat par récurrence sur n.

Pour $n = 2$, le résultat est la définition de probabilité conditionnelle par rapport à A_1 (qui est bien de probabilité non nulle car il contient l'intersection qui est de probabilité non nulle).

Soit $n \in \mathbb{N}$ tel que le résultat est établi pour toute famille de n événements d'intersection de probabilité non nulle. Considérons $(A_k)_{k\leq n+1}$ une famille d'événements telle que

$$\mathbf{P}\Big(\bigcap_{k=1}^{n+1} A_k\Big) > 0.$$

Alors, d'après la définition de probabilité conditionnelle sachant $\cap_{k=1}^{n} A_k$,

$$\mathbf{P}\Big(\bigcap_{k=1}^{n+1} A_k\Big) = \mathbf{P}\Big(A_{n+1} \cap \Big(\bigcap_{k=1}^{n} A_k\Big)\Big) = \mathbf{P}\Big(A_{n+1} \Big| \bigcap_{k=1}^{n} A_k\Big)\mathbf{P}\Big(\bigcap_{k=1}^{n} A_k\Big)$$

et il ne reste qu'à appliquer l'hypothèse de récurrence. ■

2.2. Formule des probabilités totales

Proposition 24.14. Formule des probabilités totales

Soit $(\Omega, \mathscr{P}(\Omega), \mathbf{P})$ un espace probabilisé fini, $A_1, \ldots, A_n$ une partition de Ω d'événements de probabilités non nulles. Alors, pour tout $B \subset \Omega$,

$$\mathbf{P}(B) = \sum_{k=1}^{n} \mathbf{P}(B|A_k)\mathbf{P}(A_k).$$

Démonstration

Les événements $B \cap A_k$ pour $k \in [\![1, n]\!]$ forment une partition de B donc

$$\mathbf{P}(B) = \sum_{k=1}^{n} \mathbf{P}(B \cap A_k) = \sum_{k=1}^{n} \mathbf{P}(B|A_k)\mathbf{P}(A_k).$$

■

Conseils méthodologiques

La formule des probabilités totales sert souvent à « discuter » la probabilité d'un événement B selon les différentes issues possibles $A_1, \ldots, A_n$ d'une autre expérience. C'est le cas, par exemple, lorsque l'on cherche les probabilités associées à un second tirage dans une urne : on discute selon les résultats du premier tirage pour connaître la composition de l'urne au moment du deuxième tirage.

Remarque

Dans la formule précédente, on peut s'affranchir de l'hypothèse « de probabilités non nulles » en réécrivant

$$\mathbf{P}(B) = \sum_{\substack{k=1 \\ \mathbf{P}(A_k)>0}}^{n} \mathbf{P}(B|A_k)\mathbf{P}(A_k),$$

ou, plus simplement, en adoptant la convention $\mathbf{P}(B|A_k)\mathbf{P}(A_k) = 0$ dès que $\mathbf{P}(A_k) = 0$.

Voici un cas particulier simple avec $A_1 = A$ et $A_2 = \overline{A}$.

Corollaire 24.15.

Soit $(\Omega, \mathscr{P}(\Omega), \mathbf{P})$ un espace probabilisé fini, $A \subset \Omega$ de probabilité différente de 0 et de 1.
Alors, pour tout $B \subset \Omega$,

$$\mathbf{P}(B) = \mathbf{P}(B|A)\mathbf{P}(A) + \mathbf{P}(B|\overline{A})\mathbf{P}(\overline{A}).$$

Exemple

Calculons la probabilité que la somme des résultats du lancer de deux dés usuels soit égale à 9. L'espace probabilisé est construit avec l'univers $\Omega = [\![1,6]\!]^2$ muni de la probabilité uniforme. L'événement qui nous intéresse est $B = \{(x,y) \in \Omega,\ x+y=9\}$ et on considère la partition de Ω obtenue avec les événements $A_k = \{k\} \times [\![1,6]\!]$ pour $k \in [\![1,6]\!]$. La formule des probabilités totales donne alors $\mathbf{P}(B) = \sum_{k=1}^{6} \mathbf{P}(B|A_k)\mathbf{P}(A_k)$. Or, pour tout $k \in [\![1,6]\!]$, $\mathbf{P}(A_k) = \frac{|A_k|}{|\Omega|} = \frac{1}{6}$, donc

$$\begin{aligned}\mathbf{P}(B) &= \frac{1}{6}\sum_{k=1}^{6} \mathbf{P}(B|A_k) \\ &= \frac{1}{6}\Big(\mathbf{P}(\varnothing) + \mathbf{P}(\varnothing) + \mathbf{P}(\{3,6\}) + \mathbf{P}(\{4,5\}) + \mathbf{P}(\{5,4\}) + \mathbf{P}(\{6,3\})\Big) = \frac{1}{9}.\end{aligned}$$

On aurait pu bien évidemment calculer directement cette probabilité en dénombrant B.

Corollaire 24.16.

Soit $(\Omega, \mathscr{P}(\Omega), \mathbf{P})$ un espace probabilisé fini, $A \subset \Omega$ et $A_1, \ldots, A_n$ une partition de A d'événements de probabilités non nulles. Alors, pour tout $B \subset \Omega$,

$$\mathbf{P}(B|A) = \sum_{k=1}^{n} \mathbf{P}(B|A_k)\mathbf{P}(A_k|A).$$

Démonstration

Il suffit d'appliquer la formule des probabilités totales à la probabilité $\mathbf{P}_A$; on obtient alors

$$\mathbf{P}(B|A) = \sum_{k=1}^{n} \mathbf{P}(B|A_k \cap A)\mathbf{P}(A_k|A)$$

et on conclut car $A_k \subset A$. ■

Exemple Paradoxe de Simpson

Une université recrute des professeurs à l'aide de deux commissions : l'une pour les sciences, l'autre pour les humanités. Voici les effectifs de chaque commission

Sciences	Candidats	Admis
Hommes	500	400
Femmes	100	90

Humanités	Candidats	Admis
Hommes	100	20
Femmes	300	90

La probabilité qu'une femme (respectivement un homme) soit recrutée est de $\frac{9}{10}$ (respectivement $\frac{8}{10}$) en commission Sciences et de $\frac{3}{10}$ (respectivement $\frac{2}{10}$) en commission Humanités. La probabilité qu'une femme soit recrutée est donc d'après la formule des probabilités totales

$$\frac{9}{10} \times \frac{1}{4} + \frac{3}{10} \times \frac{3}{4} = \frac{3}{10}.$$

De même, la probabilité qu'un homme soit recruté est

$$\frac{8}{10}\times\frac{5}{6}+\frac{2}{10}\times\frac{1}{6}=\frac{7}{10}.$$

Les probabilités conditionnelles sont favorables aux femmes ; les probabilités aux hommes.

2.3. Formule de Bayes

Proposition 24.17. Formule de Bayes

Soit $(\Omega,\mathscr{P}(\Omega),\mathbf{P})$ un espace probabilisé fini, A et $B\subset\Omega$ de probabilités non nulles. Alors,

$$\mathbf{P}(A|B)=\frac{\mathbf{P}(B|A)\mathbf{P}(A)}{\mathbf{P}(B)}.$$

Démonstration

Par définition des probabilités conditionnelles sachant A et sachant B,

$$\mathbf{P}(A\cap B)=\mathbf{P}(A|B)\mathbf{P}(B)=\mathbf{P}(B|A)\mathbf{P}(A).$$

■

Exemple

Étudions la population de deux lycées et sa réussite à un concours :

- d'un côté le lycée privé Sainte-Guenièvre compte 30 % des élèves et a un taux de réussite de 90 %,
- de l'autre, le lycée Louis-le-Minuscule a un taux de réussite de 50 % (pour 70 % des élèves de l'étude donc).

Déterminons la probabilité qu'un élève ayant réussi le concours provienne du lycée privé Sainte-Guenièvre. Pour cela, définissons l'espace probabilisé où l'univers Ω est l'ensemble des élèves et la probabilité $\mathbf{P}$ est uniforme sur cet ensemble. Considérons les événements suivants

- S l'ensemble des élèves du lycée privé Sainte-Guenièvre,
- R l'ensemble des élèves ayant réussi le concours.

Les hypothèses de l'énoncé sont donc $\mathbf{P}(S)=0,3$, $\mathbf{P}(R|S)=0,9$ et $\mathbf{P}(R|\overline{S})=0,5$ et on cherche $\mathbf{P}(S|R)$. Remarquons, tout d'abord, que l'on peut calculer $\mathbf{P}(R)$ avec la formule des probabilités totales associée à $(S,\overline{S})$:

$$\mathbf{P}(R)=\mathbf{P}(R|S)\mathbf{P}(S)+\mathbf{P}(R|\overline{S})\mathbf{P}(\overline{S})=0,9\times0,3+0,5\times0,7=0,62.$$

Ensuite, d'après la formule de Bayes,

$$\mathbf{P}(S|R)=\frac{\mathbf{P}(R|S)\mathbf{P}(S)}{\mathbf{P}(R)}=\frac{0,9\times0,3}{0,62}\simeq0,435.$$

En conclusion, 43,5 % (environ) des élèves ayant réussi le concours proviennent du lycée privé Sainte-Guenièvre.

Exemple

Le test d'un dépistage d'une maladie courante (un pour cent de la population) se doit d'être bon marché ; la fiabilité est alors mise en cause dans deux cas :

- 5 % des patients sains sont déclarés malades,
- 2 % des patients malades sont déclarés sains.

Déterminons la probabilité d'être sain alors que le test est positif (c'est-à-dire le risque de faux positif). Notons T l'événement « le test est positif » et S l'événement « le patient est sain ». Alors,

$$\mathbf{P}(S|T) = \frac{\mathbf{P}(T|S)\mathbf{P}(S)}{\mathbf{P}(T|S)\mathbf{P}(S)+\mathbf{P}(T|\overline{S})\mathbf{P}(\overline{S})} = \frac{0,05\times 0,99}{0,05\times 0,99+0,98\times 0,01} \simeq 0,835.$$

En conclusion, 16,5% (environ) des patients déclarés malades sont sains.

Voici la forme exacte de la formule de Bayes que l'on applique si on effectue simultanément les deux étapes de l'exemple sur les lycées.

Proposition 24.18.

Soit $(\Omega, \mathscr{P}(\Omega), \mathbf{P})$ un espace probabilisé fini, $B_1, \ldots, B_n$ un système complet d'événements de probabilités non nulles. Alors, pour tout $A \subset \Omega$ de probabilité non nulle,

$$\forall i \in [\![1,n]\!], \qquad \mathbf{P}(B_i|A) = \frac{\mathbf{P}(A|B_i)\mathbf{P}(B_i)}{\sum_{k=1}^{n} \mathbf{P}(A|B_k)\mathbf{P}(B_k)}.$$

Démonstration

Il suffit de combiner la formule précédente et la formule des probabilités totales pour le calcul de la probabilité $\mathbf{P}(A)$ qui apparaît au dénominateur. ■

Remarque

Cette formule est un peu contre-intuitive mais admet une interprétation redoutablement efficace.

Si on connaît la probabilité de réalisation d'un événement A sous différentes « causes » possibles B_i, alors on peut déterminer, en observant la réalisation de A, la probabilité pour que chaque B_i en soit la cause.

3. Indépendance

3.1. Indépendance deux à deux

Définition 24.19.

Soit $(\Omega, \mathscr{P}(\Omega), \mathbf{P})$ un espace probabilisé fini.
▹ Deux événements A, $B \subset \Omega$ sont indépendants si $\mathbf{P}(A\cap B) = \mathbf{P}(A)\mathbf{P}(B)$.
▹ Des événements $(A_k)_{k\in[1,n]}$ sont indépendants deux à deux si, pour tous $i \neq j \in [\![1,n]\!]$, A_i et A_j sont indépendants.

Exemples

▹ Un événement A est indépendant de lui-même si, et seulement si, $\mathbf{P}(A) = \mathbf{P}(A\cap A) = \mathbf{P}(A)^2$, c'est-à-dire si $\mathbf{P}(A) \in \{0,1\}$.
▹ Deux événements disjoints A et B sont indépendants si, et seulement si,

$$0 = \mathbf{P}(\varnothing) = \mathbf{P}(A\cap B) = \mathbf{P}(A)\mathbf{P}(B), \quad \text{c'est-à-dire si} \mathbf{P}(A) = 0 \text{ ou } P(B) = 0.$$

Exemple

Avec $\Omega = \{1,2,3,4\}$ muni de la probabilité uniforme $\mathbf{P}$, les événements $A = \{1,2\}$ et $B = \{1,3\}$ sont indépendants car $\mathbf{P}(A\cap B) = \frac{1}{4} = \frac{1}{2}\cdot\frac{1}{2} = \mathbf{P}(A)\mathbf{P}(B)$.

Remarque
L'indépendance est souvent une hypothèse implicite dans les situations de modélisation : deux lancers successifs d'un dé sont indépendants, deux tirages (avec remises) dans une urne sont indépendants ; cela signifie qu'un événement relatif au premier lancer/tirage est indépendant d'un événement relatif au second lancer/tirage.

Proposition 24.20.

Soit $(\Omega, \mathscr{P}(\Omega), \mathbf{P})$ un espace probabilisé fini, $A \subset \Omega$ tel que $\mathbf{P}(A) > 0$ et $B \subset \Omega$. Alors, les événements A et B sont indépendants si, et seulement si, $\mathbf{P}(B|A) = \mathbf{P}(B)$.

Démonstration
Il suffit de remarquer que la condition $\mathbf{P}(B|A) = \mathbf{P}(B)$ c'est-à-dire $\frac{\mathbf{P}(A\cap B)}{\mathbf{P}(A)} = \mathbf{P}(B)$ équivaut à la condition d'indépendance $\mathbf{P}(A \cap B) = \mathbf{P}(A)\mathbf{P}(B)$. ■

Proposition 24.21.

Soit $(\Omega, \mathscr{P}(\Omega), \mathbf{P})$ un espace probabilisé fini et A, B des événements indépendants. Alors, les événements A et $\overline{B}$ sont indépendants.

Démonstration
Il suffit de calculer avec les propriétés de la probabilité $\mathbf{P}$:

$$\mathbf{P}(A \cap \overline{B}) = \mathbf{P}(A) - P(A \cap B) = \mathbf{P}(A) - \mathbf{P}(A)\mathbf{P}(B) = \mathbf{P}(A)(1 - \mathbf{P}(B)) = \mathbf{P}(A)\mathbf{P}(\overline{B}).$$

■

En appliquant la proposition précédente à deux reprises, on obtient le corollaire suivant.

Corollaire 24.22.

Soit $(\Omega, \mathscr{P}(\Omega), \mathbf{P})$ un espace probabilisé fini et A, B des événements indépendants. Alors, les événements $\overline{A}$ et $\overline{B}$ sont indépendants.

3.2. Indépendance mutuelle

Définition 24.23.

Soit $(\Omega, \mathscr{P}(\Omega), \mathbf{P})$ un espace probabilisé fini. Des événements $A_1, \ldots, A_n$ sont (mutuellement) indépendants si, pour toute partie $I \subset [\![1, n]\!]$,

$$\mathbf{P}\Big(\bigcap_{k\in I} A_k\Big) = \prod_{k\in I} \mathbf{P}(A_k).$$

Remarque
On vérifie immédiatement que des événements mutuellement indépendants sont indépendants deux à deux en considérant les parties I de cardinal 2.
La réciproque est fausse. Considérons $\Omega = \{0,1\}^2$ muni de la probabilité uniforme et les événements

$$A = \{(0,0),(0,1)\}, \qquad B = \{(0,0),(1,0)\}, \qquad C = \{(0,0),(1,1)\}.$$

- Les intersections $A\cap B = A\cap C = B\cap C = \{(0,0)\}$ sont de probabilités $\frac{1}{4}$ et les événements A, B et C sont de probabilité $\frac{1}{2}$: les événements A, B et C sont deux à deux indépendants.
- L'intersection $A\cap B\cap C = \{(0,0)\}$ est de probabilité $\frac{1}{4}$ alors que $\mathbf{P}(A)\mathbf{P}(B)\mathbf{P}(C) = \frac{1}{8}$ donc $\mathbf{P}(A)\mathbf{P}(B)\mathbf{P}(C) \neq \mathbf{P}(A\cap B\cap C)$: les événements A, B et C ne sont pas mutuellement indépendants.

Exemple

Considérons n lancers successifs d'un dé et, pour tout $k \in [\![1,n]\!]$, A_k l'événement « le résultat du $k^{\text{ième}}$ lancer est 6 ». Les événements $A_1,\dots, A_n$ sont indépendants.

Proposition 24.24.

Soit $(\Omega, \mathscr{P}(\Omega), \mathbf{P})$ un espace probabilisé fini et $A_1, \dots, A_n \subset \Omega$ des événements mutuellement indépendants, $I \subset [\![1,n]\!]$. Alors, les événements $\cap_{k\in I} A_k$ et $\cap_{k\notin I} A_k$ sont indépendants.

Démonstration

$$\begin{aligned}\mathbf{P}\Big(\bigcap_{k\in I} A_k \cap \bigcap_{k\notin I} A_k\Big) &= \mathbf{P}\Big(\bigcap_{k=1}^{n} A_k\Big) = \prod_{k=1}^{n} \mathbf{P}(A_k) && \text{par indépendance}\\ &= \Big(\prod_{k\in I}\mathbf{P}(A_k)\Big)\Big(\prod_{k\notin I}\mathbf{P}(A_k)\Big)\\ &= \mathbf{P}\Big(\bigcap_{k\in I} A_k\Big)\mathbf{P}\Big(\bigcap_{k\notin I} A_k\Big).\end{aligned}$$

■

Donnons une caractérisation de l'indépendance. Ce résultat peu pratique sert ici à comprendre la différence avec l'indépendance deux à deux.

Proposition 24.25.

Soit $(\Omega, \mathscr{P}(\Omega), \mathbf{P})$ un espace probabilisé fini. Des événements $A_1, \dots, A_n \subset \Omega$ sont indépendants si, et seulement si, pour toute partie $I \subset [\![1,n]\!]$ telle que $\mathbf{P}\Big(\bigcap_{k\in I} A_k\Big) > 0$,

$$\forall i \notin I, \qquad \mathbf{P}\Big(A_i \,\Big|\, \bigcap_{k\in I} A_k\Big) = \mathbf{P}(A_i).$$

Démonstration

($\Rightarrow$) Soit I une telle partie et $i \notin I$. D'après la proposition précédente, les événements A_i et $\cap_{k\in I} A_k$ sont indépendants d'où le résultat.

($\Leftarrow$) Il suffit d'appliquer la formule de la formule des probabilités composées : en effet, pour toute partie $I = \{i_1,\dots,i_p\} \subset [\![1,n]\!]$, on obtient

$$\begin{aligned}\mathbf{P}\Big(\bigcap_{i\in I} A_i\Big) &= \mathbf{P}\Big(\bigcap_{k=1}^{p} A_{i_k}\Big)\\ &= \prod_{k=1}^{p} \mathbf{P}\Big(A_{i_k} \,\Big|\, \bigcap_{j=1}^{k-1} A_{i_j}\Big) && \text{d'après la formule des probabilités composées}\\ &= \prod_{k=1}^{p} \mathbf{P}(A_{i_k}) = \prod_{i\in I} \mathbf{P}(A_i) && \text{d'après l'hypothèse du sens retour}\end{aligned}$$

ce qui est bien la définition d'indépendance. ■

FICHE

SYNTHÈSE

On définit le cadre probabiliste à partir de la notion abstraite d'espace probabilisé.

Espaces probabilisés finis

Un espace probabilisé fini est un triplet $(\Omega, \mathscr{P}(\Omega), \mathbf{P})$ avec Ω un ensemble fini et $\mathbf{P} : \mathscr{P}(\Omega) \to [0,1]$ une application telle que

- $\mathbf{P}(\Omega) = 1$,
- pour toutes parties A, $B \subset \Omega$ disjointes, $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B)$.

Les éléments de Ω sont les résultats, les parties de Ω les événements.

Une fois la phase de modélisation terminée, on se retrouve dans le cadre de la théorie des probabilités avec notamment les règles de calcul suivantes.

Règles de calcul

Soit $(\Omega, \mathscr{P}(\Omega), \mathbf{P})$ un espace probabilisé fini. Alors,

▹ pour tout $A \subset \Omega$, $\mathbf{P}(\overline{A}) = 1 - \mathbf{P}(A)$;
▹ pour tous A, $B \subset \Omega$ tels que $A \subset B$, $\mathbf{P}(A) \leq \mathbf{P}(B)$;
▹ pour tous A, $B \subset \Omega$, $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B)$.

Pour gérer l'apport d'information dans les situations modélisées, on introduit la notion de conditionnement.

Probabilité conditionnelle

Soit $(\Omega, \mathscr{P}(\Omega), \mathbf{P})$ un espace probabilisé fini et $A \subset \Omega$ de probabilité non nulle. La probabilité conditionnelle sachant A est

$$\mathbf{P}_A : \left\{ \begin{array}{ccc} \mathscr{P}(\Omega) & \to & [0,1] \\ B & \mapsto & \mathbf{P}(B|A) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(A)} \end{array} \right.$$

Avec les probabilités conditionnelles, on obtient différentes formules pour le calcul de probabilités.

Formules

▹ **Formule des probabilités composées.** Soit $(\Omega, \mathscr{P}(\Omega), \mathbf{P})$ un espace probabilisé fini et $(A_k)_{k\leq n}$ une famille d'événements vérifiant

$$\mathbf{P}\Big(\bigcap_{k=1}^{n} A_k\Big) > 0.$$

Alors,

$$\mathbf{P}\Big(\bigcap_{k=1}^{n} A_k\Big) = \mathbf{P}\Big(A_n \Big| \bigcap_{k=1}^{n-1} A_k\Big)\mathbf{P}\Big(A_{n-1}\Big|\bigcap_{k=1}^{n-2} A_k\Big)\dots\mathbf{P}(A_2|A_1)\mathbf{P}(A_1).$$

▹ **Formule des probabilités totales.** Soit $(\Omega, \mathscr{P}(\Omega), \mathbf{P})$ un espace probabilisé fini, $A_1, \dots, A_n$ une partition de Ω d'événements de probabilités non nulles. Alors, pour tout $B \subset \Omega$,

$$\mathbf{P}(B) = \sum_{k=1}^{n} \mathbf{P}(B|A_k)\mathbf{P}(A_k).$$

▹ **Formule de Bayes.** Soit $(\Omega, \mathscr{P}(\Omega), \mathbf{P})$ un espace probabilisé fini, $B_1, \dots, B_n$ un système complet d'événements de probabilités non nulles. Alors, pour tout $A \subset \Omega$ de probabilité non nulle,

$$\forall i \in [\![1, n]\!], \qquad \mathbf{P}(B_i|A) = \frac{\mathbf{P}(A|B_i)\mathbf{P}(B_i)}{\sum_{k=1}^{n} \mathbf{P}(A|B_k)\mathbf{P}(B_k)}.$$

La notion d'indépendance admet deux traductions distinctes et non équivalentes dans la théorie.

Indépendance

▹ Soit $(\Omega, \mathscr{P}(\Omega), \mathbf{P})$ un espace probabilisé fini. Des événements $(A_k)_{k\in[1,n]}$ sont deux à deux indépendants si, pour tous $i \neq j \in [\![1, n]\!]$,

$$\mathbf{P}\Big(A_i \cap A_j\Big) = \mathbf{P}(A_i)\mathbf{P}(A_j).$$

▹ Soit $(\Omega, \mathscr{P}(\Omega), \mathbf{P})$ un espace probabilisé fini. Des événements $(A_k)_{k\in[1,n]}$ sont (mutuellement) indépendants si, pour tout $I \subset [\![1, n]\!]$,

$$\mathbf{P}\Big(\bigcap_{k\in I} A_k\Big) = \prod_{k\in I} \mathbf{P}(A_k).$$

Cette propriété entraîne la précédente.

EXERCICES

Vrai ou faux ?

		Vrai	Faux
a)	Soit $p_1, \dots, p_n \in \mathbb{R}_+$ de somme 1. Il existe une unique probabilité $\mathbf{P}$ sur $\Omega = [\![1, n]\!]$ telle que $\mathbf{P}(\{k\}) = p_k$.	☐	☐
b)	Pour toute probabilité $\mathbf{P}$ sur $\Omega = [\![1, n]\!]$ et tout $k \in [\![1, n]\!]$, $\mathbf{P}(\{k\}) > 0$.	☐	☐
c)	Soit $\mathbf{P}$ une probabilité sur Ω et $A \subset \Omega$. Si $\mathbf{P}(A) = 0$, alors $A = \varnothing$.	☐	☐
d)	Soit A de probabilité non nulle. Alors, $\mathbf{P}(B\|A) = 1$ entraîne $B \subset A$.	☐	☐
e)	Soit A de probabilité non nulle. Alors, pour tout $B \subset \Omega$, $\mathbf{P}(B\|A) \leq \mathbf{P}(B)$.	☐	☐
f)	Soit A de probabilité dans $]0, 1[$. Alors, pour tout $B \subset \Omega$, $\mathbf{P}(B\|A) + \mathbf{P}(B\|\overline{A}) = 1$.	☐	☐
g)	Soit A de probabilité dans $]0, 1[$. Alors, pour tout $B \subset \Omega$, $\mathbf{P}(B\|A) + \mathbf{P}(\overline{B}\|A) = 1$.	☐	☐
h)	Un événement indépendant de son complémentaire est de probabilité 0 ou 1.	☐	☐
i)	Soit A, $B \subset \Omega$. Alors, $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B)$ si, et seulement si, A et B sont indépendants.	☐	☐
j)	Deux événements disjoints sont indépendants.	☐	☐
k)	Deux événements indépendants sont disjoints.	☐	☐
l)	Soit A, B et $C \subset \Omega$ des événements tels que A et B sont indépendants et B et C sont indépendants. Alors, A et C sont indépendants.	☐	☐
m)	Trois événements indépendants sont indépendants deux à deux.	☐	☐
n)	Soit A, B et $C \subset \Omega$ tels que $\mathbf{P}(A \cap B \cap C) = \mathbf{P}(A)\mathbf{P}(B)\mathbf{P}(C)$. Alors, A, B et C sont indépendants.	☐	☐
o)	Soit A, B et $C \subset \Omega$ indépendants. Alors, $\overline{A}$, B et C sont indépendants.	☐	☐
p)	Soit A, B, C et $D \subset \Omega$ tels que A, B, C sont indépendants et B, C, D sont indépendants. Alors, A, B, C et D sont indépendants.	☐	☐
q)	Soit $A_1, \dots, A_n$ des événements indépendants de même probabilité p. Alors, $\mathbf{P}(\cup_{k=1}^{n} A_k) = np$.	☐	☐

Exercices d'application du chapitre 24

●○○ **Exercice 1**

Soit A et B deux événements d'un espace probabilisé tels que $\mathbf{P}(A) = \mathbf{P}(B) = \frac{3}{4}$. Déterminer le meilleur encadrement pour $\mathbf{P}(A \cap B)$.

●●○ **Exercice 2**

Soit A, B et C trois événements d'un espace probabilisé tels que

$$\mathbf{P}(A) = \frac{1}{2}, \qquad \mathbf{P}(B) = \frac{3}{5}, \qquad \mathbf{P}(A \cap B) = \frac{1}{5}, \qquad \mathbf{P}(C|A) = \mathbf{P}(C|B) = \frac{1}{2}.$$

Calculer $\mathbf{P}(C)$ sous l'hypothèse additionnelle que cette valeur est l'inverse d'un entier.

●○○ **Exercice 3**

Soit $(\Omega, \mathscr{P}(\Omega), \mathbf{P})$ un espace probabilisé fini, $A \subset \Omega$, $A_1, \ldots, A_n$ une partition de A d'événements de probabilités non nulles, et $B \subset \Omega$, telle que la probabilité $\mathbf{P}(B|A_k)$ ne dépende pas de k. Montrer que, pour tout $k \in [\![1, n]\!]$, $\mathbf{P}(B|A_k) = \mathbf{P}(B|A)$.

●●● **Exercice 4**

Soit un espace probabilisé $(\Omega, \mathscr{P}(\Omega), \mathbf{P})$ tel que $|\Omega|$ est un nombre premier p et $\mathbf{P}$ est la probabilité uniforme. Montrer que deux événements A et B non triviaux ne peuvent pas être indépendants.

●●● **Exercice 5 Inégalités de Boole-Fréchet**

Soit $A_1, \ldots, A_n$ des événements d'un espace probabilisé fini $(\Omega, \mathscr{P}(\Omega), \mathbf{P})$.

1. Montrer que $\mathbf{P}(A_1 \cap \ldots \cap A_n) \geq \sum_{i=1}^{n} \mathbf{P}(A_i) - (n-1)$.
2. Montrer que $\mathbf{P}(A_1 \cap \ldots \cap A_n) \leq \min_{1 \leq i \leq n} \mathbf{P}(A_i)$. Étudier le cas d'égalité.

●○○ **Exercice 6**

Soit $(\Omega, \mathscr{P}(\Omega), \mathbf{P})$ un espace probabilisé, A et B deux événements tels que $\mathbf{P}(A) > 0$. Montrer que

$$\mathbf{P}(A \cap B | A \cup B) \leq \mathbf{P}(A \cap B | A).$$

●○○ **Exercice 7**

Soit $(\Omega, \mathscr{P}(\Omega), \mathbf{P})$ un espace probabilisé fini et A, $B \subset \Omega$. Montrer que

$$|\mathbf{P}(A \cap B) - \mathbf{P}(A)\mathbf{P}(B)| \leq \frac{1}{4}.$$

●●○ **Exercice 8**

Soit $(\Omega, \mathscr{P}(\Omega), \mathbf{P})$ un espace probabilisé fini. Montrer qu'il n'existe pas d'événements $A_1, \ldots, A_n$ indépendants, de réunion Ω et de même probabilité $p < 1$.

●○○ **Exercice 9**

Soit $A_1, \ldots, A_n$ des événements indépendants d'un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$. Montrer que

$$\mathbf{P}\left(\bigcap_{k=1}^{n} \overline{A}_k\right) \leq \exp\left(-\sum_{k=1}^{n} \mathbf{P}(A_k)\right).$$

●●○ **Exercice 10**

Considérons trois tirages indépendants d'une pièce biaisée à Pile ou Face. Déterminer les valeurs de p la probabilité d'obtenir Pile à un lancer pour lesquelles les événements A « les deux premiers lancers donnent le même résultat » et B « les deux derniers lancers donnent le même résultat » sont indépendants.

●●○ **Exercice 11**

Une urne contient n boules blanches et n boules colorées. On tire les boules deux par deux sans remise. Quelle est la probabilité d'obtenir à tous les tirages une boule blanche et une boule colorée ?

●●○ **Exercice 12**

Une personne négligente extrait $2r$ chaussettes d'un tiroir où se trouvent n paires de chaussettes non triées. Quelle est la probabilité de n'obtenir aucune paire de chaussettes assortie ? d'obtenir $k \leq r$ paires assorties ?

●○○ **Exercice 13**

Un couple a deux enfants dont une fille. Quelle est la probabilité que l'autre enfant soit un garçon ?

●●○ **Exercice 14**

On considère un jeu de cartes numérotées initialement de 1 à n. Après avoir mélangé les cartes, quelle est la probabilité que

1. la carte 1 soit plus loin dans le paquet que la carte 2 ?
2. les cartes 1 et 2 soient voisines ?

●●○ **Exercice 15**

Une urne contient $2n$ boules numérotées de 1 à $2n$. On tire toutes les boules successivement et sans remise.

1. Déterminer la probabilité que l'on tire les boules de numéros impairs dans l'ordre croissant, non nécessairement consécutivement (par exemple, avec le tirage suivant pour $n = 8 : 8, 2, 1, 3, 6, 4, 5, 7$).
2. Déterminer la probabilité que l'on tire les boules de numéros impairs dans l'ordre croissant et consécutivement (par exemple, avec le tirage suivant pour $n = 8 : 8, 2, 1, 3, 5, 7, 6, 4$).

●●● **Exercice 16**

Soit $n \in \mathbb{N}^*$ un entier fixé. Une urne contient b boules blanches et c boules colorées ; tirons les boules avec les règles suivantes

- si la boule est blanche, on la retire définitivement ;
- si la boule est colorée, on la replace dans l'urne.

Déterminer la probabilité d'obtenir exactement une boule blanche en n tirages.

●●○ **Exercice 17**

Considérons l'espace probabilisé obtenu avec $\Omega = S_n$ muni de la probabilité uniforme $\mathbf{P}$. Déterminer la probabilité que 1 et n soient dans la même orbite d'une permutation.

Exercices d'approfondissement du chapitre 24

●●● **Exercice A**

Soit $\Omega = [\![1, n]\!]$, $\mathbf{P}$ la probabilité uniforme sur Ω. Notons $\mathscr{P}_n$ l'ensemble des diviseurs premiers de n et définissons, pour tout $p \in \mathscr{P}_n$, $A_p = \{a \in \Omega,\ p|a\}$.

1. Calculer $\mathbf{P}(A_p)$ pour tout $p \in \mathscr{P}_n$.
2. Montrer que les événements A_p pour $p \in \mathscr{P}_n$ sont indépendants.

3. Considérons la fonction d'Euler $\varphi : \mathbb{N}^* \to \mathbb{N}$ telle que, pour tout $n \geq 1$ par $\varphi(n)$ est le nombre d'entiers k de $[\![1, n]\!]$ tels que $k \wedge n = 1$.
   Déduire des questions précédentes la relation
$$\varphi(n) = n \prod_{p \in \mathscr{P}_n} \left(1 - \frac{1}{p}\right).$$

●●● Exercice B Loi de succession de Laplace

Considérons $n+1$ urnes numérotées de 0 à n telle que l'urne numérotée k contienne $n-k$ boules colorées et k boules blanches. On choisit une urne au hasard et on effectue des tirages avec remises au hasard dans cette urne.

1. Déterminer la probabilité p_n que les N premiers tirages amènent des boules blanches. Déterminer la limite de p_n lorsque n tend vers $+\infty$.
2. Déterminer la probabilité q_n que la $(N+1)^{\text{ième}}$ boule tirée soit colorée sachant que les N premières boules tirées étaient blanches. Déterminer la limite de q_n lorsque n tend vers $+\infty$.

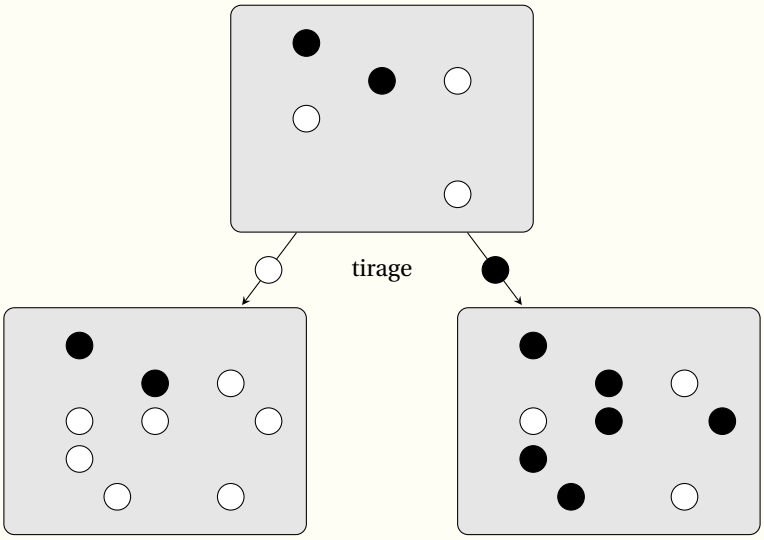
●●○ Exercice C Loi de Hardy-Weinberg

Considérons un gène qui se présente sous deux allèles (c'est-à-dire deux variantes) A et a. Un individu dispose de deux allèles d'un même gène donc AA, Aa ou aa. Un individu reçoit un allèle de chacun de ses parents au hasard. Notons p, q et r les proportions des génotypes dans une génération et P, Q et R les proportions dans la génération suivante. Montrer que $Q^2 = 4PR$.

●●● Exercice D Urne de Polya

On considère une urne contenant a boules colorées et b boules blanches. Après chaque tirage, la boule extraite est remise dans l'urne avec c boules de la même couleur.

Pour $a = 3$, $b = 2$, $c = 4$, on obtient le schéma suivant pour le premier tirage



Déterminer la probabilité que la $n^{\text{ième}}$ boule tirée soit blanche.

CORRIGÉS

Corrigés des Vrai/Faux

a) Vrai.

b) Faux.

c) Faux, cela dépend de la loi de probabilité : sur $\Omega = \{0,1\}$, les conditions $\mathbf{P}(\{0\}) = 1 = 1 - \mathbf{P}(\{1\})$ définissent une probabilité qui ne vérifie pas l'affirmation.

d) Faux en toute généralité, cela entraîne seulement $\mathbf{P}(A \cap B) = \mathbf{P}(A)$. En revanche, si la probabilité de chaque singleton est non nulle, cette condition devient que $A \cap B = A$ donc $B \subset A$.

e) Faux avec, par exemple, $B = A$ de probabilité différente de 1.

f) Faux, par exemple, avec $\Omega = \{1,2,3\}$, $\mathbf{P}$ la probabilité uniforme, $A = \{1,2\}$, $B = \{1,3\}$: $\mathbf{P}(B|A) + \mathbf{P}(B|\overline{A}) = \frac{1}{2} + 1 \neq 1$.

g) Vrai car $\mathbf{P}_A$ est une probabilité et B et $\overline{B}$ sont disjoints.

h) Vrai.

i) Faux, la condition est $\mathbf{P}(A \cap B) = 0$.

j) Faux, il faudrait qu'au moins l'un des événements soit de probabilité nulle.

k) Faux, par exemple, avec $\Omega = \{1,2,3,4\}$, $\mathbf{P}$ la probabilité uniforme, $A = \{1,2\}$, $B = \{1,3\}$: $\mathbf{P}(A \cap B) = \frac{1}{4} = \mathbf{P}(A)\mathbf{P}(B)$.

l) Faux, par exemple avec $A = C$ de probabilité différente de 1 et 0.

m) Vrai.

n) Faux, par exemple, avec $\Omega = [\![1,6]\!]$ muni de la probabilité uniforme : les événements $A = \{1,2,3\}$, $B = \{1,4,5\}$ et $C = \{1,2,3,6\}$ vérifient la condition de l'affirmation et pourtant $\mathbf{P}(A \cap C) \neq \mathbf{P}(A)\mathbf{P}(C)$.

o) Vrai.

p) Faux.

q) Faux.

Corrigés des exercices d'application du chapitre 24

Exercice 1

▹ Commençons par remarquer que $\mathbf{P}(A \cap B) \leq \mathbf{P}(A) = \frac{3}{4}$ et que ce majorant est bien atteint avec $A = B$.

▹ Par ailleurs,

$$\mathbf{P}(A \cap B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cup B) \geq \mathbf{P}(A) + \mathbf{P}(B) - 1 = \frac{1}{2}.$$

Avec la loi uniforme $\mathbf{P}$ sur l'ensemble $\Omega = \{1,2,3,4\}$ et les événements $A = \{1,2,3\}$ et $B = \{1,2,4\}$, on a bien les probabilités réalisant le cas d'égalité.

Exercice 2

▹ Remarquons tout d'abord $\mathbf{P}(A\cap C) = \mathbf{P}(A)\mathbf{P}(C|A) = \frac{1}{4}$ et $\mathbf{P}(B\cap C) = \mathbf{P}(B)\mathbf{P}(C|B) = \frac{3}{10}$.
▹ Par ailleurs, $\mathbf{P}(C) = \mathbf{P}(C\cap A) + \mathbf{P}(C\cap \overline{A}) \leq \mathbf{P}(C\cap A) + \mathbf{P}(\overline{A}) = \frac{1}{4} + (1-\frac{1}{2}) = \frac{3}{4}$. Enfin,

$$\begin{aligned}\mathbf{P}(C) &\geq \mathbf{P}(C\cap(A\cup B)) = \mathbf{P}((A\cap C)\cup(B\cap C))\\ &\geq \mathbf{P}(A\cap C) + \mathbf{P}(B\cap C) - \mathbf{P}(A\cap B\cap C)\\ &\geq \mathbf{P}(A\cap C) + \mathbf{P}(B\cap C) - \mathbf{P}(A\cap B) = \frac{7}{20}.\end{aligned}$$

Ainsi, $\frac{1}{3} < \mathbf{P}(C) < 1$ donc $\mathbf{P}(C) = \frac{1}{2}$ avec l'hypothèse additionnelle.

Exercice 3

Notons $p = \mathbf{P}(B|A_k)$. La formule des probabilités totales donne

$$\mathbf{P}(B|A) = \sum_{k=1}^{n} \mathbf{P}(B|A_k)\mathbf{P}(A_k|A) = p\sum_{k=1}^{n}\mathbf{P}(A_k|A) = p.$$

Exercice 4

Soit A et B deux événements non triviaux de Ω et indépendants. Notons $a = |A|$ et $b = |B|$. Comme la probabilité est uniforme, $\mathbf{P}(A) = \frac{a}{p}$, et $\mathbf{P}(B) = \frac{b}{p}$. Par indépendance,

$$\frac{ab}{p^2} = \mathbf{P}(A)\mathbf{P}(B) = \mathbf{P}(A\cap B) = \frac{|A\cap B|}{p},$$

donc $p|A\cap B| = ab$. D'après le lemme d'Euclide, p divise a ou b ; supposons sans perte de généralité que p divise a. Comme $a \leq p$, $a = 0$ ou $a = p$. Donc A est l'événement certain ou impossible, c'est-à-dire un élément trivial de l'univers.

Exercice 5

1. Procédons par récurrence sur n, le nombre d'événements.
▹ La propriété est vraie au rang 1.
▹ Soit $n \in \mathbb{N}\setminus\{0\}$ tel que pour tout n-uplets d'événements, on ait l'inégalité de l'énoncé. Considérons des événements $A_1, \ldots, A_{n+1}$ et posons $B_n = A_1\cap\ldots\cap A_n$. Alors,

$$\begin{aligned}\mathbf{P}(A_1\cap\ldots\cap A_{n+1}) &= \mathbf{P}(B_n\cap A_{n+1})\\ &= \mathbf{P}(B_n) + \mathbf{P}(A_{n+1}) - \mathbf{P}(B_n\cup A_{n+1})\\ &\geq \mathbf{P}(B_n) + \mathbf{P}(A_{n+1}) - 1\\ &\geq \sum_{i=1}^{n}\mathbf{P}(A_i) - (n-1) + \mathbf{P}(A_{n+1}) - 1\end{aligned}$$

en utilisant l'hypothèse de récurrence, ce qui constitue l'inégalité au rang $n+1$.
2. Comme $A_1\cap\ldots\cap A_n$ est inclus dans chacun des A_i,

$$\mathbf{P}(A_1\cap\ldots\cap A_n) \leq \min_{1\leq i\leq n}\mathbf{P}(A_i).$$

Il y a égalité dès que l'un des A_i est inclus dans tous les autres.

Exercice 6

En utilisant la croissance de la probabilité, on a $\mathbf{P}(A\cup B)\geq \mathbf{P}(A)$ donc

$$\mathbf{P}(A\cap B|A\cup B)=\frac{\mathbf{P}(A\cap B\cap(A\cup B))}{\mathbf{P}(A\cup B)}=\frac{\mathbf{P}(A\cap B)}{\mathbf{P}(A\cup B)}\leq\frac{\mathbf{P}(A\cap B)}{\mathbf{P}(A)}=\mathbf{P}(A\cap B|A).$$

Exercice 7

Notons $a=\mathbf{P}(A\cap\overline{B})$, $b=\mathbf{P}(\overline{A}\cap B)$ et $x=\mathbf{P}(A\cap B)$. Alors,

$$\mathbf{P}(A\cap B)-\mathbf{P}(A)\mathbf{P}(B)=x-(x+a)(x+b)=-ab+x(1-x-a-b).$$

Encadrons

$$\frac{1}{4}\geq(x+a+b)(1-x-a-b)\geq -ab+x(1-x-a-b)\geq -ab\geq -a(1-a)\geq -\frac{1}{4},$$

car pout tout $y\in[0,1]$, $0\leq y(1-y)\leq\frac{1}{4}$, $a\in[0,1]$ et $x+a+b\in[0,1]$.

Exercice 8

Pour tout $i\in[\![1,n]\!]$, l'indépendance mutuelle entraîne l'indépendance entre A_i et $\cap_{\substack{k=1\\k\neq i}}^{n}\overline{A_k}$ donc

$$\mathbf{P}(A_i)=\mathbf{P}\Big(A_i\Big|\bigcap_{\substack{k=1\\k\neq i}}^{n}\overline{A_k}\Big)=1-\mathbf{P}\Big(\overline{A_i}\Big|\bigcap_{\substack{k=1\\k\neq i}}^{n}\overline{A_k}\Big)=1-\frac{\mathbf{P}\Big(\bigcap_{k=1}^{n}\overline{A_k}\Big)}{\mathbf{P}\Big(\bigcap_{\substack{k=1\\k\neq i}}^{n}\overline{A_k}\Big)}=1,$$

car $\cap_{k=1}^{n}\overline{A_k}=\overline{\Omega}=\varnothing$. On contredit ainsi l'hypothèse $p<1$.

Exercice 9

Les événements $A_1,\dots,A_n$ sont indépendants donc leurs complémentaires $\overline{A_1},\dots,\overline{A_n}$ le sont également : ainsi,

$$\begin{aligned}\mathbf{P}\Big(\bigcap_{k=1}^{n}\overline{A}_k\Big)&=\prod_{k=1}^{n}\mathbf{P}(\overline{A}_k)\\&=\prod_{k=1}^{n}\big(1-\mathbf{P}(A_k)\big)\leq\prod_{k=1}^{n}e^{-\mathbf{P}(A_k)}\end{aligned}$$

en utilisant la majoration $1-x\leq e^{-x}$ valable pour tout $x\in\mathbb{R}$. On conclut avec la propriété de morphisme de l'exponentielle.

Exercice 10

Commençons par fixer l'espace probabilisé avec $\Omega=\{P,F\}^3$ et $\mathbf{P}$ la probabilité uniforme. Alors,

$$\begin{aligned}\mathbf{P}(A)&=\mathbf{P}(\{(P,P,P),(P,P,F),(F,F,P),(F,F,F)\})=p^2+(1-p)^2\\\mathbf{P}(B)&=\mathbf{P}(\{(P,P,P),(F,P,P),(P,F,F),(F,F,F)\})=p^2+(1-p)^2\\\mathbf{P}(A\cap B)&=\mathbf{P}(\{(P,P,P),(F,F,F)\})=p^3+(1-p)^3.\end{aligned}$$

Ainsi, $\mathbf{P}(A\cap B)=\mathbf{P}(A)\mathbf{P}(B)$ si, et seulement si, $p^3+(1-p)^3=(p^2+(1-p)^2)^2$. Des solutions évidentes de cette équation sont 0 et 1. En factorisant, on trouve qu'il existe une unique autre solution : $\frac{1}{2}$. Les solutions recherchées sont donc 0, $\frac{1}{2}$ et 1.

Exercice 11

Notons, pour tout $k \leq n$, A_k l'événement « au $k^{\text{ième}}$ tirage » on obtient une boule blanche et une boule colorée.

▹ Soit $k \leq n$. Remarquons, dans un premier temps que si les $k-1$ premiers tirages ont donné une boule blanche et une boule colorée, il reste $n+1-k$ boules de chaque couleur dans l'urne pour le $k^{\text{ième}}$ tirage. Ainsi,

$$\mathbf{P}\Big(A_k \Big| \bigcap_{j=1}^{k-1} A_j\Big) = \frac{(n-k+1)^2}{\binom{2(n-k+1)}{2}} = \frac{n-k+1}{2n-2k+1}.$$

▹ Calculons désormais la probabilité de l'intersection de tous les événements A_k en utilisant la formule des probabilités composées

$$\begin{aligned}\mathbf{P}\Big(\bigcap_{j=1}^{n} A_j\Big) &= \prod_{k=1}^{n} \mathbf{P}\Big(A_k \Big| \bigcap_{j=1}^{k-1} A_j\Big)\\ &= \prod_{k=1}^{n} \frac{n-k+1}{2n-2k+1} = \prod_{j=1}^{n} \frac{j}{2j-1} = \frac{2^n n!^2}{(2n)!}.\end{aligned}$$

Exercice 12

Commençons par fixer l'espace probabilisé avec Ω l'ensemble des tirages possibles de chaussettes dans le tiroir et $\mathbf{P}$ la probabilité uniforme. On est alors ramené à un problème de dénombrement.

▹ Pour n'avoir aucune paire, il faut prendre une chaussette dans $2r$ paires différentes : il y a $\binom{n}{2r}$ choix pour les paires et 2 choix de chaussettes par paire. La probabilité de n'avoir aucune paire est donc

$$\frac{\binom{n}{2r}2^{2r}}{\binom{2n}{2r}}.$$

▹ Pour avoir k paires, il faut prendre k paires complètes puis $2(r-k)$ chaussettes non assorties dans les $n-k$ paires restantes (dont la probabilité a été calculée ci-dessus). La probabilité est donc

$$\frac{\binom{n}{k}\binom{n-k}{2(r-k)}2^{2(r-k)}}{\binom{2n}{2r}}.$$

Exercice 13

Sans hypothèse additionnelle, l'espace probabilisé est $\Omega = \{FF, FG, GF, GG\}$ muni de la probabilité uniforme. Notons $A = \{FF, FG, GF\}$ et $B = \{FG, GF\}$. La probabilité recherchée est alors

$$\mathbf{P}(B|A) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(A)} = \frac{2}{3}.$$

Exercice 14

Une modélisation du problème est l'espace probabilisé obtenu avec $\Omega = S_n$ muni de la probabilité uniforme $\mathbf{P}$.

1. On cherche dans cette première question à dénombrer l'événement $A = \{\sigma \in S_n,\ \sigma(1) > \sigma(2)\}$.

▹ Une première méthode consiste à partitionner A selon la valeur $k = \sigma(1)$: une fois $k \geq 2$ choisi, il reste $k-1$ choix pour $\sigma(2)$ puis une bijection entre $[\![3, n]\!]$ et $[\![1, n]\!] \setminus \{\sigma(1), \sigma(2)\}$. Ainsi,

$$|A| = \sum_{k=2}^{n} (k-1)(n-2)!, \qquad \mathbf{P}(A) = \frac{\sum_{k=2}^{n}(k-1)(n-2)!}{n!} = \frac{n(n-1)(n-2)!}{2n!} = \frac{1}{2}.$$

▹ Une deuxième méthode repose sur le fait que l'application $\sigma \mapsto \sigma \circ (1\ 2)$ réalise une bijection de A dans $\overline{A}$ donc $|A| = |\overline{A}| = \frac{|\Omega|}{2}$. Ainsi, $\mathbf{P}(A) = \frac{1}{2}$.
2. On cherche à dénombrer l'événement $B = \{\sigma \in S_n,\ |\sigma(1) - \sigma(2)| = 1\}$. Pour cela partitionnons selon la valeur de $k = \sigma(1)$ puis celle de $\sigma(2)$ (une possibilité si $k = 1$ ou $k = n$, deux possibilités sinon) et enfin une bijection entre $[\![3, n]\!]$ et $[\![1, n]\!] \setminus \{\sigma(1), \sigma(2)\}$. Ainsi,

$$|B| = (n-2)! + \sum_{k=2}^{n-1} 2(n-2)! + (n-2)! = (n-2)!(2n-2) = 2(n-1)!, \qquad \mathbf{P}(B) = \frac{2}{n}.$$

Exercice 15

Pour cet exercice, l'espace probabilisé naturel est S_{2n} muni de la tribu pleine et de la probabilité uniforme. Il suffit donc de dénombrer les permutations satisfaisant les conditions considérées dans chaque question.
1. Il suffit ici de choisir les indices des tirages des n boules de numéros impairs (le plus petit indice correspondra à la boule 1, le suivant à 3...), soit $\binom{2n}{n}$ possibilités, puis d'attribuer aux n autres tirages les numéros des boules paires correspondantes, soit $n!$ possibilités. La probabilité recherchée est donc

$$\frac{\binom{2n}{n} n!}{(2n)!} = \frac{1}{n!}.$$

Ce résultat est cohérent avec l'intuition suivante : il y a $n!$ façons d'ordonner les boules de numéros impairs dans l'ensemble des tirages et chacune est équiprobable.
2. Il suffit ici de choisir les indices des tirages des n boules de numéros impairs soit $n+1$ possibilités (il suffit de choisir le premier indice entre 1 et $n+1$) puis d'attribuer aux n autres tirages les numéros des boules paires correspondantes, soit $n!$ possibilités. La probabilité recherchée est donc

$$\frac{(n+1)n!}{(2n)!} = \frac{1}{(2n)(2n-1)\dots(n+2)}.$$

Exercice 16

Notons, pour tout $j \in [\![1, n]\!]$, B_j l'événement « la boule obtenue au $j^{\text{ième}}$ tirage est blanche ».
▹ Soit $k \in [\![1, n]\!]$. Calculons la probabilité d'obtenir exactement une boule blanche en n tirages et que cette boule sorte au $k^{\text{ième}}$ tirage. Pour cela utilisons la formule des probabilités totales

$$\begin{aligned}
\mathbf{P}\Big(\bigcap_{j=1}^{k-1} \overline{B_j} \cap B_k \cap \bigcap_{j=k+1}^{n} \overline{B_j}\Big) &= \Big(\prod_{i=1}^{k-1} \mathbf{P}\Big(\overline{B_i} \Big| \bigcap_{j=1}^{i-1} \overline{B_j}\Big)\Big) \times \mathbf{P}\Big(B_k \Big| \bigcap_{j=1}^{k-1} \overline{B_j}\Big) \times \Big(\prod_{i=k+1}^{n} \mathbf{P}\Big(\overline{B_i} \Big| \bigcap_{j=1}^{k-1} \overline{B_j} \cap B_k \cap \bigcap_{j=k+1}^{i-1} \overline{B_j}\Big)\Big) \\
&= \Big(\prod_{i=1}^{k-1} \frac{c}{b+c}\Big) \times \frac{b}{b+c} \times \Big(\prod_{i=k+1}^{n} \frac{c}{b+c-1}\Big) \\
&= \frac{b\,c^{n-1}}{(b+c)^k (b+c-1)^{n-k}}.
\end{aligned}$$

▹ Pour revenir à la question initiale, il suffit de remarquer que l'événement « obtenir exactement une boule blanche en n tirages » est la réunion disjointe des événements dont on vient de calculer la probabilité. Sa probabilité est donc égale à la somme (géométrique) suivante

$$\sum_{k=1}^{n} \frac{b\,c^{n-1}}{(b+c)^k (b+c-1)^{n-k}} = \frac{b\,c^{n-1}}{(b+c-1)^{n-1}} \Big(1 - \Big(\frac{b+c-1}{b+c}\Big)^n\Big).$$

On vérifie rapidement que pour $n = 1$, la probabilité obtenue est bien $\frac{b}{b+c}$.

Exercice 17

Partitionnons cet événement selon la taille l de l'orbite $\mathscr{O}$ commune de 1 et n. Une fois l fixé, il faut choisir les $l-2$ autres éléments de cette orbite (soit $\binom{n-2}{l-2}$ choix), la permutation circulaire de ces l éléments (soit $(l-1)!$ choix) puis une permutation de $[\![1,n]\!]\setminus\mathscr{O}$ (soit $(n-l)!$ choix). La probabilité recherchée est donc (en remplaçant le coefficient binomial par son écriture avec les factorielles)

$$\frac{\sum_{l=2}^{n}\binom{n-2}{l-2}(l-1)!(n-l)!}{n!}=\frac{\sum_{l=2}^{n}(n-2)!(l-1)}{n!}=\frac{1}{2}.$$

Corrigés des exercices d'approfondissement du chapitre 24

Exercice A

1. Soit $p\in\mathscr{P}_n$. Comme $\mathbf{P}$ est la probabilité uniforme, il suffit de dénombrer A_p. Or, il y a $\frac{n}{p}$ multiples de p dans $[\![1,n]\!]$ donc $|A_p|=\frac{n}{p}$ et par conséquent, $\mathbf{P}(A_p)=\frac{1}{p}$.
2. Soit $p_1,\,p_2,\,\ldots,\,p_k\in\mathscr{P}_n$ deux à deux distincts. D'après le lemme d'Euclide, un multiple de $p_1,\,\ldots,\,p_k$ est un multiple de $p_1\ldots p_k$. Traduite en terme d'ensembles, cette propriété devient $A_{p_1}\cap\ldots\cap A_{p_k}=A_{p_1\ldots p_k}$ donc

$$\mathbf{P}(A_{p_1}\cap\ldots\cap A_{p_k})=\mathbf{P}(A_{p_1\ldots p_k})=\frac{1}{p_1\cdots p_k}=\mathbf{P}(A_{p_1})\ldots\mathbf{P}(A_{p_k})$$

en vérifiant que le calcul de la première question reste valable pour tout diviseur de n. En conclusion, les événements A_p pour $p\in\mathscr{P}_n$ sont mutuellement indépendants.
3. Calculons de deux manières la probabilité qu'un entier de Ω soit premier avec n.
▹ D'après la définition de la fonction d'Euler, cette probabilité est $\frac{\varphi(n)}{n}$.
▹ Par ailleurs, un entier de Ω est premier avec n si, et seulement s'il est premier avec tous les facteurs premiers de n, c'est-à-dire qu'il n'est multiple d'aucun diviseur premier de n. La probabilité recherchée est donc, en utilisant l'indépendance,

$$\mathbf{P}\Big(\bigcap_{p\in\mathscr{P}_n}\overline{A_p}\Big)=\prod_{p\in\mathscr{P}_n}\mathbf{P}(\overline{A_p})=\prod_{p\in\mathscr{P}_n}\Big(1-\frac{1}{p}\Big).$$

La formule recherchée est obtenue en identifiant ces deux expressions.

Exercice B

Notons U_k l'événement « la $k^{\text{ième}}$ urne est choisie » et B_i l'événement « la $i^{\text{ème}}$ boule tirée est blanche ».
1. On calcule $\mathbf{P}(B_1\cap\ldots\cap B_N)$ avec la formule des probabilités totales.

$$p_n=\mathbf{P}(B_1\cap\ldots\cap B_N)=\sum_{k=0}^{n}\mathbf{P}(B_1\cap\ldots\cap B_N|U_k)\mathbf{P}(U_k).$$

Le choix de l'urne étant *a priori* uniforme, $\mathbf{P}(U_k)=\frac{1}{n+1}$; par ailleurs, les tirages sont indépendants une fois l'urne choisie. Alors,

$$p_n=\sum_{k=0}^{n}\Big(\frac{k}{n}\Big)^N\frac{1}{n+1}.$$

On reconnaît une quantité proche d'une somme de Riemann

$$p_n = \frac{n}{n+1}.\frac{1}{n}\sum_{k=1}^{n}\left(\frac{k}{n}\right)^N \to \int_0^1 x^N \,\mathrm{d}x = \frac{1}{N+1}.$$

2. On se ramène à la question précédente :

$$q_n = \mathbf{P}(B_{N+1}|B_1 \cap \ldots \cap B_N) = \frac{\mathbf{P}(B_1 \cap \ldots \cap B_N \cap B_{N+1})}{\mathbf{P}(B_1 \cap \ldots \cap B_N)}.$$

Avec la question précédente, on a immédiatement que $\lim q_n = \frac{N+1}{N+2}$.

Exercice C

Pour calculer la probabilité de l'événement B « un individu de la deuxième génération est de génotype AA », utilisons la formule des probabilités totales avec les événements suivants :

- A_1 « les deux parents ont un génotype AA »,
- A_2 « le père est AA et la mère est Aa »,
- A_3 « le père est Aa et la mère est AA »,
- A_4 « les deux parents sont Aa »,
- A_5 « l'un des parents est aa ».

$$\begin{aligned} P = \mathbf{P}(B) &= \sum_{k=1}^{5}\mathbf{P}(B|A_k)\mathbf{P}(A_k) \\ &= \mathbf{P}(A_1) + \frac{1}{2}\mathbf{P}(A_2) + \frac{1}{2}\mathbf{P}(A_3) + \frac{1}{4}\mathbf{P}(A_4) + 0\mathbf{P}(A_5) \\ &= \left(p + \frac{q}{2}\right)^2. \end{aligned}$$

On calcule de même

$$Q = 2\left(p + \frac{q}{2}\right)\left(r + \frac{q}{2}\right), \qquad R = \left(r + \frac{q}{2}\right)^2.$$

Par conséquent, $Q^2 = 4RS$.

Exercice D

Notons B_n l'événement « la $n^{\text{ième}}$ boule tirée est blanche » et $p_n = \mathbf{P}(B_n)$.
Montrons par récurrence que $p_n = \frac{b}{a+b}$.

▹ On a immédiatement $p_1 = \frac{b}{a+b}$.

▹ Calculons p_2 avec la formule des probabilités totales :

$$\begin{aligned} p_2 = \mathbf{P}(B_2) &= \mathbf{P}(B_2|B_1)\mathbf{P}(B_1) + \mathbf{P}(B_2|\overline{B_1})\mathbf{P}(\overline{B_1}) \\ &= \frac{b+c}{a+b+c}p_1 + \frac{b}{a+b+c}(1-p_1) = \frac{b}{a+b}. \end{aligned}$$

▹ Soit $n \geq 2$ tel que $p_k = \frac{b}{a+b}$ pour tout $k \in [\![1, n]\!]$. Considérons les événements A_k « exactement k boules blanches ont été tirées dans les $n-1$ premiers tirages » et exprimons p_n avec la formule des probabilités

CORRIGÉS

totales

$$\begin{aligned}
p_n = \mathbf{P}(B_n) &= \sum_{k=0}^{n-1} \mathbf{P}(B_n|A_k)\mathbf{P}(A_k) \\
&= \sum_{k=0}^{n-1} \frac{b+kc}{a+b+(n-1)c}\mathbf{P}(A_k) \\
&= \frac{1}{a+b+(n-1)c}\sum_{k=0}^{n-1}(b+kc)\mathbf{P}(A_k).
\end{aligned}$$

Par hypothèse de récurrence, $p_n = \frac{b}{a+b}$ donc

$$\sum_{k=0}^{n-1}(b+kc)\mathbf{P}(A_k) = \frac{b(a+b+(n-1)c)}{a+b}.$$

Exploitons cette relation en calculant p_{n+1} avec la formule des probabilités totales :

$$\begin{aligned}
p_{n+1} = \mathbf{P}(B_{n+1}) &= \sum_{k=0}^{n-1}\mathbf{P}(B_{n+1}|B_n\cap A_k)\mathbf{P}(B_n\cap A_k) + \sum_{k=0}^{n-1}\mathbf{P}(B_{n+1}|\overline{B_n}\cap A_k)\mathbf{P}(\overline{B_n}\cap A_k) \\
&= \sum_{k=0}^{n-1}\frac{b+(k+1)c}{a+b+nc}\mathbf{P}(B_n\cap A_k) + \sum_{k=0}^{n-1}\frac{b+kc}{a+b+nc}\mathbf{P}(\overline{B_n}\cap A_k) \\
&= \sum_{k=0}^{n-1}\frac{c}{a+b+nc}\mathbf{P}(B_n\cap A_k) + \sum_{k=0}^{n-1}\frac{b+kc}{a+b+nc}\Big(\mathbf{P}(B_n\cap A_k)+\mathbf{P}(\overline{B_n}\cap A_k)\Big) \\
&= \frac{c}{a+b+nc}\sum_{k=0}^{n-1}\mathbf{P}(B_n\cap A_k) + \sum_{k=0}^{n-1}\frac{b+kc}{a+b+nc}\mathbf{P}(A_k) \\
&= \frac{c}{a+b+nc}\mathbf{P}(B_n) + \frac{1}{a+b+nc}\sum_{k=0}^{n-1}(b+kc)\mathbf{P}(A_k) \\
&= \frac{c}{a+b+nc}\frac{b}{a+b} + \frac{1}{a+b+nc}\frac{b(a+b+(n-1)c)}{a+b} = \frac{b}{a+b}.
\end{aligned}$$

En conclusion, la probabilité de tirer une boule blanche ne dépend pas du numéro du tirage effectué.

Chapitre 25

Variables aléatoires

1. Généralités — **2.** Indépendance de variables aléatoires — **3.** Espérance, moments d'ordre supérieur

Objectifs et compétences du programme

- Calculer la loi d'une variable aléatoire finie.
- Identifier et exploiter une propriété d'indépendance.
- Obtenir espérance ou variance par la définition ou par une décomposition comme somme.
- Savoir utiliser les inégalités en probabilité.

Siméon Denis Poisson, mathématicien français (1781-1840)
Jeune étudiant à l'École polytechnique en 1798, il y sera ensuite professeur. Il s'intéresse aussi bien à la physique qu'aux mathématiques allant de la théorie de l'élasticité (coefficient de Poisson) aux séries de Fourier pour l'équation de la chaleur en passant par la théorie des probabilités. On peut d'ailleurs toujours lire avec intérêt ses « recherches sur la probabilité des jugements en matières criminelles et matière civile ».
Malheureusement pour lui, il laisse aussi un mauvais souvenir à la communauté mathématique comme auteur d'un rapport négatif sur le célèbre mémoire d'Évariste Galois.

COURS

Maintenant que le cadre formel d'espace probabilisé est fixé, la notion de variable aléatoire permet de soutenir l'intuition de fonctions associant des valeurs à une expérience probabiliste. Le « fonctionnement » d'une variable est alors décrit par une mesure de probabilité sur l'ensemble des valeurs prises : sa loi. Ensuite, l'indépendance des variables aléatoires étend la propriété introduite au chapitre précédent pour les événements avec les mêmes avantages dans les calculs. Enfin, parmi les valeurs associées à une loi, on insiste sur les moments (principalement espérance et variance) et leurs utilisations pratiques.

1. Généralités

1.1. Définition des variables aléatoires

Définition 25.1.

Soit $(\Omega, \mathscr{P}(\Omega), \mathbf{P})$ un espace probabilisé fini. Une variable aléatoire sur cet espace et à valeurs dans l'ensemble E est une application $X : \Omega \to E$. Lorsque $E \subset \mathbb{R}$, la variable aléatoire est dite réelle.

Exemple

Considérons l'expérience qui correspond au lancer de deux dés identifiables. Un espace probabilisé adapté est alors défini par $\Omega = [\![1,6]\!]^2$ muni de la probabilité uniforme. Voici quelques exemples de variables aléatoires

- $X_1 : (a,b) \mapsto a$ le résultat du premier dé,
- $X_2 : (a,b) \mapsto b$ le résultat du deuxième dé,
- $X_3 : (a,b) \mapsto a+b$ la somme des résultats,
- $X_4 : (a,b) \mapsto 1 + \mathbb{1}_{a \leq b}$ le numéro du dé qui renvoie le plus grand résultat (et le numéro 2 en cas d'égalité des résultats).

Exemples

▹ Soit $(\Omega, \mathscr{P}(\Omega), \mathbf{P})$ un espace probabilisé fini, X une variable aléatoire à valeurs dans E et $f : E \to F$. Alors, $f(X)$ est une variable aléatoire à valeurs dans F.
▹ Soit $(\Omega, \mathscr{P}(\Omega), \mathbf{P})$ un espace probabilisé fini et $X_1, \ldots, X_p$ des variables aléatoires à valeurs dans $E_1, \ldots, E_p$. Alors, $(X_1, \ldots, X_p)$ est une variable aléatoire à valeurs dans $F = E_1 \times \ldots \times E_p$.

Définition 25.2.

Soit $(\Omega, \mathscr{P}(\Omega), \mathbf{P})$ un espace probabilisé fini et X une variable aléatoire à valeurs dans E. Pour tout $A \subset E$, l'événement noté $\{X \in A\}$, ou plus simplement $X \in A$, est défini par

$$X^{-1}(A) = \{\omega \in \Omega,\ X(\omega) \in A\}.$$

Si $A = \{a\}$, l'événement $X \in A$ est noté $X = a$. Pour une variable aléatoire réelle X et $A = [x, +\infty[$, l'événement $X \in A$ est noté $X \geq x$.

Exemple

Avec les notations du premier exemple de ce chapitre, l'événement « la somme des résultats est paire » est donc noté $X_3 \in 2\mathbb{N}$ ou $X_3 = 0[2]$ en gardant l'esprit de cette notation.

1.2. Loi d'une variable

Proposition 25.3.

Soit $(\Omega, \mathscr{P}(\Omega), \mathbf{P})$ un espace probabilisé fini et X une variable aléatoire à valeurs dans l'ensemble E fini. L'application

$$\mathbf{P}_X : \left\{ \begin{array}{ccc} \mathscr{P}(E) & \to & [0,1] \\ A & \mapsto & \mathbf{P}(X \in A) \end{array} \right.$$

est une probabilité sur E, appelée loi de la variable X.

Démonstration

Vérifions les deux propriétés de probabilité.

▹ Par définition, $\mathbf{P}_X(E) = \mathbf{P}(X \in E) = 1$.

▹ Soit A et $B \subset E$ disjointes. Alors,

$$\mathbf{P}_X(A \cup B) = \mathbf{P}(X \in A \cup B) = \mathbf{P}(X \in A \cup X \in B) = \mathbf{P}(X \in A) + \mathbf{P}(X \in B) = \mathbf{P}_X(A) + \mathbf{P}_X(B),$$

en utilisant que $X \in A$ et $X \in B$ sont des parties disjointes de Ω. ■

Comme une probabilité sur un espace probabilisé fini est uniquement déterminée par l'image des singletons, on obtient le résultat suivant.

Proposition 25.4.

Soit $(\Omega, \mathscr{P}(\Omega), \mathbf{P})$ un espace probabilisé fini et X une variable aléatoire à valeurs dans l'ensemble E. La loi de X est uniquement déterminée par la donnée de $X(\Omega)$ et des valeurs $\mathbf{P}(X = x)$ pour $x \in X(\Omega)$.

Conseils méthodologiques

Lorsque l'on cherche la loi d'une variable aléatoire X, on commence toujours par préciser l'ensemble $X(\Omega)$ des valeurs prises par X puis on calcule les $\mathbf{P}(X = x)$ pour $x \in X(\Omega)$.

Exemple

Soit $(\Omega, \mathscr{P}(\Omega), \mathbf{P})$ un espace probabilisé fini, X une variable aléatoire à valeurs dans E et $f : E \to F$. Alors, la loi de la variable $f(X)$ vérifie, pour tout $B \subset F$,

$$\mathbf{P}_{f(X)}(B) = \mathbf{P}(f(X) \in B) = \mathbf{P}(X \in f^{-1}(B)) = \mathbf{P}_X(f^{-1}(B)).$$

Dorénavant, on suppose implicitement l'existence d'un espace probabilisé fini $(\Omega, \mathscr{P}(\Omega), \mathbf{P})$ sous-jacent.

1.3. Exemples

Définition 25.5.

Une variable X à valeurs dans E suit la loi uniforme si, pour tout $x \in E$, $\mathbf{P}(X = x) = \frac{1}{|E|}$.
On note $X \hookrightarrow \mathscr{U}(E)$.

Exemple

La variable aléatoire correspondant au résultat du lancer d'un dé usuel suit une loi uniforme sur l'ensemble fini $[\![1,6]\!]$.

Définition 25.6.

Soit $p \in [0,1]$. Une variable X à valeurs dans $\{0,1\}$ suit la loi de Bernoulli de paramètre p si

$$\mathbf{P}(X=1) = p = 1 - \mathbf{P}(X=0).$$

On note $X \hookrightarrow \mathscr{B}(p)$.

Exemple

Pour une expérience à deux issues notées 0 et 1, la variable aléatoire qui renvoie le résultat suit une loi de Bernoulli. Par extension, on dit quelquefois que toute variable renvoyant le résultat d'une expérience à deux issues (victoire ou défaite, pile ou face) suit une loi de Bernoulli (en identifiant implicitement les deux issues à 0 et 1).

Remarque

Toute variable aléatoire à valeurs dans $\{0,1\}$ suit une loi de Bernoulli de paramètre sa probabilité de valoir 1.

Proposition 25.7.

Soit $A \subset \Omega$ un événement de probabilité p. Alors, la fonction indicatrice $\mathbb{1}_A$ est une variable aléatoire de loi de Bernoulli de paramètre p.

Démonstration

Il suffit de noter que $\mathbf{P}(\mathbb{1}_A = 1) = \mathbf{P}(A) = p$ et $\mathbf{P}(\mathbb{1}_A = 0) = \mathbf{P}(\overline{A}) = 1 - p$. ■

Définition 25.8.

Soit $n \in \mathbb{N} \setminus \{0\}$ et $p \in [0,1]$. Une variable X à valeurs dans $[\![0,n]\!]$ suit la loi binomiale de paramètres n et p si, pour tout $k \in [\![0,n]\!]$,

$$\mathbf{P}(X=k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

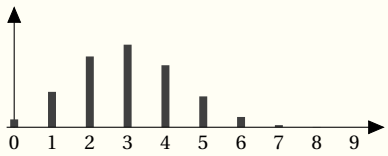
On note $X \hookrightarrow \mathscr{B}(n,p)$.

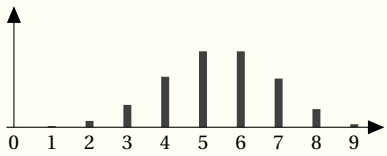
Remarque

Remarquons que cette définition est bien correcte car, avec la formule du binôme,

$$\sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = (p+1-p)^n = 1.$$

Voici les représentations en « histogramme » de la suite finie $(\mathbf{P}(X=k))_k$ lorsque X suit la loi binomiale de paramètres 9 et $\frac{1}{3}$ ou de paramètres 9 et $\frac{6}{10}$.





Remarque
La loi binomiale de paramètres 1 et p n'est autre que la loi de Bernoulli de paramètre p.

Exemple

Considérons une expérience à deux issues menée à n reprises dans les mêmes conditions (lancers de pile ou face, tirages dans une urne avec remises...). Le nombre X d'issues victorieuses définit une variable aléatoire de loi binomiale de paramètres n et p, où p est la probabilité de victoire en une expérience unique. Notons A_i l'événement « victoire à la $i^{\text{ème}}$ expérience » pour $i \in [\![1,n]\!]$. Alors, pour tout $k \in [\![0,n]\!]$,

$$\begin{aligned}
\mathbf{P}(X=k) &= \mathbf{P}\Big(\bigcup_{I\in\mathscr{P}_k([1,n])}\Big(\bigcap_{i\in I}A_i \cap \bigcap_{j\notin I}\overline{A_j}\Big)\Big)\\
&= \sum_{I\in\mathscr{P}_k([1,n])}\mathbf{P}\Big(\bigcap_{i\in I}A_i \cap \bigcap_{j\notin I}\overline{A_j}\Big)\\
&= \sum_{I\in\mathscr{P}_k([1,n])}\Big(\prod_{i\in I}\mathbf{P}(A_i)\Big)\Big(\prod_{j\notin I}\mathbf{P}(\overline{A_j})\Big)\\
&= \sum_{I\in\mathscr{P}_k([1,n])}p^k(1-p)^{n-k}\\
&= \binom{n}{k}p^k(1-p)^{n-k}.
\end{aligned}$$

2. Indépendance de variables aléatoires

2.1. Couples, n-uplets de variables

Définition 25.9.

▹ Soit $X_1, \dots, X_p$ des variables aléatoires à valeurs dans $E_1, \dots, E_p$. La loi jointe de ces variables est la loi de la variable aléatoire $X = (X_1, \dots, X_p)$.
▹ Soit $X = (X_1, \dots, X_p)$ une variable aléatoire à valeurs dans $E_1 \times \dots \times E_p$. Pour tout $k \in [\![1,p]\!]$, la $k^{\text{ième}}$ loi marginale de X est la loi de la variable projetée X_k.

Proposition 25.10.

Soit $X = (X_1, \dots, X_p)$ une variable aléatoire à valeurs dans $E_1 \times \dots \times E_p$ de loi $\mathbf{P}_X$. Alors, la loi marginale de X_1 est donnée par

$$\forall A \subset E_1, \qquad \mathbf{P}_{X_1}(A) = \mathbf{P}_X(A \times E_2 \times \dots \times E_p).$$

Démonstration
Il suffit de remarquer que l'événement $X_1 \in A$ se réécrit $X \in A \times E_2 \times \dots \times E_p$. ■

Remarque

Cette proposition indique que l'on peut retrouver les lois marginales à partir de la loi jointe. La réciproque est fausse : pour le prouver construisons deux vecteurs aléatoires (X_1, X_2) et (Y_1, Y_2) de mêmes lois marginales mais de lois distinctes.

Considérons une urne avec 2 boules numérotées 0 et 2 boules numérotées 1.

▹ Effectuons deux tirages avec remises et notons X_1 et X_2 les variables correspondant aux numéros tirés. Alors, les deux variables suivent des lois de Bernoulli de paramètre $\frac{1}{2}$.

▹ Effectuons deux tirages sans remise et notons Y_1 et Y_2 les variables correspondant aux numéros tirés. La variable Y_1 suit une loi de Bernoulli de paramètre $\frac{1}{2}$ et

$$\mathbf{P}(Y_2 = 1) = \mathbf{P}(Y_2 = 1 | Y_1 = 1)\mathbf{P}(Y_1 = 1) + \mathbf{P}(Y_2 = 1 | Y_1 = 0)\mathbf{P}(Y_1 = 0) = \frac{1}{3}\frac{1}{2} + \frac{2}{3}\frac{1}{2} = \frac{1}{2}.$$

Ainsi, Y_2 suit aussi une loi de Bernoulli de paramètre $\frac{1}{2}$.

En revanche, $\mathbf{P}((X_1, X_2) = (1,1)) = \frac{1}{4} \neq \frac{1}{6} = \mathbf{P}((Y_1, Y_2) = (1,1))$. Ainsi, les deux variables (X_1, X_2) et (Y_1, Y_2) ne suivent pas la même loi alors qu'elles ont les mêmes lois marginales.

On peut aussi déduire de la loi jointe les « lois conditionnelles ».

Proposition 25.11.

Soit $X = (X_1, \ldots, X_p)$ une variable aléatoire à valeurs dans $E_1 \times \ldots \times E_p$ de loi $\mathbf{P}_X$. Alors, pour tout $(A_1, \ldots, A_p) \in \mathscr{P}(E_1) \times \ldots \times \mathscr{P}(E_p)$ tel que $\mathbf{P}(X_2 \in A_2, \ldots, X_p \in A_p) > 0$,

$$\mathbf{P}(X_1 \in A_1 | X_2 \in A_2, \ldots, X_p \in A_p) = \frac{\mathbf{P}_X(A_1 \times A_2 \times \ldots \times A_p)}{\mathbf{P}_X(E_1 \times A_2 \times \ldots \times A_p)}.$$

Démonstration

Il suffit d'utiliser la définition de probabilité conditionnelle et la proposition précédente. ■

2.2. Indépendance

Définition 25.12.

▹ Deux variables X et Y à valeurs dans E et F sont indépendantes si, pour tout $A \subset E$ et tout $B \subset F$,

$$\mathbf{P}(X \in A, Y \in B) = \mathbf{P}(X \in A)\mathbf{P}(Y \in B),$$

c'est-à-dire si, pour tout $A \subset E$ et tout $B \subset F$, les événements $X \in A$ et $Y \in B$ sont indépendants.

▹ Des variables aléatoires $X_1, \ldots, X_p$ à valeurs dans $E_1, \ldots, E_p$ sont mutuellement indépendantes si

$$\forall (A_1, \ldots, A_p) \in \mathscr{P}(E_1) \times \ldots \times \mathscr{P}(E_p), \qquad \mathbf{P}(X_1 \in A_1, \ldots, X_p \in A_p) = \prod_{k=1}^{p} \mathbf{P}(X_k \in A_k),$$

c'est-à-dire, pour tout $(A_1, \ldots, A_p) \in \mathscr{P}(E_1) \times \ldots \times \mathscr{P}(E_p)$, les événements $X_1 \in A_1, \ldots, X_p \in A_p$ sont (mutuellement) indépendants.

Comme pour les événements, l'indépendance (mutuelle) entraîne l'indépendance deux à deux mais la réciproque est fausse.

Remarque

D'après cette définition, il est facile de retrouver la loi jointe à partir des lois marginales pour des variables (mutuellement) indépendantes.

Plus précisément, si $X = (X_1, \dots, X_p)$ avec $X_1, \dots, X_p$ des variables (mutuellement) indépendantes à valeurs dans $E_1, \dots, E_p$, alors, pour tout $(A_1, \dots, A_p) \in \mathscr{P}(E_1) \times \dots \times \mathscr{P}(E_p)$,

$$\mathbf{P}_X(A_1 \times \dots \times A_p) = \prod_{k=1}^{p} \mathbf{P}_{X_k}(A_k).$$

Exemple

Le cadre habituel pour l'obtention de variables aléatoires mutuellement indépendantes est celui où chaque X_i correspond au résultat d'une expérience aléatoire dans une succession d'expériences indépendantes.

Proposition 25.13.

Soit X_1 et X_2 deux variables indépendantes à valeurs dans E_1 et E_2, $f_1 : E_1 \to F_1$ et $f_2 : E_2 \to F_2$. Alors, les variables $f_1(X_1)$ et $f_2(X_2)$ sont indépendantes.

Démonstration

Pour tout $A_1 \subset E_1$ et tout $A_2 \subset E_2$,

$$\begin{aligned}\mathbf{P}(f_1(X_1) \in A_1, f_2(X_2) \in A_2) &= \mathbf{P}(X_1 \in f_1^{-1}(A_1), X_2 \in f_2^{-1}(A_2)) \\ &= \mathbf{P}(X_1 \in f_1^{-1}(A_1))\mathbf{P}(X_2 \in f_2^{-1}(A_2)) \\ &= \mathbf{P}(f_1(X_1) \in A_1)\mathbf{P}(f_2(X_2) \in A_2).\end{aligned}$$

■

De manière plus générale, on a le résultat suivant.

Proposition 25.14. Lemme des coalitions

Soit $X_1, \dots, X_p$ des variables mutuellement indépendantes. Alors, une fonction de $X_1, \dots, X_k$ et une fonction de $X_{k+1}, \dots, X_p$ définissent des variables indépendantes.

Exemple

Lors de 2000 lancers d'une pièce à Pile ou Face, le nombre de Piles obtenu dans les 50 premiers lancers est une variable aléatoire indépendante du nombre de Piles obtenus dans les 1950 lancers suivants.

Exemple

Soit $X_1, \dots, X_n$ des variables mutuellement indépendantes de loi uniforme sur $[\![1, m]\!]$ et $M = \max(X_1, \dots, X_n)$. Calculons la loi de M.

▹ Remarquons tout d'abord que $M(\Omega) = [\![1, m]\!]$.

▹ Soit $k \in [\![1, m]\!]$. Calculons $\mathbf{P}(M \leq k)$:

$$\begin{aligned}\mathbf{P}(M \leq k) &= \mathbf{P}(X_1 \leq k, \dots, X_n \leq k) \\ &= \mathbf{P}(X_1 \leq k) \dots \mathbf{P}(X_n \leq k) = \left(\frac{k}{m}\right)^n\end{aligned}$$

par indépendance.

Déduisons $\mathbf{P}(M=k)$ en remarquant que l'événement $(M \leq k)$ est la réunion disjointe des événements $(M=k)$ et $(M \leq k-1)$

$$\mathbf{P}(M=k)=\mathbf{P}(M \leq k)-\mathbf{P}(M \leq k-1)=\left(\frac{k}{m}\right)^n-\left(\frac{k-1}{m}\right)^n.$$

Conseils méthodologiques

On peut noter comme dans cet exemple que pour déterminer la loi d'une variable X définie comme maximum de variables indépendantes, il est plus facile d'étudier l'événements $(X \leq k)$ (car il se traduit par « toutes les variables sont inférieures à k ») que directement l'événement $(X=k)$. On peut conclure en utilisant que $\mathbf{P}(X=k)=\mathbf{P}(X \leq k)-\mathbf{P}(X \leq k-1)$.
De même, avec une variable X définie comme minimum de variables indépendantes, on privilégiera les événements de la forme $(X \geq k)$ (car il se traduit par « toutes les variables sont supérieures à k »).

Proposition 25.15.

La somme de n variables indépendantes de loi de Bernoulli de paramètre p est une variable aléatoire de loi binomiale de paramètres n et p.

Démonstration

Procédons par récurrence sur n.

▹ Le résultat est déjà établi pour $n=1$.

▹ Soit $n \in \mathbb{N} \setminus \{0\}$. Supposons le résultat établi pour la somme de n variables indépendantes de loi de Bernoulli de paramètre p et considérons $X_1, \ldots, X_{n+1}$ des variables indépendantes de loi de Bernoulli de paramètre p. Alors, pour tout $k \in [\![1, n+1]\!]$,

$$\begin{aligned}
\mathbf{P}\Big(\sum_{j=1}^{n+1} X_j=k\Big) &= \mathbf{P}\Big(\sum_{j=1}^{n+1} X_j=k \,\Big|\, X_{n+1}=1\Big)\mathbf{P}(X_{n+1}=1)+\mathbf{P}\Big(\sum_{j=1}^{n+1} X_j=k \,\Big|\, X_{n+1}=0\Big)\mathbf{P}(X_{n+1}=0) \\
&= p\mathbf{P}\Big(\sum_{j=1}^{n} X_j=k-1\Big)+(1-p)\mathbf{P}\Big(\sum_{j=1}^{n} X_j=k\Big) \\
&= p\binom{n}{k-1}p^{k-1}(1-p)^{n-k+1}+(1-p)\binom{n}{k}p^k(1-p)^{n-k} \\
&= \binom{n+1}{k}p^k(1-p)^{n+1-k}.
\end{aligned}$$

Par ailleurs,

$$\mathbf{P}\Big(\sum_{j=1}^{n+1} X_j=0\Big)=\mathbf{P}(X_1=0,\ldots,X_{n+1}=0)=\prod_{i=1}^{n+1}\mathbf{P}(X_i=0)=(1-p)^{n+1}.$$

En conclusion, $X_1+\ldots+X_{n+1} \hookrightarrow \mathscr{B}(n+1,p)$. ■

3. Espérance, moments d'ordre supérieur

3.1. Espérance

Définition 25.16.

L'espérance d'une variable aléatoire réelle X est le réel

$$\mathbf{E}(X) = \sum_{x \in X(\Omega)} x\mathbf{P}(X = x).$$

Une variable aléatoire réelle est centrée si son espérance est nulle.

Remarque

L'espérance de X est la moyenne des valeurs possibles pour X pondérées par leurs probabilités d'apparaître. Cette quantité dépend de la loi de X mais pas de la réalisation de cette loi (autrement dit deux variables de même loi ont la même espérance).

Exemple

Pour une variable aléatoire X constante, il existe $c \in E$ tel que $\mathbf{P}(X = c) = 1$ donc

$$\mathbf{E}(X) = c\mathbf{P}(X = c) = c.$$

Exemple

▹ Pour une variable aléatoire X qui suit la loi de Bernoulli de paramètre p,

$$\mathbf{E}(X) = 1\mathbf{P}(X = 1) + 0\mathbf{P}(X = 0) = p.$$

▹ Plus généralement, si A est un événement, la variable indicatrice $\mathbb{1}_A$ suit la loi de Bernoulli de paramètre $\mathbf{P}(A)$ donc $\mathbf{E}(\mathbb{1}_A) = \mathbf{P}(A)$.

Exemple

Pour une variable aléatoire X qui suit la loi uniforme sur $[\![1, n]\!]$,

$$\mathbf{E}(X) = \sum_{k=1}^{n} k\mathbf{P}(X = k) = \sum_{k=1}^{n} k\frac{1}{n} = \frac{n+1}{2}.$$

Exemple

Soit X une variable aléatoire à valeurs dans $[\![0, n]\!]$. Réécrivons la formule donnant l'espérance en remarquant que, pour tout $k \in [\![0, n]\!]$, $\mathbf{P}(X = k) = \mathbf{P}(X \geq k) - \mathbf{P}(X \geq k+1)$. Ainsi,

$$\begin{aligned}
\mathbf{E}(X) &= \sum_{k=0}^{n} k(\mathbf{P}(X \geq k) - \mathbf{P}(X \geq k+1)) \\
&= \sum_{k=0}^{n} k\mathbf{P}(X \geq k) - \sum_{k=1}^{n+1} (k-1)\mathbf{P}(X \geq k) \\
&= \sum_{k=1}^{n} \mathbf{P}(X \geq k).
\end{aligned}$$

Cette nouvelle formule est pratique, notamment lorsque l'on considère des variables pour lesquelles les événements $(X \geq k)$ sont de probabilités facilement calculables (par exemple, si X est le minimum de variables indépendantes.

Proposition 25.17.

Soit X et Y deux variables aléatoires réelles, λ, $\mu \in \mathbb{R}$. Alors, $\mathbf{E}(\lambda X + \mu Y) = \lambda \mathbf{E}(X) + \mu \mathbf{E}(Y)$.

Démonstration

Décomposons l'espérance de $\lambda X + \mu Y$ avec la loi de (X, Y) (ce qui revient à décomposer un événement $\lambda X + \mu Y = z$ comme réunion des $(X = x) \cap (Y = y)$ pour les (x, y) tels que $\lambda x + \mu y = z$). Ainsi

$$\begin{aligned}
\mathbf{E}(\lambda X + \mu Y) &= \sum_{(x,y) \in X(\Omega) \times Y(\Omega)} (\lambda x + \mu y) \mathbf{P}(X = x, Y = y) \\
&= \sum_{x \in X(\Omega)} \sum_{y \in Y(\Omega)} (\lambda x + \mu y) \mathbf{P}(X = x, Y = y) \\
&= \lambda \sum_{x \in X(\Omega)} \sum_{y \in Y(\Omega)} x \mathbf{P}(X = x, Y = y) + \mu \sum_{x \in X(\Omega)} \sum_{y \in Y(\Omega)} y \mathbf{P}(X = x, Y = y) \\
&= \lambda \sum_{x \in X(\Omega)} x \sum_{y \in Y(\Omega)} \mathbf{P}(X = x, Y = y) + \mu \sum_{y \in Y(\Omega)} y \sum_{x \in X(\Omega)} \mathbf{P}(X = x, Y = y) \\
&= \lambda \sum_{x \in X(\Omega)} x \mathbf{P}(X = x) + \mu \sum_{y \in Y(\Omega)} y \mathbf{P}(Y = y) = \lambda \mathbf{E}(X) + \mu \mathbf{E}(Y).
\end{aligned}$$

■

Proposition 25.18.

Soit X et Y deux variables aléatoires réelles.
▹ Si X est à valeurs positives, alors $\mathbf{E}(X) \geq 0$.
▹ Si $\mathbf{P}(X \leq Y) = 1$, alors $\mathbf{E}(X) \leq \mathbf{E}(Y)$.

Démonstration

▹ Tous les termes $xP(X = x)$ pour $x \in X(\Omega)$ de la somme définissant $\mathbf{E}(X)$ sont positifs donc $\mathbf{E}(X) \geq 0$.

▹ Il suffit d'appliquer le point précédent en remarquant que la variable $Y - X$ est presque sûrement positive. Ainsi, $\mathbf{E}(Y) - \mathbf{E}(X) = \mathbf{E}(Y - X) \geq 0$. ■

Remarque

Soit X une variable aléatoire réelle positive d'espérance nulle. Comme la somme des $xP(X = x)$ pour $x \in X(\Omega)$ est nulle, tous les termes sont nuls, c'est-à-dire si $\mathbf{P}(X = x) = 0$ dès que $x > 0$ soit $\mathbf{P}(X = 0) = 1$.

Exemple

L'espérance d'une variable aléatoire qui suit la loi binomiale de paramètres n et p est np. En effet, une telle variable peut être obtenue comme somme de n variables $X_1, \ldots, X_n$ indépendantes qui suivent toutes la loi de Bernoulli de paramètre p donc son espérance est

$$\mathbf{E}\Big(\sum_{k=1}^{n} X_k\Big) = \sum_{k=1}^{n} \mathbf{E}(X_k) = \sum_{k=1}^{n} p = np.$$

Proposition 25.19. Inégalité de Markov

Soit X une variable aléatoire réelle positive. Alors, pour tout $a \in \mathbb{R}$,

$$\mathbf{E}(X) \geq a\mathbf{P}(X \geq a).$$

Démonstration
Décomposoons $X = X\mathbb{1}_{X\geq a} + X\mathbb{1}_{X<a}$ et minorons les espérances en utilisant notamment que $\mathbf{E}(X\mathbb{1}_{X<a}) \geq 0$ car la variable X est positive

$$\mathbf{E}(X) = \mathbf{E}(X\mathbb{1}_{X\geq a}) + \mathbf{E}(X\mathbb{1}_{X<a}) \geq \mathbf{E}(X\mathbb{1}_{X\geq a}) \geq a\mathbf{E}(\mathbb{1}_{X\geq a}) = a\mathbf{P}(X \geq a).$$

■

Quitte à diviser par la valeur a (quand celle-ci est strictement positive pour préserver l'inégalité), on obtient la formulation suivante de l'inégalité de Markov.

Corollaire 25.20.

Soit X une variable aléatoire réelle positive. Alors, pour tout $a \in \mathbb{R}_+^*$, $\mathbf{P}(X \geq a) \leq \frac{\mathbf{E}(X)}{a}$.

Mentionnons enfin que l'espérance se comporte bien pour les produits de variables indépendantes.

Proposition 25.21.

Soit $X_1, \ldots, X_p$ des variables mutuellement indépendantes. Alors,

$$\mathbf{E}(X_1 \times \ldots \times X_p) = \mathbf{E}(X_1) \times \ldots \times \mathbf{E}(X_p).$$

Démonstration
Utilisons la définition de l'espérance en se ramenant à la loi de $(X_1, \ldots, X_p)$ puis l'indépendance pour simplifier les probabilités d'intersection

$$\begin{aligned}
\mathbf{E}(X_1 \times \ldots \times X_p) &= \sum_{(x_1,\ldots,x_p)\in X_1(\Omega)\times\ldots\times X_p(\Omega)} x_1 \times \ldots \times x_p \mathbf{P}_X(\{(x_1,\ldots,x_p)\}) \\
&= \sum_{x_1\in X_1(\Omega)} \ldots \sum_{x_p\in X_p(\Omega)} x_1 \times \ldots \times x_p \mathbf{P}_{X_1}(\{x_1\}) \times \ldots \times \mathbf{P}_{X_p}(\{x_p\}) \\
&= \sum_{x_1\in X_1(\Omega)} x_1 \mathbf{P}_{X_1}(\{x_1\}) \times \ldots \times \sum_{x_p\in X_p(\Omega)} x_p \mathbf{P}_{X_p}(\{x_p\}) \\
&= \mathbf{E}(X_1) \times \ldots \times \mathbf{E}(X_p).
\end{aligned}$$

■

On verra plus tard que la réciproque est en général fausse.

3.2. Formule de transfert

Énonçons désormais une formule très pratique pour calculer des espérances de variables définies comme fonction d'une ou plusieurs variables dont on connaît la loi (on a déjà utilisé ce résultat dans des démonstrations de la section précédente).

Proposition 25.22. Formule de transfert

Soit X une variable aléatoire à valeurs dans E et $f : E \to F$. Alors,

$$\mathbf{E}(f(X)) = \sum_{x \in X(\Omega)} f(x)\mathbf{P}(X = x).$$

Démonstration

Exploitons la définition de l'espérance de $f(X)$ puis décomposons chaque événement $f(X) = y$ comme la réunion (disjointe) des événements $X = x$ pour $x \in f^{-1}(\{y\})$:

$$\begin{aligned}
\mathbf{E}(f(X)) &= \sum_{y \in f(X(\Omega))} y\mathbf{P}(f(X) = y) \\
&= \sum_{y \in f(X(\Omega))} y \sum_{\substack{x \in X(\Omega) \\ f(x)=y}} \mathbf{P}(X = x) \\
&= \sum_{y \in f(X(\Omega))} \sum_{\substack{x \in X(\Omega) \\ f(x)=y}} f(x)\mathbf{P}(X = x) = \sum_{x \in X(\Omega)} f(x)\mathbf{P}(X = x)
\end{aligned}$$

en utilisant l'associativité de la somme pour la dernière égalité. ■

Exemple

Soit X une variable de loi binomiale de paramètres n et p. Calculons $\mathbf{E}\left(\frac{1}{X+1}\right)$ avec la formule de transfert.

$$\begin{aligned}
\mathbf{E}\left(\frac{1}{X+1}\right) &= \sum_{k=0}^{n} \frac{1}{k+1}\mathbf{P}(X = k) = \sum_{k=0}^{n} \frac{1}{k+1}\binom{n}{k} p^k (1-p)^{n-k} \\
&= \sum_{k=0}^{n} \frac{1}{n+1}\binom{n+1}{k+1} p^k (1-p)^{n-k} \\
&= \frac{1}{p(n+1)} \sum_{k=0}^{n} \binom{n+1}{k+1} p^{k+1} (1-p)^{n+1-k-1} \\
&= \frac{1}{p(n+1)}\left((p + (1-p))^{n+1} - (1-p)^{n+1}\right) = \frac{1-(1-p)^{n+1}}{p(n+1)}.
\end{aligned}$$

Proposition 25.23. Inégalité de Markov

Soit X une variable aléatoire réelle et $f : \mathbb{R} \to \mathbb{R}_+$ croissante. Alors, pour tout $a \in \mathbb{R}$ tel que $f(a) > 0$,

$$\mathbf{P}(X \geq a) \leq \frac{\mathbf{E}(f(X))}{f(a)}.$$

Démonstration

On remarque que l'événement $X \geq a$ est inclus dans l'événement $f(X) \geq f(a)$ (grâce à la croissance de f) puis on applique l'inégalité de Markov précédente à la variable $f(X)$

$$\mathbf{P}(X \geq a) \leq \mathbf{P}\big(f(X) \geq f(a)\big) \leq \frac{\mathbf{E}(f(X))}{f(a)}.$$

■

3.3. Variance

Définition 25.24.

Soit X une variable aléatoire réelle.
▹ Le moment d'ordre $k \in \mathbb{N}$ de X est le réel $\mathbf{E}(X^k)$.
▹ La variance de X est le réel positif $\mathbf{V}(X) = \mathbf{E}((X - \mathbf{E}(X))^2)$.
▹ L'écart-type de X est $\sigma(X) = \sqrt{\mathbf{V}(X)}$.
Une variable centrée est réduite si sa variance est égale à 1.

Remarque

La variance donne une mesure de la dispersion de la variable (en fait de la loi) car c'est la moyenne pondérée des écarts à la moyenne au carré. On remarque immédiatement que la variance d'une variable certaine est nulle. Réciproquement, pour une variable X de variance nulle, $(X - \mathbf{E}(X))^2$ est certaine égale à 0 donc X est certaine (égale à son espérance).

Conseils méthodologiques

Pour montrer qu'une variable est certaine, il suffit de montrer que sa variance est nulle.

Proposition 25.25. Formule de König-Huygens

Soit X une variable aléatoire réelle. Alors, $\mathbf{V}(X) = \mathbf{E}(X^2) - \mathbf{E}(X)^2$.

Démonstration

Développons le carré dans la définition de variance

$$\begin{aligned}\mathbf{V}(X) &= \mathbf{E}(X^2 - 2\mathbf{E}(X)X + \mathbf{E}(X)^2)\\ &= \mathbf{E}(X^2) - 2\mathbf{E}(X)\mathbf{E}(X) + \mathbf{E}(X)^2\mathbf{E}(1) = \mathbf{E}(X^2) - \mathbf{E}(X)^2.\end{aligned}$$

■

On trouve avec cette expression que $\mathbf{E}(X^2) - \mathbf{E}(X)^2 \geq 0$. De manière plus générale, on obtient l'inégalité de Cauchy-Schwarz.

Proposition 25.26. Inégalité de Cauchy-Schwarz

Soit X et Y deux variables aléatoires réelles. Alors, $\mathbf{E}(XY)^2 \leq \mathbf{E}(X^2)\mathbf{E}(Y^2)$.

Démonstration

Considérons la fonction polynomiale $P : \lambda \mapsto \mathbf{E}((\lambda X + Y)^2)$. Cette fonction est positive (comme espérance d'une variable positive) et s'écrit en développant

$$\forall \lambda \in \mathbb{R}, \qquad P(\lambda) = \mathbf{E}(X^2)\lambda^2 + 2\mathbf{E}(XY)\lambda + \mathbf{E}(Y^2).$$

▹ Si $\mathbf{E}(X^2) = 0$, la fonction P est de degré au plus 1 et de signe constant donc est une constante : ainsi $\mathbf{E}(XY) = 0$ et l'inégalité est vérifiée.
▹ Sinon, P est de degré 2 et de signe constant : son discriminant est négatif soit

$$4\mathbf{E}(XY)^2 - 4\mathbf{E}(X^2)\mathbf{E}(Y^2) \leq 0$$

ce qui est bien l'inégalité attendue. ■

Exemple

Pour une variable aléatoire X constante égale à c, $\mathbf{V}(X) = \mathbf{E}(X^2) - \mathbf{E}(X)^2 = c^2 - c^2 = 0$.

Exemple

Pour une variable aléatoire X qui suit la loi de Bernoulli de paramètre p,

$$\mathbf{V}(X) = \mathbf{E}(X^2) - \mathbf{E}(X)^2 = p - p^2 = p(1-p).$$

Proposition 25.27.

Soit X une variable aléatoire réelle, λ et $\mu \in \mathbb{R}$. Alors, $\mathbf{V}(\lambda X + \mu) = \lambda^2 \mathbf{V}(X)$.

Démonstration

Utilisons la définition de variance

$$\begin{aligned}\mathbf{V}(\lambda X + \mu) &= \mathbf{E}((\lambda X + \mu - \mathbf{E}(\lambda X + \mu))^2)\\ &= \mathbf{E}((\lambda X - \lambda \mathbf{E}(X))^2)\\ &= \lambda^2 \mathbf{E}((X - \mathbf{E}(X))^2) = \lambda^2 \mathbf{V}(X).\end{aligned}$$

■

Remarque

Soit X une variable aléatoire réelle non constante. Alors, la variable $Y = \frac{X - \mathbf{E}(X)}{\sigma(X)}$ est centrée et réduite. En effet,

$$\mathbf{E}(Y) = \frac{1}{\sigma(X)}(\mathbf{E}(X) - \mathbf{E}(\mathbf{E}(X))) = 0, \qquad \mathbf{V}(Y) = \frac{1}{\sigma(X)^2}\mathbf{V}(X) = 1.$$

Proposition 25.28. Inégalité de Bienaymé-Tchebychev

Soit X une variable aléatoire réelle. Alors, pour tout $a \in \mathbb{R}_+^*$,

$$\mathbf{P}(|X - \mathbf{E}(X)| \geq a) \leq \frac{\mathbf{V}(X)}{a^2}.$$

Démonstration

Il suffit d'appliquer l'inégalité de Markov avec la fonction $f : x \mapsto x^2$ croissante sur $\mathbb{R}_+$ et la variable aléatoire positive $|X - \mathbf{E}(X)|$. ■

Proposition 25.29.

La variance d'une somme de variables aléatoires indépendantes est la somme des variances de ces variables.

Démonstration

Il suffit de la montrer pour $n = 2$ variables puis d'effectuer une récurrence immédiate.

Soit X et Y indépendantes.

$$\begin{aligned}\mathbf{V}(X+Y) &= \mathbf{E}((X+Y)^2)-(\mathbf{E}(X)+\mathbf{E}(Y))^2\\ &= \mathbf{E}(X^2)+2\mathbf{E}(XY)+\mathbf{E}(Y^2)-\mathbf{E}(X)^2-2\mathbf{E}(X)\mathbf{E}(Y)-\mathbf{E}(Y)^2\\ &= \mathbf{V}(X)+\mathbf{V}(Y)+2(\mathbf{E}(XY)-\mathbf{E}(X)\mathbf{E}(Y))\\ &= \mathbf{V}(X)+\mathbf{V}(Y),\end{aligned}$$

en utilisant que l'espérance d'un produit de variables indépendantes est le produit des espérances de ces variables. ■

Exemple

La variance d'une variable aléatoire qui suit la loi binomiale de paramètres n et p est la somme des variances de n variables indépendantes suivant la loi de Bernoulli de paramètre p donc est égale à $np(1-p)$.

3.4. Covariance

Définition 25.30.

Soit X et Y deux variables aléatoires réelles. La covariance de X et Y est le réel

$$\mathbf{Cov}(X,Y)=\mathbf{E}((X-\mathbf{E}(X))(Y-\mathbf{E}(Y))).$$

Les variables X et Y son non corrélées si $\mathbf{Cov}(X,Y)=0$.

La covariance de X et de X n'est autre que la variance de X. Ceci permet d'obtenir l'identité remarquable suivante.

Proposition 25.31. Identité remarquable — Identité de polarisation

Soit X et Y deux variables aléatoires réelles. Alors,

$$\mathbf{V}(X+Y)=\mathbf{V}(X)+2\mathbf{Cov}(X,Y)+\mathbf{V}(Y),$$

ou, de manière équivalente,

$$\mathbf{Cov}(X,Y)=\frac{1}{2}\Big(\mathbf{V}(X+Y)-\mathbf{V}(X)-\mathbf{V}(Y)\Big).$$

Démonstration

Il suffit de développer le carré dans l'espérance définissant la variance de $X+Y$

$$\begin{aligned}\mathbf{V}(X+Y) &= \mathbf{E}((X+Y-\mathbf{E}(X+Y))^2)\\ &= \mathbf{E}((X-\mathbf{E}(X)+Y-\mathbf{E}(Y))^2)\\ &= \mathbf{E}((X-\mathbf{E}(X))^2)+2\mathbf{E}((X-\mathbf{E}(X))(Y-\mathbf{E}(Y)))+\mathbf{E}((Y-\mathbf{E}(Y))^2)\\ &= \mathbf{V}(X)+2\mathbf{Cov}(X,Y)+\mathbf{V}(Y).\end{aligned}$$

■

Proposition 25.32.

Soit X et Y deux variables aléatoires réelles. La covariance de X et Y est le réel

$$\mathbf{Cov}(X,Y)=\mathbf{E}(XY)-\mathbf{E}(X)\mathbf{E}(Y).$$

Démonstration

Avec l'identité de polarisation ci-dessus et la formule de König-Huygens, on obtient

$$\begin{aligned}\mathbf{Cov}(X,Y)&=\tfrac{1}{2}\Big(\mathbf{V}(X+Y)-\mathbf{V}(X)-\mathbf{V}(Y)\Big)\\&=\tfrac{1}{2}\Big(\mathbf{E}((X+Y)^2)-\mathbf{E}(X+Y)^2-\mathbf{E}(X^2)+\mathbf{E}(X)^2-\mathbf{E}(Y^2)+\mathbf{E}(Y)^2\Big)\\&=\mathbf{E}(XY)-\mathbf{E}(X)\mathbf{E}(Y).\end{aligned}$$

■

Proposition 25.33.

Deux variables indépendantes sont non corrélées.

Démonstration

Soit X et Y des variables indépendantes. Alors,

$$\mathbf{Cov}(X,Y)=\mathbf{E}((X-\mathbf{E}(X))(Y-\mathbf{E}(Y)))=\mathbf{E}(X-\mathbf{E}(X))\mathbf{E}(Y-\mathbf{E}(Y))=0,$$

car les variables $X-\mathbf{E}(X)$ et $Y-\mathbf{E}(Y)$ sont centrées. ■

En revanche, la réciproque est fausse comme on peut le voir avec l'exemple suivant.

Exemple

Considérons X de loi uniforme sur $\{-1,0,1\}$ et Y la fonction indicatrice de l'événement $X=0$.
▹ Comme $XY=0$ (Y est non nulle si, et seulement si, X est nulle), $\mathbf{E}(XY)=0$. Par ailleurs, X est une variable centrée donc

$$\mathbf{Cov}(X,Y)=\mathbf{E}(XY)-\mathbf{E}(X)\mathbf{E}(Y)=0.$$

Les variables X et Y ne sont pas corrélées.
▹ Comme $XY=0$, $\mathbf{P}(X=1,Y=1)=0$. Or $\mathbf{P}(X=1)=\frac{1}{3}$ et $\mathbf{P}(Y=1)=\mathbf{P}(X=0)=\frac{1}{3}$ donc

$$\mathbf{P}(X=1,Y=1)\neq\mathbf{P}(X=1)\mathbf{P}(Y=1),$$

les variables X et Y ne sont pas indépendantes.

Remarque

Pour étudier la corrélation de variables X et Y à valeurs dans E et F, on définit quelquefois le coefficient de corrélation de X et Y par

$$\rho(X,Y)=\frac{\mathbf{Cov}(X,Y)}{\sigma(X)\sigma(Y)}.$$

D'après l'inégalité de Cauchy-Schwarz, $\mathbf{Cov}(X,Y)^2\leq\mathbf{V}(X)\mathbf{V}(Y)$ donc $|\rho(X,Y)|\leq 1$.

FICHE

SYNTHÈSE

Ce chapitre permet d'étudier les résultats d'une expérience aléatoire en introduisant le concept de variable aléatoire.

Loi

▹ Soit $(\Omega, \mathscr{P}(\Omega), \mathbf{P})$ un espace probabilisé fini. Une variable aléatoire à valeurs dans E est une application de Ω dans E.
▹ Soit $(\Omega, \mathscr{P}(\Omega), \mathbf{P})$ un espace probabilisé fini et X une variable aléatoire à valeurs dans l'ensemble E fini. L'application

$$\mathbf{P}_X : \left\{ \begin{array}{ccc} \mathscr{P}(E) & \longrightarrow & [0,1] \\ A & \longmapsto & \mathbf{P}(X \in A) \end{array} \right.$$

est une probabilité sur E, appelée loi de la variable X.

Comme pour les événements, on dispose de la notion d'indépendance (en fait de deux notions d'indépendance).

Indépendance

▹ Deux variables X et Y à valeurs dans E et F sont indépendantes si, pour tout $A \subset E$ et tout $B \subset F$,

$$\mathbf{P}(X \in A, Y \in B) = \mathbf{P}(X \in A)\mathbf{P}(Y \in B),$$

c'est-à-dire si, pour tout $A \subset E$ et tout $B \subset F$, les événements $X \in A$ et $Y \in B$ sont indépendants.
▹ Des variables aléatoires $X_1, \ldots, X_p$ à valeurs dans $E_1, \ldots, E_p$ sont mutuellement indépendantes si

$$\forall (A_1, \ldots, A_p) \in \mathscr{P}(E_1) \times \ldots \times \mathscr{P}(E_p), \qquad \mathbf{P}(X_1 \in A_1, \ldots, X_p \in A_p) = \prod_{k=1}^{p} \mathbf{P}(X_k \in A_k),$$

c'est-à-dire, pour tout $(A_1, \ldots, A_p) \in \mathscr{P}(E_1) \times \ldots \times \mathscr{P}(E_p)$, les événements $X_1 \in A_1, \ldots, X_p \in A_p$ sont (mutuellement) indépendants.

L'espérance est un paramètre associé à la loi d'une variable aléatoire réelle qui indique la moyenne des valeurs prises par cette variable.

Espérance

▹ L'espérance d'une variable aléatoire X à valeurs dans $E \subset \mathbb{R}$ est le réel

$$\mathbf{E}(X) = \sum_{x \in E} x\mathbf{P}(X = x).$$

▹ (Formule de transfert) Pour toute application $f : E \longrightarrow F$, $\mathbf{E}(f(X)) = \sum_{x \in E} f(x)\mathbf{P}(X = x)$.
▹ Soit $X_1, \ldots, X_p$ des variables mutuellement indépendantes. Alors, $\mathbf{E}(X_1 \times \ldots \times X_p) = \mathbf{E}(X_1) \times \ldots \times \mathbf{E}(X_p)$.

La variance est un paramètre associé à la loi d'une variable aléatoire réelle qui indique la dispersion de

celle-ci.

Variance

▹ La variance d'une variable aléatoire réelle X est le réel positif

$$\mathbf{V}(X) = \mathbf{E}((X - \mathbf{E}(X))^2) = \mathbf{E}(X^2) - \mathbf{E}(X)^2.$$

L'écart-type de X est $\sigma(X) = \sqrt{\mathbf{V}(X)}$.
La variance d'une somme de variables indépendantes est la somme de leurs variances.
▹ La covariance de deux variables réelles X et Y est le réel

$$\mathbf{Cov}(X, Y) = \mathbf{E}((X - \mathbf{E}(X))(Y - \mathbf{E}(Y))) = \mathbf{E}(XY) - \mathbf{E}(X)\mathbf{E}(Y).$$

La covariance de deux variables indépendantes est nulle.

Outre les définitions, il est bon de connaître ces paramètres pour les lois usuelles (uniforme, Bernoulli, binomiale).

EXERCICES

Vrai ou faux ?

		Vrai	Faux
a)	La somme de deux variables indépendantes de loi uniforme suit encore une loi uniforme.	☐	☐
b)	Si $X \hookrightarrow \mathscr{U}([\![0,n]\!])$, alors $n-X \hookrightarrow \mathscr{U}([\![0,n]\!])$.	☐	☐
c)	Si $X \hookrightarrow \mathscr{B}(n,p)$, alors $n-X \hookrightarrow \mathscr{B}(n,p)$.	☐	☐
d)	Deux variables X et Y sont indépendantes si, et seulement si, les événements $X=x$ et $Y=y$ sont indépendants pour tout x et tout y.	☐	☐
e)	Si deux variables X et Y sont de somme constante, alors elles ne sont pas indépendantes.	☐	☐
f)	Si X et Y sont indépendantes et X et Z sont indépendantes alors Y et Z sont indépendantes.	☐	☐
g)	Si X et Y sont indépendantes et X et Z sont indépendantes alors X et (Y,Z) sont indépendantes.	☐	☐
h)	Si X et (Y,Z) sont indépendantes, alors X et Y sont indépendantes.	☐	☐
i)	Soit X une variable (finie) à valeurs dans $\mathbb{N}$. Alors, $\mathbf{E}(X)=\sum_{k\in\mathbb{N}}\mathbf{P}(X>k)$.	☐	☐
j)	Pour toute variable réelle X, $\mathbf{E}(\lvert X\rvert)\leq\lvert\mathbf{E}(X)\rvert$.	☐	☐
k)	Pour toute variable réelle X, $\mathbf{E}(X)^2\leq\mathbf{E}(X^2)$.	☐	☐
l)	Soit une variable X telle que $\mathbf{V}(X)=0$, alors $\mathbf{Cov}(X,Y)=0$ pour toute variable Y.	☐	☐
m)	Soit une variable X telle que $\mathbf{E}(X^2)=0$, alors $\mathbf{E}(XY)=0$ pour toute variable Y.	☐	☐
n)	Des variables indépendantes deux à deux sont non corrélées.	☐	☐

Exercices d'application du chapitre 25

●○○ **Exercice 1**

On considère une variable aléatoire à valeurs dans $\{0,1,2\}$ telle que $\mathbf{E}(X)=1$ et $\mathbf{V}(X)=\frac{1}{2}$. Déterminer la loi de X.

●○○ **Exercice 2**

Soit X une variable aléatoire à valeurs dans $[\![0,n]\!]$ suivant une loi arithmétique $\mathbf{P}(X=k)=a\,k$. Déterminer a pour que X soit bien une variable aléatoire. En déduire $\mathbf{E}(X)$.

●●○ **Exercice 3**

Soit X et Y deux variables aléatoires indépendantes de loi uniforme sur $[\![1,n]\!]$.

1. Calculer $\mathbf{P}(X=Y)$ et $P(X \geq Y)$.
2. Déterminer la loi de $X-Y$.

●○○ **Exercice 4**

Soit $X_1, X_2, \ldots, X_n$ des variables aléatoires indépendantes suivant des lois de Bernoulli respectivement de paramètres $1, \frac{1}{2}, \ldots, \frac{1}{n}$. Soit N la variable aléatoire définie comme 0 si $X_1 = \ldots = X_n = 1$, $\min\{k \in [\![1,n]\!], X_k = 0\}$ sinon. Calculer la loi de N.

●○○ **Exercice 5**

Soit X une variable de loi binomiale de paramètres n et p. Déterminer l'entier j tel que la probabilité $\mathbf{P}(X=j)$ soit maximale.

●○○ **Exercice 6**

Soit X une variable aléatoire positive d'espérance nulle. Montrer que $\mathbf{P}(X=0)=1$.

●●○ **Exercice 7**

Soit $n \geq 2$, $m \in [\![1,n]\!]$, X_1 et X_2, deux variables aléatoires indépendantes de loi uniforme sur $[\![1,n]\!]$. Définissons la variable aléatoire Z par

$$\forall \omega \in \Omega, \qquad Z(\omega) = \begin{cases} X_1(\omega) & \text{si } X_2(\omega) \leq m, \\ X_2(\omega) & \text{sinon.} \end{cases}$$

1. Comparer les espérances des variables X_1, X_2 et Z.
2. Déterminer les valeurs de m qui maximisent l'espérance de Z.

●○○ **Exercice 8**

Soit $A_1, \ldots, A_n$ des événements d'un espace probabilisé fini $(\Omega, \mathscr{P}(\Omega), \mathbf{P})$. Définissons, pour tout $k \in [\![1,n]\!]$, B_k l'événement « au moins k événements parmi $A_1, \ldots, A_n$ sont réalisés ».

Montrer, en calculant l'espérance de la variable $X = \sum_{k=1}^{n} \mathbb{1}_{A_k}$, que

$$\sum_{k=1}^{n} \mathbf{P}(A_k) = \sum_{k=1}^{n} \mathbf{P}(B_k).$$

●●● **Exercice 9**

Soit $X_1, \ldots, X_n$ des variables aléatoires indépendantes et de même loi et N une variable aléatoire à valeurs dans $[\![0,n]\!]$ indépendante de $(X_1, \ldots, X_n)$.

Déterminer l'espérance et la variance de la variable $S_N = X_1 + \ldots + X_N$.

●●○ **Exercice 10**

Soit $n \geq 3$. Une urne contient 2 boules colorées et $n-2$ boules blanches. On tire successivement sans remise toutes les boules de l'urne.

1. Déterminer la loi de la variable aléatoire T_1 égale au numéro du tirage de la première boule colorée.
2. Déterminer la loi de la variable aléatoire T_2 égale au numéro du tirage de la seconde boule colorée.

●●○ **Exercice 11**

Une urne contient n boules numérotées de 1 à n. Tirons une poignée de boules (les 2^n poignées sont équiprobables) et notons X la somme des valeurs des boules extraites. Déterminer l'espérance de X.

●●○ **Exercice 12**

Soit X et Y deux variables aléatoires telles que

- les variables $X-Y$ et X sont indépendantes,
- les variables $X-Y$ et Y sont indépendantes.

Montrer que la variable $X-Y$ est constante.

●●○ **Exercice 13**

Soit $\varepsilon_1, \dots, \varepsilon_n$ des variables aléatoires indépendantes de loi uniforme sur $\{\pm 1\}$ et, pour tout $k \in [\![1, n]\!]$, $X_k = \varepsilon_1 \varepsilon_2 \dots \varepsilon_k$. Montrer que les variables $X_1, \dots, X_n$ sont mutuellement indépendantes.

●○○ **Exercice 14 Jeu de dés non transitif**

Considérons trois dés équilibrés à 6 faces étiquetés respectivement par $\{1,5,6,13,14,15\}$, $\{2,7,8,9,16,17\}$ et $\{3,4,10,11,12,18\}$ et notons X_1, X_2 et X_3 les variables aléatoires correspondant au lancer de chacun de ces dés.

1. Calculer $\mathbf{E}(X_i)$ pour $i \in \{1,2,3\}$.
2. Montrer que pour $j = i+1[3]$, $\mathbf{P}(X_i > X_j) < \frac{1}{2}$.

On a montré que le dé numéro 2 gagne plus d'une fois sur deux sur le dé 1, le dé 3 plus d'une fois sur deux sur le dé 2 et le dé 1 plus d'une fois sur deux sur le dé 3.

●●○ **Exercice 15**

Soit X et Y deux variables aléatoires indépendantes suivant la même loi de Bernoulli de paramètre p.

1. Déterminer la loi du couple $(U,V) = (X+Y, X-Y)$, en déduire les lois marginales.
2. Déterminer la loi de U conditionnée à $V=1$.
3. Les variables U et V sont-elles indépendantes ?

●●● **Exercice 16**

Pour toute variable aléatoire finie X à valeurs dans $\mathbb{N}$, on pose $f_X : t \mapsto \mathbf{E}(t^X)$.

1. Montrer que si les variables aléatoires X et Y à valeurs dans $\mathbb{N}$ sont indépendantes, alors
$$f_{X+Y} = f_X f_Y.$$
2. En déduire qu'il n'existe pas de moyen de piper deux dés indépendants à six faces (numérotées de 1 à 6) de sorte à ce que la somme des résultats d'un lancer des deux dés suive la loi uniforme sur l'ensemble $[\![2, 12]\!]$.

EXERCICES

Exercices d'approfondissement du chapitre 25

●●○ **Exercice A**

Soit $X_1, \dots, X_n$ des variables aléatoires indépendantes et de même loi, $\lambda > 0$. Montrer que $\mathbf{P}(X_1 > \lambda) > 0$ si, et seulement si, $\mathbf{P}\left(\sum_{k=1}^n X_k > n\lambda\right) > 0$.

●●○ **Exercice B**

Soit $n \geq 2$, $X_1, \dots, X_n$ des variables indépendantes de loi de Bernoulli de paramètre $p \in]0,1[$, L la matrice ligne dont les coefficients sont les X_j et $M = L^{\mathrm{T}} L$.

1. Déterminer les lois des variables $\det(M)$, $\mathrm{rg}(M)$ et $\mathrm{tr}(M)$.
2. Quelle est la probabilité que M soit une matrice de projection ?

●●● **Exercice C**

Soient $(X_{i,j})_{1\leq i,j\leq n}$ une suite de variables aléatoires indépendantes à valeurs dans $\{-1,1\}$, de loi donnée par

$$\mathbf{P}(X_{i,j}=1)=\mathbf{P}(X_{i,j}=-1)=\frac{1}{2}.$$

Déterminer l'espérance du déterminant de la matrice $M=(X_{i,j})_{1\leq i\leq j\leq n}$.

●●○ **Exercice D Allumettes de Banach**

Considérons deux boîtes d'allumettes A et B contenant chacune n allumettes. On choisit aléatoirement une boîte et on en retire une allumette de cette boîte puis on recommence jusqu'à ce que l'une des deux boîtes soit vide. Déterminer la loi de X, la variable définie par le nombre d'allumettes restant dans l'autre boîte.

●●● **Exercice E**

Soit X et Y des variables aléatoires indépendantes sur $[\![1,n]\!]$. Notons $p_k=\mathbf{P}(X=k)$ et $q_k=\mathbf{P}(Y=k)$ pour tout $k\in[\![1,n]\!]$. Montrer que

$$2\mathbf{P}(X\neq Y)\geq\sum_{k=1}^{n}|p_k-q_k|.$$

●●○ **Exercice F**

Soit $p>2$ un nombre premier, X et Y deux variables aléatoires indépendantes et de loi uniforme sur l'ensemble (fini) $[\![0,p-1]\!]$. Pour tout $r\in[\![0,p-1]\!]$, posons $Z_r=X+rY[p]$ Montrer que les variables aléatoires Z_r pour $r\in[\![0,p-1]\!]$ sont deux à deux indépendantes.

Problèmes du chapitre 25

●○○ **Problème 1**

On considère un mobile se déplaçant sur les points à coordonnées entières de l'axe des abscisses avec les règles suivantes :

- initialement, le mobile est à l'abscisse 0 ;
- si, au temps n, le mobile est en position k, il sera au temps $n+1$ en position $k+1$ avec probabilité $\frac{k+1}{k+2}$ et en position 0 avec la probabilité $\frac{1}{k+2}$.

On note, pour tout $n\in\mathbb{N}$, X_n la variable aléatoire correspondant à la position du mobile au temps n et u_n la probabilité que X_n soit égale à 0.

1. Quelles sont les valeurs prises par la variable aléatoire X_n ?
2. a. Montrer que, pour tout $n\in\mathbb{N}$ et tout $k\in\{1,\dots,n\}$,

$$\mathbf{P}(X_n=k)=\frac{k}{k+1}\mathbf{P}(X_{n-1}=k-1).$$

   b. En déduire que, pour tout $n\in\mathbb{N}$ et tout $k\in\{1,\dots,n\}$,

$$\mathbf{P}(X_n=k)=\frac{1}{k+1}u_{n-k}.$$

   c. Montrer que,

$$\forall n,\qquad \sum_{j=0}^{n}\frac{u_j}{n+1-j}=1.$$

   d. En déduire la valeur de u_2 et u_3.

3. a. Déterminer l'espérance de X_n en fonction des termes de la suite $(u_n)_n$.
   b. Calculer
$$\sum_{j=0}^{n-1} \frac{u_j}{n-j}.$$
   c. Montrer que
$$u_n = \sum_{j=0}^{n-1} \frac{u_j}{(n-j)(n+1-j)}.$$
   d. Montrer que $(n+1)u_n \geq 1$ pour tout n puis calculer la limite de $\mathbf{E}(X_n)$.

●●○ **Problème 2 Inégalité maximale de Kolmogorov**

Soit $n \geq 1$ un entier, $a > 0$ et $X_1, \ldots, X_n$ des variables aléatoires réelles discrètes indépendantes centrées (mais on ne suppose pas ces variables de même loi).

Posons $S_k = X_1 + \ldots + X_k$ pour tout $k \leq n$ avec la convention usuelle $S_0 = 0$. Ce court problème établit l'inégalité maximale de Kolmogorov :
$$\mathbf{P}\Big(\max_{1\leq k\leq n} |S_k| \geq a\Big) \leq \frac{1}{a^2} \sum_{k=1}^{n} \mathbf{V}(X_k).$$

1. Soit A l'événement $(\max_{1\leq k\leq n} |S_k| \geq a)$. Pour $1 \leq k \leq n$, on introduit l'événement
$$A_k = \Big(\bigcap_{1\leq j<k} (|S_j| < a)\Big) \cap (|S_k| \geq a).$$
   Montrer que les événements $A_1, \ldots, A_n$ sont deux à deux disjoints et que $A = A_1 \cup A_2 \cup \ldots \cup A_n$.
2. Montrer que
$$\mathbf{E}(S_n^2) \geq \sum_{k=1}^{n} \mathbf{E}\big(S_n^2 \mathbb{1}_{A_k}\big).$$
3. Montrer que
$$\mathbf{E}\big(S_n^2 \mathbb{1}_{A_k}\big) \geq \mathbf{E}\big(S_k^2 \mathbb{1}_{A_k}\big) + 2\mathbf{E}(S_n - S_k)\mathbf{E}\big(S_k \mathbb{1}_{A_k}\big).$$
4. Conclure.

●●● **Problème 3**

Définition 25.34.

Soit $a_1, \ldots, a_n \in \mathbb{R}$ deux à deux distincts. Une sous-suite croissante (respectivement décroissante) est une suite $a_{i_1} < a_{i_2} < \ldots < a_{i_k}$ (respectivement $a_{i_1} > a_{i_2} > \ldots > a_{i_k}$) avec des indices en ordre strictement croissant.
Notons $C(a_1, \ldots, a_n)$ la longueur de la plus longue sous-suite croissante de $a_1, \ldots, a_n$, et $D(a_1, \ldots, a_n)$ la longueur de la plus longue sous-suite décroissante.

Partie A

Cette partie n'utilise pas de connaissance de probabilités ; elle est seulement utile pour la seconde partie.

Soit $a = (a_1, \ldots, a_n)$ une suite finie de réels distincts, r et $s \geq 1$ des entiers tels que $rs < n$.

Pour tout $i \in [\![1, n]\!]$, considérons (c_i, d_i), où c_i est la longueur de la plus longue sous-suite croissante qui se termine par a_i et d_i est la longueur de la plus longue sous-suite décroissante qui se termine par a_i.

1. Pour $1 \leq i < j \leq n$, montrer que $(c_i, d_i) \neq (c_j, d_j)$.
2. En déduire qu'il existe $1 \leq i \leq n$ tel que $c_i > r$ ou $d_i > s$.
3. Montrer que $C(a)D(a) \geq n$.

PROBLÈMES

Partie B

Soit $X_n = (X_n(1), X_n(2), \ldots, X_n(n))$ une variable aléatoire de loi uniforme sur le groupe symétrique S_n.

1. Justifier que $\mathbf{E}(C(X_n)) = \mathbf{E}(D(X_n))$.
2. Montrer que $\mathbf{E}(C(X_n)) \geq \sqrt{n}$.
3. Soit $1 \leq i_1 < i_2 < \ldots < i_k \leq n$. Calculer la probabilité
$$\mathbf{P}(X_n(i_1) < X_n(i_2) < \ldots < X_n(i_k)).$$
4. Notons L_k^n le nombre de sous-suites croissantes de longueur k de X_n. Calculer $\mathbf{E}(L_k^n)$.
5. Montrer que, pour tout $\varepsilon > 0$, $\mathbf{P}(C(X_n) \geq (e+\varepsilon)\sqrt{n})$ décroît exponentiellement vite lorsque $n \to \infty$.
6. Montrer que, pour tout $\varepsilon > 0$, pour tout n assez grand,
$$\frac{\mathbf{E}(C(X_n))}{\sqrt{n}} \leq e + \varepsilon.$$

CORRIGÉS

Corrigés des Vrai/Faux

a) Faux.

b) Vrai.

c) Faux.

d) Vrai, c'est la définition.

e) Faux, par exemple avec des variables constantes.

f) Faux, par exemple avec $Y = Z$.

g) Faux, par exemple, avec Y, Z indépendantes de loi uniforme sur $\{\pm 1\}$ et $X = YZ$.

h) Vrai.

i) Vrai.

j) Faux, avec par exemple une variable centrée, non constante égale à 0.

k) Vrai, c'est l'inégalité de Cauchy-Schwarz.

l) Vrai, d'après l'inégalité Cauchy-Schwarz, $0 \leq \mathbf{Cov}(X, Y)^2 \leq \mathbf{V}(X)\mathbf{V}(Y) = 0$ donc $\mathbf{Cov}(X, Y) = 0$.

m) Vrai, d'après l'inégalité Cauchy-Schwarz, $0 \leq (\mathbf{E}(XY))^2 \leq \mathbf{E}(X^2)\mathbf{E}(Y^2) = 0$ donc $\mathbf{E}(XY) = 0$.

n) Vrai.

Corrigés des exercices d'application du chapitre 25

Exercice 1

Calculons $p_0 = \mathbf{P}(X = 0)$, $p_1 = \mathbf{P}(X = 1)$ et $p_2 = \mathbf{P}(X = 2)$.

▹ Par définition de la loi de X, on a $p_0 + p_1 + p_2 = 1$.

▹ Par définition de l'espérance de X, on a $p_1 + 2p_2 = 1$.

▹ Enfin, par définition de la variance, $\mathbf{E}(X^2) = \mathbf{V}(X) + \mathbf{E}(X)^2 = \frac{3}{2}$ donc on a $p_1 + 4p_2 = \frac{3}{2}$.

En résolvant ce système, on obtient $p_2 = \frac{1}{4}$, $p_1 = \frac{1}{2}$ et enfin $p_0 = \frac{1}{4}$. On reconnaît la loi binomiale de paramètres 2 et $\frac{1}{2}$.

Exercice 2

▹ Pour que l'on obtienne effectivement une loi de probabilité, il faut et il suffit que $a \geq 0$ et $\sum_{k=0}^{n} \mathbf{P}(X=k)=1$: la seule valeur convenant est $a=\frac{2}{n(n+1)}$.

▹ L'espérance de la variable X est alors $\mathbf{E}(X)=\frac{2}{n(n+1)}\sum_{k=0}^{n} k^2=\frac{2n+1}{3}$.

Exercice 3

1. En écrivant l'événement $X=Y$ comme la réunion des $(X=k)$ et $(Y=k)$,

$$\mathbf{P}(X=Y)=\sum_{k=1}^{n}\mathbf{P}(X=k,Y=k)=\sum_{k=1}^{n}\mathbf{P}(X=k)\mathbf{P}(Y=k)=\sum_{k=1}^{n}\frac{1}{n^2}=\frac{1}{n}.$$

Par ailleurs, l'univers est la réunion des événements $X=Y$, $X>Y$ et $X<Y$ donc

$$1=\mathbf{P}(X=Y)+\mathbf{P}(X>Y)+\mathbf{P}(X<Y)=\frac{1}{n}+2\mathbf{P}(X>Y).$$

En conclusion, $\mathbf{P}(X>Y)=\frac{n-1}{2n}$ et $\mathbf{P}(X\geq Y)=\mathbf{P}(X>Y)+\mathbf{P}(X=Y)=\frac{n+1}{2n}$.

2. La variable $X-Y$ est à valeurs dans $[\![-(n-1),n-1]\!]$. Pour tout $k\in[\![-(n-1),n-1]\!]$,

$$\begin{aligned}\mathbf{P}(X-Y=k)&=\sum_{j=1}^{n-k}\mathbf{P}(X=k+j,Y=j)\\&=\sum_{k=1}^{n-k}\mathbf{P}(X=k+j)\mathbf{P}(Y=j)\\&=\sum_{k=1}^{n-k}\frac{1}{n^2}=\frac{n-k}{n^2}.\end{aligned}$$

Exercice 4

Pour tout $k\in[\![1,n]\!]$,

$$\begin{aligned}\mathbf{P}(N=k)&=\mathbf{P}(X_1=1,\dots,X_{k-1}=1,X_k=0)\\&=\mathbf{P}(X_1=1)\dots\mathbf{P}(X_{k-1}=1)\mathbf{P}(X_k=0)=\frac{1}{(k-1)!}\Big(1-\frac{1}{k}\Big).\end{aligned}$$

Par télescopage, $\sum_{k=1}^{n}\mathbf{P}(N=k)=1-\frac{1}{n!}$, donc $\mathbf{P}(N=0)=\frac{1}{n!}$.

Exercice 5

Pour tout $k\in[\![0,n-1]\!]$, $\frac{\mathbf{P}(X=k+1)}{\mathbf{P}(X=k)}=\frac{n-k}{k+1}\frac{p}{1-p}$.
Ce terme est supérieur ou égal à 1 si, et seulement si, $p(n-k)\geq(1-p)(k+1)$, c'est-à-dire $(k+1)\leq p(n+1)$.
Par conséquent, la suite $(\mathbf{P}(X=j))_j$ est croissante jusqu'à l'indice $\lfloor p(n+1)\rfloor$ puis décroissante. L'entier recherché est $\lfloor p(n+1)\rfloor$.

Exercice 6

Raisonnons par l'absurde en supposant que $\mathbf{P}(X=0)<1$: notons $m=\min X(\Omega)\setminus\{0\}$. Alors,

$$0=\mathbf{E}(X)=\sum_{x\in X(\Omega)}x\mathbf{P}(X=x)=\sum_{x\in X(\Omega)\setminus\{0\}}x\mathbf{P}(X=x)\geq m\mathbf{P}(X>0)=m(1-\mathbf{P}(X=0))>0,$$

d'où la contradiction.

Exercice 7

1. Par définition, Z est à valeurs dans $[\![1,n]\!]$. Commençons par déterminer la loi de Z, c'est-à-dire les $\mathbf{P}(Z=k)$ pour $k\in[\![1,n]\!]$. Pour cela, écrivons l'événement $Z=k$ sous la forme

$$(X_1=k \text{ et } X_2\leq m) \text{ ou } (X_2=k \text{ et } X_2>m).$$

▹ Si $k\leq m$, en utilisant l'indépendance de X_1 et de X_2,

$$\mathbf{P}(Z=k)=\mathbf{P}(X_1=k,X_2\leq m)=\mathbf{P}(X_1=k)\mathbf{P}(X_2\leq m)=\frac{1}{n}\frac{m}{n}=\frac{m}{n^2},$$

car X_1 et X_2 suivent des lois uniformes.
▹ Si $k>m$,

$$\begin{aligned}\mathbf{P}(Z=k)&=\mathbf{P}(X_1=k,X_2\leq m)+\mathbf{P}(X_2=k,X_2>m)\\&=\mathbf{P}(X_1=k)\mathbf{P}(X_2\leq m)+\mathbf{P}(X_2=k)\\&=\frac{m}{n^2}+\frac{1}{n}.\end{aligned}$$

Alors,

$$\begin{aligned}\mathbf{E}(Z)&=\sum_{k=1}^{n}k\mathbf{P}(Z=k)\\&=\sum_{k=1}^{m}k\frac{m}{n^2}+\sum_{k=m+1}^{n}k\Big(\frac{m}{n^2}+\frac{1}{n}\Big)=\sum_{k=1}^{n}k\frac{m}{n^2}+\sum_{k=m+1}^{n}k\frac{1}{n}\\&=\frac{(n+1)m}{2n}+\frac{(n-m)(m+1+n)}{2n}=\frac{n+1}{2}+\frac{m(n-m)}{2n}.\end{aligned}$$

Comme X_1 et X_2 suivent la même loi uniforme, $\mathbf{E}(X_1)=\mathbf{E}(X_2)=\frac{n+1}{2}$; on en déduit que $\mathbf{E}(Z)\geq\mathbf{E}(X_1)$.
2. D'après le calcul précédent,

$$\mathbf{E}(Z)=\frac{1}{2n}\Big(\frac{5}{4}n^2+n-\Big(m-\frac{n}{2}\Big)^2\Big)$$

donc l'espérance $\mathbf{E}(Z)$ est maximale lorsque $\left|m-\frac{n}{2}\right|$ est minimale. Si n est pair, c'est le cas quand $m=\frac{n}{2}$. Si n est impair quand $m=\frac{n-1}{2}$ et $m=\frac{n+1}{2}$.

Exercice 8

▹ Commençons déjà par remarquer que par linéarité de l'espérance, $\mathbf{E}(X)=\sum_{k=1}^{n}\mathbf{P}(A_k)$.
▹ Calculons la loi de X de sorte à obtenir une seconde expression de son espérance. La variable X prend pour valeurs les entiers entre 0 et n (puisque chaque indicatrice ne peut prendre que les valeurs 0 ou 1). Par ailleurs, pour tout $k\in[\![0,n]\!]$, $\mathbf{P}(X\geq k)=\mathbf{P}(B_k)$. Alors,

$$\begin{aligned}\mathbf{E}(X)&=\sum_{k=0}^{n}k\big(\mathbf{P}(X\geq k)-\mathbf{P}(X\geq k+1)\big)\\&=\sum_{k=0}^{n}k\mathbf{P}(B_k)-\sum_{k=0}^{n}k\mathbf{P}(B_{k+1})\\&=\sum_{k=0}^{n}k\mathbf{P}(B_k)-\sum_{k=1}^{n+1}(k-1)\mathbf{P}(B_k)\\&=\sum_{k=1}^{n}k\mathbf{P}(B_k).\end{aligned}$$

CORRIGÉS

Par conséquent,

$$\sum_{k=1}^{n}\mathbf{P}(A_k)=\sum_{k=1}^{n}\mathbf{P}(B_k).$$

Exercice 9

▹ Commençons par utiliser la formule des probabilités totales. Pour tout x,

$$\begin{aligned}\mathbf{P}(S_N=x)&=\sum_{k=0}^{n}\mathbf{P}(S_N=x|N=k)\mathbf{P}(N=k)\\&=\sum_{k=0}^{n}\mathbf{P}(S_k=x|N=k)\mathbf{P}(N=k)\\&=\sum_{k=0}^{n}\mathbf{P}(S_k=x)\mathbf{P}(N=k)\end{aligned}$$

en utilisant que les variables S_k et N sont indépendantes. Par conséquent,

$$\begin{aligned}\mathbf{E}(S_N)&=\sum_{x}x\mathbf{P}(S_N=x)\\&=\sum_{k=0}^{n}\sum_{x}x\mathbf{P}(S_k=x)\mathbf{P}(N=k)\\&=\sum_{k=0}^{n}\mathbf{E}(S_k)\mathbf{P}(N=k).\end{aligned}$$

Or, par linéarité de l'espérance, $\mathbf{E}(S_k)=k\mathbf{E}(X_1)$,

$$\mathbf{E}(S_N)=\mathbf{E}(X_1)\sum_{k=0}^{n}k\mathbf{P}(N=k)=\mathbf{E}(X_1)\mathbf{E}(N).$$

▹ Après calculs, on obtient en utilisant l'indépendance :

$$\mathbf{E}(S_N^2)=\sum_{k=0}^{n}(k\mathbf{V}(X_1)+k^2\mathbf{E}(X_1)^2)\mathbf{P}(N=k)=\mathbf{E}(N)\mathbf{V}(X_1)+\mathbf{E}(X_1)^2\mathbf{E}(N^2).$$

En conclusion, $\mathbf{V}(S_N)=\mathbf{E}(N)\mathbf{V}(X_1)+\mathbf{E}(X_1)^2\mathbf{V}(N)$.

Exercice 10

1. Commençons par indiquer que $T_1(\Omega)=[\![1,n-1]\!]$. Ensuite, pour tout $k\in[\![1,n-1]\!]$, on dénombre $n-k$ tirages tels que la première boule colorée soit tirée au $k^{\text{ième}}$ tirage (autant que de choix pour la position du tirage de la seconde boule colorée après k). La probabilité que $T_1=k$ est donc

$$\mathbf{P}(T_1=k)=\frac{n-k}{\binom{n}{2}}=\frac{2(n-k)}{n(n-1)}.$$

Remarque

On aurait également pu discuter selon les premiers tirages. En notant, pour tout $j\in[\![1,n]\!]$, B_j l'événement « le $j^{\text{ième}}$ tirage donne une boule blanche », on obtient avec la formule des probabilités composées

$$\begin{aligned}\mathbf{P}(T_1=k)&=\mathbf{P}(B_1\cap\ldots\cap B_{k-1}\cap\overline{B_k})\\&=\Big(\prod_{j=1}^{k-1}\mathbf{P}(B_j|B_1\cap\ldots\cap B_{j-1})\Big)\mathbf{P}(\overline{B_k}\big|B_1\cap\ldots\cap B_{k-1})\\&=\Big(\prod_{j=1}^{k-1}\frac{n-1-j}{n+1-j}\Big)\frac{2}{n-j+1}=\frac{2(n-k)}{n(n-1)}.\end{aligned}$$

On remarque que le calcul ainsi mené est plus laborieux.

2. Commençons par indiquer que $T_2(\Omega) = [\![2, n]\!]$. Ensuite, pour tout $k \in [\![2, n]\!]$, on dénombre $k-1$ tirages tels que la seconde boule colorée soit tirée au $k^{\text{ième}}$ tirage (autant que de choix pour la position du tirage de la première boule colorée avant k). La probabilité que $T_2 = k$ est donc

$$\mathbf{P}(T_2 = k) = \frac{k-1}{\binom{n}{2}} = \frac{2(k-1)}{n(n-1)}.$$

Remarque
D'après ces calculs, les variables T_1 et $n+1-T_2$ suivent la même loi ce qui peut s'expliquer par un argument de symétrie en « retournant le temps. »

Exercice 11

Notons, pour tout $k \in [\![1, n]\!]$, A_k l'événement « la boule de numéro k est tirée ». La variable X est alors égale à

$$X = \sum_{k=1}^{n} k \mathbb{1}_{A_k}.$$

Par linéarité de l'espérance, on obtient

$$\mathbf{P}(X) = \sum_{k=1}^{n} k \mathbf{P}(A_k).$$

D'après l'hypothèse d'uniformité, $\mathbf{P}(A_k) = \frac{1}{2}$ pour tout k donc $\mathbf{E}(X) = \frac{n(n+1)}{4}$.

Exercice 12

Montrons que la variance de $X - Y$ est nulle.

$$\begin{aligned}
\mathbf{V}(X-Y) &= \mathbf{E}((X-Y)^2) - (\mathbf{E}(X-Y))^2 \\
&= \mathbf{E}((X-Y)(X-Y)) - (\mathbf{E}(X-Y))^2 \\
&= \mathbf{E}((X-Y)X) - \mathbf{E}((X-Y)Y) - (\mathbf{E}(X-Y))^2 \\
&= \mathbf{E}(X-Y)\mathbf{E}(X) - \mathbf{E}(X-Y)\mathbf{E}(Y) - (\mathbf{E}(X-Y))^2 \\
&= (\mathbf{E}(X-Y))^2 - (\mathbf{E}(X-Y))^2 = 0.
\end{aligned}$$

Exercice 13

Pour tout $(x_1, \dots, x_n) \in \{\pm 1\}^n$,

$$\begin{aligned}
\mathbf{P}(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) &= \mathbf{P}(\varepsilon_1 = x_1, \varepsilon_2 = x_2 x_1, \dots, \varepsilon_n = x_n x_{n-1}) \\
&= \mathbf{P}(\varepsilon_1 = x_1)\mathbf{P}(\varepsilon_2 = x_2 x_1) \dots, \mathbf{P}(\varepsilon_n = x_n x_{n-1}) = \frac{1}{2^n}.
\end{aligned}$$

De cette loi jointe, on en déduit que, pour tout $x_1 \in \{\pm 1\}$, $\mathbf{P}(X_1 = x_1) = \frac{1}{2}$ car c'est la somme de 2^{n-1} termes de probabilité $\frac{1}{2^n}$.
Ainsi, pour tout $(x_1, \dots, x_n) \in \{\pm 1\}^n$,

$$\mathbf{P}(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) = \mathbf{P}(X_1 = x_1)\mathbf{P}(X_2 = x_2) \dots \mathbf{P}(X_n = x_n).$$

En conclusion, les variables $X_1, \dots, X_n$ sont mutuellement indépendantes.

Exercice 14

1. En effectuant la moyenne, on trouve $\mathbf{E}(X_1)=\frac{54}{6}$, $\mathbf{E}(X_2)=\frac{59}{6}$ et $\mathbf{E}(X_3)=\frac{58}{6}$.
2. Calculons ces probabilités avec la formule des probabilités totales pour les événements $X_j=k$.

$$\begin{aligned}
\mathbf{P}(X_1>X_2)&=\frac{\mathbf{P}(X_1>2)+\mathbf{P}(X_1>7)+\mathbf{P}(X_1>8)+\mathbf{P}(X_1>9)+\mathbf{P}(X_1>16)+\mathbf{P}(X_1>17)}{6}\\
&=\frac{1}{6}\Big(\frac{5}{6}+\frac{3}{6}+\frac{3}{6}+\frac{3}{6}+\frac{0}{6}+\frac{0}{6}\Big)=\frac{14}{36}<\frac{1}{2}.\\
\mathbf{P}(X_2>X_3)&=\frac{\mathbf{P}(X_2>3)+\mathbf{P}(X_2>4)+\mathbf{P}(X_2>10)+\mathbf{P}(X_2>11)+\mathbf{P}(X_2>12)+\mathbf{P}(X_2>18)}{6}\\
&=\frac{1}{6}\Big(\frac{5}{6}+\frac{5}{6}+\frac{2}{6}+\frac{2}{6}+\frac{2}{6}+\frac{0}{6}\Big)=\frac{16}{36}<\frac{1}{2}.\\
\mathbf{P}(X_3>X_1)&=\frac{\mathbf{P}(X_3>1)+\mathbf{P}(X_3>5)+\mathbf{P}(X_3>6)+\mathbf{P}(X_3>13)+\mathbf{P}(X_3>14)+\mathbf{P}(X_3>15)}{6}\\
&=\frac{1}{6}\Big(\frac{6}{6}+\frac{4}{6}+\frac{4}{6}+\frac{1}{6}+\frac{1}{6}+\frac{1}{6}\Big)=\frac{17}{36}<\frac{1}{2}.
\end{aligned}$$

Exercice 15

1. Il y a quatre valeurs possibles pour la variable (X,Y) qui donnent chacune une valeur différente pour le couple (U,V) donc

$$\begin{aligned}
\mathbf{P}((U,V)=(2,0))&=\mathbf{P}((X,Y)=(1,1))=p^2,\\
\mathbf{P}((U,V)=(0,0))&=\mathbf{P}((X,Y)=(0,0))=(1-p)^2,\\
\mathbf{P}((U,V)=(1,1))&=\mathbf{P}((X,Y)=(1,0))=p(1-p),\\
\mathbf{P}((U,V)=(1,-1))&=\mathbf{P}((X,Y)=(0,1))=p(1-p).
\end{aligned}$$

2. Clairement conditionnellement à $V=1$, U est la variable constante égale à 1 car

$$\mathbf{P}(U=1|V=1)=\frac{\mathbf{P}(U=1,V=1)}{\mathbf{P}(V=1)}=\frac{p(1-p)}{p(1-p)}=1.$$

3. Comme $\mathbf{P}(V=1)=p(1-p)$, $\mathbf{P}(U=1)=2p(1-p)$ et $\mathbf{P}((U,V)=(1,1))=p(1-p)$,

$$\mathbf{P}((U,V)=(1,1))=\mathbf{P}(U=1)\mathbf{P}(V=1)$$

si, et seulement si, $p=0$, $p=1$ ou $2p(1-p)=1$, c'est-à-dire $p\in\{0,1\}$. Ainsi, les variables U et V sont indépendantes si, et seulement si, les variables X et Y sont constantes.

Exercice 16

1. Soit $t\in\mathbb{R}_+^*$. Comme les variables X et Y sont indépendantes, les variables t^X et t^Y sont elles aussi indépendantes. Par conséquent, $\mathbf{E}(t^{X+Y})=\mathbf{E}(t^Xt^Y)=\mathbf{E}(t^X)\mathbf{E}(t^Y)$. Ainsi, $f_{X+Y}(t)=f_X(t)f_Y(t)$.
2. Soit deux dés (pipés) à six faces numérotées de 1 à 6, X_1 et X_2 les variables aléatoires correspondant au lancer de ces dés. Supposons que la variable X_1+X_2 suive la loi uniforme sur $[\![2,12]\!]$. Alors, avec la question précédente, pour tout $t>0$,

$$f_{X_1}(t)f_{X_2}(t)=\frac{1}{11}\Big(t^2+t^3+\ldots+t^{12}\Big).$$

Notons $f_{X_1}(t)=tP_1(t)$ et $f_{X_2}(t)=tP_2(t)$ avec P_1 et P_2 des polynômes de degré 5 pour obtenir l'identité

$$\forall t\geq 0,\qquad 11(1-t)P_1(t)P_2(t)=1-t^{11}.$$

Cette identité entre polynômes est vérifiée pour une infinité de t donc est vérifiée pour tout $t \in \mathbb{R}$. Remarquons que le membre de gauche admet au moins 3 racines réelles (car P_1 et P_2 sont de degré impair) alors que le membre de droite n'admet qu'une racine réelle simple : contradiction.

Corrigés des exercices d'approfondissement du chapitre 25

Exercice A

Montrons la propriété équivalente : $\mathbf{P}(X_1 > \lambda) = 0$ si, et seulement si, $\mathbf{P}(\sum_{k=1}^n X_k > n\lambda) = 0$.

▹ Pour le sens direct, utilisons l'inclusion $(\sum_{k=1}^n X_k > n\lambda) \subset \cup_{k=1}^n (X_k > \lambda)$ (pour qu'une somme de n termes soit supérieure à $n\lambda$, il faut qu'au moins un terme soit supérieur à λ). Par conséquent

$$\mathbf{P}\Big(\sum_{k=1}^n X_k > n\lambda\Big) \leq \mathbf{P}\Big(\bigcup_{k=1}^n (X_k > \lambda)\Big) \leq \sum_{k=1}^n \mathbf{P}(X_k > \lambda) = n\mathbf{P}(X_1 > \lambda).$$

Ainsi, le terme de droite est nul par hypothèse donc $\mathbf{P}(\sum_{k=1}^n X_k > n\lambda) = 0$.

▹ Utilisons cette fois-ci l'inclusion $\cap_{k=1}^n (X_k > \lambda) \subset (\sum_{k=1}^n X_k > n\lambda)$ (si les n termes sont supérieurs à λ, la somme est supérieure à $n\lambda$). Ainsi, par indépendance des X_k,

$$\mathbf{P}\Big(\sum_{k=1}^n X_k > n\lambda\Big) \geq \mathbf{P}\Big(\bigcap_{k=1}^n (X_k > \lambda)\Big) = \prod_{k=1}^n \mathbf{P}(X_k > \lambda) = (\mathbf{P}(X_1 > \lambda))^n.$$

Comme le membre de gauche est nul par hypothèse, $\mathbf{P}(X_1 > \lambda) = 0$.

Exercice B

1. ▹ Comme la matrice M est de rang au plus 1 (toutes les colonnes sont proportionnelles à L^{T}), la variable aléatoire $\det(M)$ est certaine égale à 0.

▹ D'après la remarque précédente, la matrice M est de rang 0 ou 1. Elle est de rang 0 si, et seulement si, L est la ligne nulle, c'est-à-dire qi $X_1 = \ldots = X_n = 0$. Ainsi, $\mathbf{P}(\mathrm{rg}(M) = 0) = (1-p)^n$ et donc $\mathrm{rg}(M)$ suit la loi de Bernoulli $\mathscr{B}(1-(1-p)^n)$.

▹ D'après les propriétés de la trace,

$$\mathrm{tr}(M) = \mathrm{tr}\big(L^{\mathrm{T}} L\big) = \mathrm{tr}\big(L L^{\mathrm{T}}\big) = \sum_{k=1}^n X_k^2 = \sum_{k=1}^n X_k.$$

Comme les variables $X_1, \ldots, X_n$ sont indépendantes de loi de Bernoulli de paramètre p, $\mathrm{tr}(M)$ suit donc la loi binomiale de paramètres n et p.

2. Remarquons que $M^2 = \mathrm{tr}(M)M$. Ainsi, M est de projection si, et seulement si, $\mathrm{tr}(M) = 1$ ou $M = 0_n$. Par conséquent, la probabilité recherchée est

$$np(1-p)^{n-1} + (1-p)^n = (1-p)^{n-1}(1+(n-1)p).$$

Exercice C

Notons $\Delta_{i,j}$ le cofacteur correspondant au terme en position i, j. Développons par rapport à la première ligne $\det M = \sum_{j=1}^{n} X_{1,j}\Delta_{1,j}$ puis passons à l'espérance

$$\mathbf{E}(\det M) = \mathbf{E}\Big(\sum_{j=1}^{n} X_{1,j}\Delta_{1,j}\Big) = \sum_{j=1}^{n} \mathbf{E}(X_{1,j}\Delta_{1,j}).$$

La variable $\Delta_{1,j}$ s'écrit comme somme et produits de $X_{i,k}$, avec $i \geq 2$ et $k \neq j$. Ce sont toutes des variables aléatoires indépendantes de $X_{1,j}$. Donc $X_{1,j}$ et $\Delta_{1,j}$ sont indépendants. Ainsi,

$$\mathbf{E}(\det M) = \sum_{j=1}^{n} \mathbf{E}(X_{1,j})\mathbf{E}(\Delta_{1,j}).$$

Pour tout $j \in [\![1, n]\!]$, $\mathbf{E}(X_{1,j}) = \frac{1}{2} - \frac{1}{2} = 0$, donc $\mathbf{E}(\det M) = 0$.

Exercice D

Remarquons tout d'abord que X est à valeurs dans $[\![0, n]\!]$. Fixons $k \in [\![0, n]\!]$ et calculons $\mathbf{P}(X = k)$ en décomposant l'événement $X = k$ en deux événements « la boîte A est vide et il en reste k dans la boîte B » et « la boîte B est vide et il en reste k dans la boîte A ».

▹ On suppose qu'on choisit chaque boîte avec la même probabilité $\frac{1}{2}$.

Pour chacun de ces événements, il y a eu $2n-k$ tirages, le dernier est forcément dans la boîte vidée et il y en a eu $n-1$ parmi les $2n-k-1$ précédents dans cette même boîte. On reconnaît une loi binomiale dans un schéma de tirages indépendants de même probabilité $\frac{1}{2}$. Par conséquent,

$$\mathbf{P}(X = k) = 2 \times \frac{1}{2} \times \binom{2n-k-1}{n-1}\Big(\frac{1}{2}\Big)^{n-1}\Big(1-\frac{1}{2}\Big)^{n-k} = \binom{2n-k-1}{n-1}\frac{1}{2^{2n-k-1}}.$$

▹ On suppose qu'on choisit chaque boîte avec la probabilité p et $(1-p)$. Le même raisonnement donne

$$\begin{aligned}\mathbf{P}(X = k) &= p\binom{2n-k-1}{n-1}p^{n-1}(1-p)^{n-k} + (1-p)\binom{2n-k-1}{n-1}(1-p)^{n-1}p^{n-k}\\ &= \binom{2n-k-1}{n-1}(p^n(1-p)^{n-k} + p^{n-k}(1-p)^n).\end{aligned}$$

Exercice E

Partitionnons l'événement $X \neq Y$ selon la valeur de Y puis utilisons l'indépendance :

$$\mathbf{P}(X \neq Y) = \sum_{k=1}^{n} \mathbf{P}(X = k, Y \neq k) = \sum_{k=1}^{n} p_k(1-q_k).$$

D'où

$$2\mathbf{P}(X \neq Y) = \sum_{k=1}^{n}\Big(p_k(1-q_k) + q_k(1-p_k)\Big).$$

Or, pour tous $x, y \in [0,1]$, $x(1-y) + y(1-x) \geq x - y$ (car $2(1-x)y \geq 0$) et, par symétrie,

$$x(1-y) + y(1-x) \geq |x-y|.$$

En conclusion, $2\mathbf{P}(X \neq Y) \geq \sum_{k=1}^{n} |p_k - q_k|$.

Exercice F

Soit r et $s \in [\![0, p-1]\!]$ distincts ; montrons que les variables Z_r et Z_s sont indépendantes. Considérons a et $b \in [\![0, p-1]\!]$.

▹ Calculons $\mathbf{P}(Z_r = a, Z_s = b)$. Le système

$$\begin{cases} x + ry = a[p] \\ x + sy = b[p] \end{cases}$$

admet une unique solution (x_0, y_0) dans $[\![0, p-1]\!]^2$ (car $(r-s)y = a-b$ admet une unique solution modulo p). Par conséquent, avec l'indépendance de X et Y,

$$\mathbf{P}(Z_r = a, Z_s = b) = \mathbf{P}(X = x_0, Y = y_0) = \mathbf{P}(X = x_0)\mathbf{P}(Y = y_0) = \frac{1}{p^2}.$$

▹ Calculons $\mathbf{P}(Z_r = a)$ en discutant selon la valeur prise par la variable Y

$$\mathbf{P}(Z_r = a) = \sum_{k=0}^{p-1} \mathbf{P}(X + rk = a[p], Y = k) = \sum_{k=0}^{p-1} \mathbf{P}(X + rk = a[p])\mathbf{P}(Y = k) = \sum_{k=0}^{p-1} \frac{1}{p^2} = \frac{1}{p}.$$

De même, $\mathbf{P}(Z_s = b) = \frac{1}{p}$.

En conclusion, pour tous a et $b \in [\![0, p-1]\!]$, $\mathbf{P}(Z_r = a, Z_s = b) = \mathbf{P}(Z_r = a)\mathbf{P}(Z_s = b)$ donc les variables Z_r et Z_s sont indépendantes.

Corrigés des problèmes du chapitre 25

Problème 1

1. La variable aléatoire X_n peut prendre toutes les valeurs entières de 0 à n.
2. a. Soit $k \in \{1, \dots, n\}$; alors, en remarquant que $(X_n = k) \subset (X_{n-1} = k-1)$,

$$\begin{aligned} \mathbf{P}(X_n = k) &= \mathbf{P}(X_n = k, X_{n-1} = k-1) \\ &= \mathbf{P}(X_n = k | X_{n-1} = k-1)\mathbf{P}(X_{n-1} = k-1) \\ &= \frac{k}{k+1}\mathbf{P}(X_{n-1} = k-1). \end{aligned}$$

b. Notons $a_{n,k} = (k+1)\mathbf{P}(X_n = k)$ pour tous $k \leq n$. D'après le point précédent, $a_{n,k} = a_{n-1,k-1}$ donc par une récurrence immédiate, $a_{n,k} = a_{n-k,0} = \mathbf{P}(X_{n-k} = 0) = u_{n-k}$. Ainsi,

$$\mathbf{P}(X_n = k) = \frac{1}{k+1} u_{n-k}.$$

c. Écrivons que la loi de X_n est une probabilité :

$$\begin{aligned} 1 &= \sum_{k=0}^{n} \mathbf{P}(X_n = k) \\ &= u_n + \sum_{k=1}^{n} \frac{1}{k+1} u_{n-k} = \sum_{k=0}^{n} \frac{1}{k+1} u_{n-k} \\ &= \sum_{j=0}^{n} \frac{1}{n+1-j} u_j \qquad \text{avec } j = n-k. \end{aligned}$$

CORRIGÉS

d. Remarquons tout d'abord que, par définition, $u_0 = 1$ et $u_1 = \frac{1}{2}$. Écrivons ensuite la relation précédente pour $n = 2$,

$$\frac{u_0}{3} + \frac{u_1}{2} + \frac{u_2}{1} = 1$$

donc $u_2 = \frac{5}{12}$. En recommençant de même avec $n = 3$, on obtient

$$\frac{u_0}{4} + \frac{u_1}{3} + \frac{u_2}{2} + \frac{u_3}{1} = 1$$

donc $u_3 = \frac{3}{8}$.

3. a. Partons de la définition d'espérance

$$\begin{aligned}
\mathbf{E}(X_n) &= \sum_{k=0}^{n} k\mathbf{P}(X_n = k) = \sum_{k=1}^{n} \frac{k}{k+1} u_{n-k} && \text{avec la question 2.}(a)\\
&= \sum_{k=1}^{n} u_{n-k} - \sum_{k=1}^{n} \frac{1}{k+1} u_{n-k} && \text{avec } \frac{k}{k+1} = 1 - \frac{1}{k+1}\\
&= \sum_{k=1}^{n} u_{n-k} - (1 - u_n) && \text{avec la question 2.}(b)\\
&= \sum_{j=0}^{n} u_j - 1.
\end{aligned}$$

b. D'après la question 2.(a),

$$\sum_{j=0}^{n-1} \frac{u_j}{n-j} = \sum_{j=0}^{n-1} \frac{u_j}{n-1+1-j} = 1.$$

c. Repartons de la question précédente,

$$\begin{aligned}
1 &= \sum_{j=0}^{n-1} \frac{u_j}{n+1-j} \frac{n+1-j}{n-j}\\
&= \sum_{j=0}^{n-1} \frac{u_j}{n+1-j} \left(1 + \frac{1}{n-j}\right)\\
&= 1 - u_n + \sum_{j=0}^{n-1} \frac{u_j}{(n+1-j)(n-j)}
\end{aligned}$$

d'où

$$u_n = \sum_{j=0}^{n-1} \frac{u_j}{(n-j)(n+1-j)}.$$

d. ▹ Montrons par récurrence que $(n+1)u_n \geq 1$ pour tout n.

- Le résultat est immédiat pour $n = 0$ puisque $u_0 = 1$.
- Soit $n \geq 1$ tel que $(j+1)u_j \geq 1$ pour tout $j < n$. Utilisons l'identité précédente pour obtenir le résultat

au rang n.

$$\begin{aligned}(n+1)u_n &= \frac{n+1}{n}\sum_{j=0}^{n-1}\frac{n\,u_j}{(n-j)(n+1-j)} \geq \frac{n+1}{n}\sum_{j=0}^{n-1}\frac{(j+1)u_j}{(n-j)(n+1-j)}\\ &\geq \frac{n+1}{n}\sum_{j=0}^{n-1}\frac{1}{(n-j)(n+1-j)} && \text{par hypothèse de récurrence}\\ &\geq \frac{n+1}{n}\sum_{j=0}^{n-1}\Big(\frac{1}{n-j}-\frac{1}{n+1-j}\Big)\\ &\geq \frac{n+1}{n}\Big(1-\frac{1}{n+1}\Big)=1.\end{aligned}$$

▹ Avec cette minoration et l'expression de l'espérance trouvée à la question 3.(a), on obtient

$$\mathbf{E}(X_n) \geq \sum_{j=1}^{n}\frac{1}{j+1}.$$

Par comparaison à la série harmonique, $\mathbf{E}(X_n) \to +\infty$.

Problème 2

1. L'événement A_k est l'intersection de A et de l'événement $(\min\{j,\ S_j \geq a\} = k)$. Comme les événements $(\min\{j,\ S_j \geq a\} = k)$ sont deux à deux disjoints, les événements A_k sont eux aussi deux à deux disjoints. Par ailleurs, l'événement A se réécrit « $\{j,\ S_j \geq a\}$ est non vide » ; la décomposition proposée correspond donc à la discussion selon la valeur du plus petit élément de cet ensemble.
2. La partition de la première question donne

$$\mathbb{1}_A = \sum_{k=1}^{n}\mathbb{1}_{A_k}.$$

Ainsi,

$$S_n^2 \geq S_n^2\mathbb{1}_A = \sum_{k=1}^{n}S_n^2\mathbb{1}_{A_k}.$$

On obtient l'inégalité souhaitée par croissance et linéarité de l'espérance.
3. Remarquons tout d'abord que

$$\begin{aligned}S_n^2 &= (S_n - S_k + S_k)^2\\ &= (S_n - S_k)^2 + 2S_k(S_n - S_k) + S_k^2\\ &\geq 2S_k(S_n - S_k) + S_k^2.\end{aligned}$$

Ainsi, par croissance de l'espérance,

$$\mathbf{E}(S_n^2\mathbb{1}_{A_k}) \geq 2\mathbf{E}(S_k(S_n - S_k)\mathbb{1}_{A_k}) + \mathbf{E}(S_k^2\mathbb{1}_{A_k}).$$

Remarquons alors que les variables aléatoires $S_n - S_k$ (somme des variables X_j pour $j > k$) et $S_k\mathbb{1}_{A_k}$ (qui ne dépend que des variables X_j pour $j \leq k$) sont indépendantes. Ainsi,

$$\mathbf{E}(S_k(S_n - S_k)\mathbb{1}_{A_k}) = \mathbf{E}(S_n - S_k)\mathbf{E}(S_k\mathbb{1}_{A_k}),$$

ce qui donne exactement l'inégalité souhaitée.

4. Remarquons que la variable $S_n - S_k$ est centrée donc l'inégalité ci-dessus devient $\mathbf{E}(S_n^2 \mathbb{1}_{A_k}) \geq \mathbf{E}(S_k^2 \mathbb{1}_{A_k})$. Reportons ceci dans le résultat de la deuxième question

$$\begin{aligned}\mathbf{E}(S_n^2) &\geq \sum_{k=1}^{n} \mathbf{E}(S_k^2 \mathbb{1}_{A_k}) \\ &\geq \sum_{k=1}^{n} \mathbf{E}(a^2 \mathbb{1}_{A_k}) = a^2 \sum_{k=1}^{n} \mathbf{P}(A_k) = a^2 \mathbf{P}(A).\end{aligned}$$

Pour conclure, utilisons que les variables $X_1, \ldots, X_n$ sont centrées (donc S_n est centrée) et indépendantes

$$\mathbf{E}(S_n^2) = \mathbf{V}(S_n) = \sum_{k=1}^{n} \mathbf{V}(X_k).$$

En conclusion, $a^2 \mathbf{P}(A) \leq \sum_{k=1}^{n} \mathbf{V}(X_k)$.

Problème 3

Partie A

1. Remarquons tout d'abord que $a_i \neq a_j$.
▹ Si $a_i < a_j$, il est possible de continuer une sous-suite croissante finissant en a_i, de sorte que $c_i < c_j$.
▹ Si $a_i > a_j$, il est possible de continuer une sous-suite décroissante finissant en a_i, de sorte que $d_i < d_j$.
Dans les deux cas, $(c_i, d_i) \neq (c_j, d_j)$.
2. Raisonnons par l'absurde en supposant que $c_i \leq r$ et $d_i \leq s$ pour tout $i \in [\![1, n]\!]$. Ainsi le nombre de couples (c_i, d_i) possibles est au plus $rs < n$. Or, d'après la question précédente, il y a au moins n tels couples distincts : contradiction.
3. Comme à la question précédente, il y a au plus $C(a)D(a)$ valeurs distinctes pour les couples (c_i, d_i); or, d'après la première question, il existe n couples (c_i, d_i) différents. En conclusion, $n \leq C(a)D(a)$.

Partie B

1. Les variables aléatoires $X_n' = (X_n(n), X_n(n-1), \ldots, X_n(1))$ et X_n suivent la même loi, et $C(X_n) = D(X_n')$. Ainsi,

$$\mathbf{E}(C(X_n)) = \mathbf{E}(D(X_n')) = \mathbf{E}(D(X_n)).$$

2. D'après la dernière question de la partie A, $C(X_n)D(X_n) \geq n$. En utilisant la question précédente, l'inégalité arithmético-géométrique et la croissance de l'espérance, on obtient

$$\mathbf{E}(C(X_n)) = \mathbf{E}\left(\frac{C(X_n) + D(X_n)}{2}\right) \geq \mathbf{E}\left(\sqrt{C(X_n)D(X_n)}\right) \geq \sqrt{n}.$$

3. Une permutation $\sigma \in S_n$ telle que $\sigma(i_1) < \sigma(i_2) < \ldots < \sigma(i_k)$ est caractérisée par
- les k entiers images de $i_1, i_2, \ldots, i_k$, soit $\binom{n}{k}$ choix,
- une permutation entre les $n-k$ éléments restants au départ et à l'arrivée, soit $(n-k)!$ choix.

Ainsi,

$$\mathbf{P}(X_n(i_1) < X_n(i_2) < \ldots < X_n(i_k)) = \frac{1}{n!}\binom{n}{k}(n-k)! = \frac{1}{k!}.$$

Remarque
Ce résultat est intuitif : il y a $k!$ façons d'ordonner les valeurs correspondant aux indices $i_1, i_2, \ldots, i_k$ et une seule les place dans l'ordre croissant.

4. Par linéarité de l'espérance, $\mathbf{E}(L_k^n)$ vaut

$$\sum_{1\le i_1<i_2<\ldots<i_k\le n} \mathbf{P}(X_n(i_1)<X_n(i_2)<\ldots<X_n(i_k)) = \sum_{1\le i_1<i_2<\ldots<i_k\le n} \frac{1}{k!} = \frac{\binom{n}{k}}{k!}.$$

5. Tout d'abord, si $\alpha>0$ est fixé, la formule de Stirling garantit l'existence d'une constante $C>0$ telle que

$$\frac{1}{\lfloor\alpha\sqrt{n}\rfloor!}\binom{n}{\lfloor\alpha\sqrt{n}\rfloor} = C\exp\bigl(2\alpha(1-\ln(\alpha))\sqrt{n}(1+o(1))\bigr).$$

Ainsi, pour $\alpha>e$, $\mathbf{E}(L^n_{\lfloor\alpha\sqrt{n}\rfloor})$ décroit exponentiellement vite. Il suffit alors de remarquer que

$$\mathbf{P}(C(X_n)\ge(e+\varepsilon)\sqrt{n}) = \mathbf{P}(L^n_{\lfloor(e+\varepsilon)\sqrt{n}\rfloor}\ge 1)\le \mathbf{E}(L^n_{\lfloor(e+\varepsilon)\sqrt{n}\rfloor}).$$

6. Fixons $\varepsilon>0$. L'idée est de majorer $C(X_n)$ par n lorsque l'événement $\{C(X_n)>(e+\varepsilon/2)\sqrt{n}\}$ est réalisé. Plus précisément, écrivons

$$\begin{aligned}\mathbf{E}(C(X_n)) &= \mathbf{E}(C(X_n)\mathbb{1}_{C(X_n)>(e+\varepsilon/2)\sqrt{n}}) + \mathbf{E}(C(X_n)\mathbb{1}_{C(X_n)\le(e+\varepsilon/2)\sqrt{n}})\\ &\le n\mathbf{P}(C(X_n)>(e+\varepsilon/2)\sqrt{n}) + (e+\varepsilon/2)\sqrt{n}\mathbf{P}(C(X_n)\le(e+\varepsilon/2)\sqrt{n})\\ &\le n\mathbf{P}(C(X_n)>(e+\varepsilon/2)\sqrt{n}) + (e+\varepsilon/2)\sqrt{n}.\end{aligned}$$

Ainsi,

$$\frac{\mathbf{E}(C(X_n))}{\sqrt{n}} \le \sqrt{n}\mathbf{P}\bigl(C(X_n)>(e+\varepsilon/2)\sqrt{n}\bigr) + e + \frac{\varepsilon}{2}.$$

On conclut en remarquant que, d'après la question précédente, le premier terme de ce majorant tend vers 0.

Index

affixe, 70
algorithme d'Euclide, 453
algorithme de Horner, 489
angle orienté, 795
anneau, 425
anneau intègre, 428
application, 21
application linéaire, 558
arccos, arcsin, arctan, 170
argument, 78
argument principal, 79
automorphisme, 559

base, 587
base directe, 735
base duale, 602
bijection, 23
bijection réelle, 156
borne inférieure, 32
borne supérieure, 32

cardinal, 908
changement de variables, 202, 835
classe $\mathscr{C}^n$, 154, 344
classe $\mathscr{C}^\infty$, 154, 344
coefficient binomial, 123, 913
cofacteur, 740
comatrice, 740
combinaison linéaire, 552
congruence, 448
continuité, 289
continuité uniforme, 305
corps, 66, 428
covariance, 977
critères de divisibilité, 449
croissances comparées, 165, 247, 385
cycle, 707

degré, 485, 525
dichotomie, 300
dimension, 592
diviseur de zéro, 427
division euclidienne, 450
division euclidienne dans $_K[X]$, 494
décimaux, 253
décomposition en cycles disjoints, 709
décomposition en éléments simples, 528
dérivabilité, 338
dérivée partielle, 155
déterminant, 733
déterminant de Vandermonde, 743
développement asymptotique, 395
développement limité, 385

écart-type, 975
écriture décimale, 252
élément neutre, 25
endomorphisme, 558
entiers de Gauss, 426
équation différentielle, 196
espace dual, 559
espace euclidien, 774
espace probabilisable, 936
espace probabilisé, 937
espace vectoriel, 550
espérance, 971
événement, 936
événement certain, 937
événement négligeable, 937

famille libre, 586
fonction en escalier, 820
fonction indicatrice, 29
fonction polynomiale, 488
fonctions symétriques élémentaires, 493
forme linéaire, 559
formule de Bayes, 945
formule de Grassmann, 596
formule de König-Huygens, 975
formule de la comatrice, 742
formule de Leibniz, 345, 488
formule de Moivre, 76
formule de Pascal, 123, 914
formule de Stirling, 248, 878
formule de Taylor, 491
formule de Taylor avec reste intégral, 838
formule de Taylor-Young, 387, 840
formule de transfert, 974
formule de Vandermonde, 915
formule des probabilités composées, 942
formule des probabilités totales, 943
formule du binôme de Newton, 125
formule du crible, 910
formule du rang, 598, 667
formules d'Euler, 75, 168
formules de Cramer, 745
fraction rationnelle, 524

groupe, 77, 412
groupe linéaire, 559
groupe orthogonal, 790
groupe symétrique, 706

homothétie, 567
hyperplan, 603

identité du parallèlogramme, 775

identités de polarisation, 775
identités remarquables, 775
image directe, 22
image réciproque, 22
indépendance, 946, 968
indépendance mutuelle, 947, 968
injection, 23
interpolation de Lagrange, 490
intégrale, 822
intégrales de Wallis, 202, 834
intégration par parties, 201, 834
inégalité de Bienaymé-Tchebychev, 976
inégalité de Cauchy-Schwarz, 824
inégalité de Markov, 973, 974
inégalité de Taylor-Lagrange, 839
inégalité des accroissements finis, 352
isomorphisme, 558
isométrie, 790
isométrie positive, 793

lemme d'encadrement, 250, 294
lemme d'Euclide, 457
Lemme de Gauss, 455
lemme des coalitions, 969
limite de fonction, 284
loi associative, 25
loi binomiale, 966
loi commutative, 25
loi d'une variable aléatoire, 965
loi de Bernoulli, 966
loi de composition interne, 25
loi distributive, 26
loi jointe, 967
loi marginale, 967
loi uniforme, 965
lois de De Morgan, 30

matrice, 630
matrice antisymétrique, 636
matrice de passage, 664
matrice nilpotente, 645
matrice orthogonale, 792
matrice par blocs, 631
matrice symétrique, 636
matrice triangulaire, 635
matrices semblables, 680
matrices échelonnées, 671
matrices équivalentes, 668
multiplicité d'une racine, 491

nombre algébrique, 757
noyau, 561
numération en base b, 451

opérations élémentaires, 669, 739
orbite, 706
orthogonal d'un sous-espace, 785

partie dense, 256
partie entière, 251, 527
partition, 27
permutation, 706
permutation circulaire, 707
PGCD, 452, 495
pivot de Gauss, 672
plus grand/petit élément, 32
polynôme, 484
polynôme irréductible, 498
polynôme scindé, 348, 492
polynômes de Tchebychev, 127
PPCM, 452, 495
premiers, 456
premiers entre eux, 454
primitive, 199, 830
principe de superposition, 198
probabilité, 937
probabilité conditionnelle, 941
problème de Cauchy, 198
procédé de Gram-Schmidt, 782
produit cartésien, 20
produit de Cauchy, 235
produit scalaire, 774
projecteur, 564
puissances, 163

quantificateurs, 15

racine, 489
racines de l'unité, 86
rang, 596, 597, 746
relation d'ordre, 30
relation d'équivalence, 28
relation de Chasles, 825
relations coefficients-racines, 493
relations de comparaison, 244, 383
rotation, 795
règle de Sarrus, 733
récurrence, 34
réflexion, 792

signature, 712
somme de Riemann, 826
somme de sous-espaces, 556
somme directe, 556
somme géométrique, 129
sous-anneau, 428
sous-espace affine, 562
sous-espace vectoriel, 551
sous-espace vectoriel engendré, 553
sous-groupe, 416
sous-groupes de $\mathbb{R}$, 257
sous-groupes de $\mathbb{Z}$, 450
suite convergente, 236
suite extraite, 257
suite stationnaire, 236
suites adjacentes, 254
support, 706
surjection, 23

symbole de Kronecker, 485
symétrie, 566
système linéaire, 676, 677
série absolument convergente, 868
série convergente, 868
série de Riemann, 874
série géométrique, 870
série harmonique, 870
série numérique, 868

théorème de Bolzano-Weierstrass, 262
théorème de Bézout, 455
théorème de D'Alembert-Gauss, 88, 493
théorème de Fermat, 457
théorème de Heine, 306
théorème de la base extraite, 591
théorème de la base incomplète, 591
théorème de la bijection, 156, 343
théorème de limite monotone, 254
théorème de Rolle, 348
théorème des accroissements finis, 349
théorème des bornes, 303
théorème des valeurs intermédiaires, 299
théorème fondamental de l'arithmétique, 458
trace, 641
transposition, 707
transposée, 635

univers, 936

variable aléatoire, 964
variance, 975
variation de la constante, 206
vecteurs orthogonaux, 780

Crédits iconographiques
p. 13. School of Mathematics and Statistics, University of St Andrews, Scotland (Wikimedia Commons) – p. 65. bibmaths.fr – p. 109. Wikimedia Commons – p. 149. Wikimedia Commons – p. 195. alchetron.com – p. 229. bibmaths.fr – p. 283. Wikimedia Commons – p. 337. redsearch.org – p. 381. bibmaths.fr – p. 411. bibmaths.fr – p. 447. popsciences.universite.lyon.fr – p. 483. britannica.com – p. 523. Wikimedia Commons – p. 549. newworldencyclopedia.org – p. 585. redresearch.org – p. 629. Wikimedia Commons – p. 663. apprendre-maths.info – p. 705. sspiprint.com – p. 729. capture d'une vidéo youtube.com – Inversibilité d'une matrice de Vandermonde – p. 773. prun.net – p. 819. Wikimedia Commons – p. 867. Wikimedia Commons – p. 905. Wikimedia Commons – p. 935. britannica.com – p. 963. bibmaths.fr